Taylor & Francis
Taylor & Francis Group

Research Article

Influence of different species range types on the perception of macroecological patterns

CLAUDIA RAEDIG[1,2] & HOLGER KREFT[3,4]

[1]Institute for Technology and Resources Management in the Tropics and Subtropics, Cologne University of Applied Sciences, Betzdorfer Straße 2, 50679 Cologne, Germany
[2]Department of Systematic Botany, Institute of Biology I, University of Leipzig, Johannisallee 21, 04103 Leipzig, Germany
[3]Biodiversity, Macroecology & Conservation Biogeography Group, Georg-August University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany
[4]Nees Institute for the Biodiversity of Plants, University of Bonn, Meckenheimer Allee 170, 53115 Bonn, Germany

In the face of increasing availability and use of distribution data, large-scale approaches of mapping species distribution patterns have become a central component of development of large-scale conservation policies. Particularly in tropical regions and for non-vertebrate taxa, knowledge on distribution patterns at large spatial extents remains woefully limited. Datasets are often geographically and taxonomically incomplete, have presence-only character and lack abundance information. One intermediate step for the application of such data common to most approaches is the construction of species geographic ranges. In this study, we investigated the effects of different methods for constructing species ranges on range attributes and species richness. We selected the Neotropical palm genus *Bactris* as a typical example for a diverse tropical plant taxon that is comparatively well researched, but under collected. For 48 *Bactris* species, we compared point-to-grid ranges, expert ranges, convex polygons and modelled ranges. Range attributes and resulting species richness patterns differed tremendously and were constrained by incompleteness of the respective data type and by the errors associated to the method itself. The consequences of applying different methods to construct species ranges highlighted here can aid in selecting appropriate methods for analysing distribution data at large geographic scales.

Key words: *Bactris*, data scarcity, large-scale distribution data, MAXENT, plants, spatial patterns, species geographic ranges, species richness, tropics

Introduction

Geographic ranges of species represent the basis of most macroecological and biogeographical studies (e.g. Jetz & Rahbek, 2002; Graham & Hijmans, 2006; Grenyer *et al.*, 2006; Kreft *et al.*, 2006; Jetz *et al*., 2008; Raedig *et al.*, 2010) as well as a fundamental criterion to assess the conservational importance of species and regions (Willis *et al.*, 2003; Orme *et al.*, 2005; IUCN Standards and Petitions Subcommittee, 2010).

Recent studies indicate that different range attributes are of crucial importance towards our perception of richness patterns as well as of their environmental correlates and causal drivers (e.g. Lennon *et al.*, 2004; Luoto *et al.*, 2005; Kreft *et al.*, 2006; McPherson & Jetz, 2007). Basically, a geographic range can be defined as the area wherein

occurrences of individuals of a species are located (Rapoport, 1982; Gaston, 1991). Whereas the extent of occurrence (EOO) describes the area between the outmost borders of a species range, the area of occupancy (AOO) describes the area where a species is actually present (Gaston, 1991, 2003). These two parameters are used by the IUCN to derive biologically meaningful properties of geographic ranges (Gaston & Fuller, 2009; IUCN Standards and Petitions Subcommittee, 2010). In practice, species distributions at large spatial extents and at coarse grain are usually depicted using EOO (e.g. Buckley & Jetz, 2007; Hurlbert & Jetz, 2007).

Apart from this distinction, species ranges used to derive distribution maps rely on different data sources and different methods to construct species ranges from primary occurrence data. Generally, species ranges are in one way or another related to field collections, observations, or museum specimens. At large spatial extents, these data are

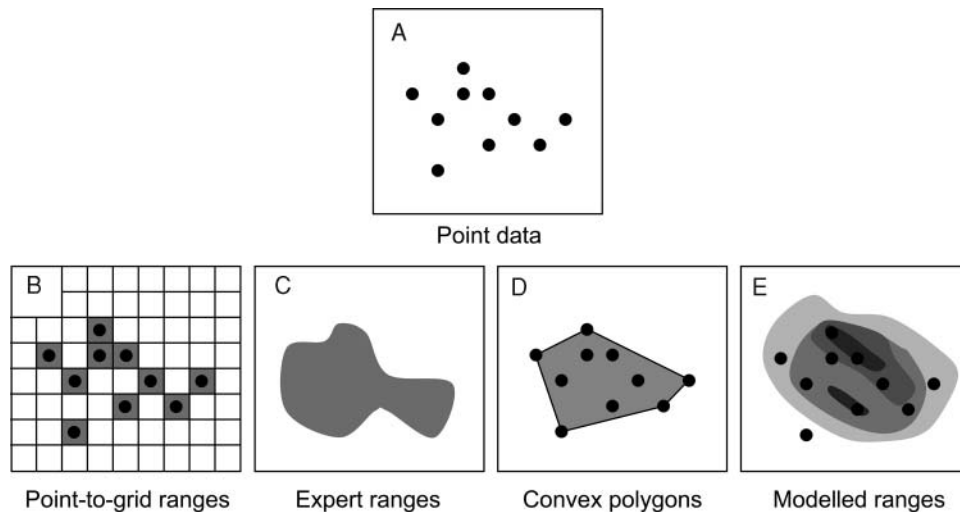Correspondence to: Claudia Raedig. E-mail: claudia.raedig@fh-koeln.de

Fig. 1. Schematic illustration of frequently used methods to depict the geographic range of a species. (A) Point data form the basis of the different types of species ranges. (B) Point-to-grid ranges are created by binning point data into a gridded mapping scheme. (C) Expert ranges are deduced by experts based on knowledge of specimens, field experience and other non-quantifiable information. (D) Convex polygons represent the minimal convex set including all underlying point data. (E) Modelled ranges are based on statistical relationships between observed occurrences and climatic and other environmental variables.

mostly presence-only data, lacking the information whether a species is truly absent from areas where it has not been recorded (e.g. Graham *et al*., 2004; Elith *et al*., 2006). Such georeferenced data can be used to generate dot maps (Fig. 1a). Often, point data are transformed into gridded mapping templates yielding point-to-grid maps (Fig. 1b). A second major data source originates from distribution atlases (Fig. 1c). Such expert-drawn range maps are usually constructed by complementing occurrence data with extensive, but not quantified, expert knowledge on species and their environmental requirements (Graham & Hijmans, 2006; Hurlbert & Jetz, 2007). This latter data source has been extensively exploited to analyse large-scale richness patterns of taxa – such as terrestrial vertebrates – which are generally well-known and mapped at a global scale (e.g. Orme *et al.*, 2005; Grenyer *et al.*, 2006; Buckley & Jetz, 2007, Schipper *et al.*, 2008). For non-vertebrate taxa, no comparable studies based on expert ranges exist at the global scale.

Furthermore, two main approaches exist to construct species ranges. First, methods for range construction as utilised by the IUCN to estimate EOO (IUCN Standards and Petitions Subcommittee, 2010) interpolate species ranges using convex polygons which flank the outermost known edges of documented occurrences (Willis *et al.*, 2003; Fig. 1d). This approach is based on geometric rules and generally does not consider ecological variables that determine the distribution of species or phenomena such as disjunct distributions. Secondly, species ranges can be produced by extrapolating over the limits of documented occurrences applying different extrapolation constraints. Extrapolation can be performed using criteria like the distance

to neighbouring occurrences (Hopkins, 2007), a standard diameter as buffer width (Schulman *et al*., 2007), or the average distance of a minimum spanning tree between occurrences as buffer width (Rapoport, 1982; Willis *et al.*, 2003; Hernández & Navarro, 2007; Moat, 2007). Another extrapolation approach is species distribution modelling which has attracted massive scientific interest in the past decade (e.g. Guisan & Zimmermann, 2000; Thuiller *et al*., 2003; Segurado & Araújo, 2004; Guisan & Thuiller, 2005; Araújo & Guisan, 2006; Elith *et al.*, 2006; Araújo & New, 2007; Botkin *et al*., 2007; Tsoar *et al.*, 2007; Cayuela *et al.*, 2009). Statistical relationships between species occurrences and environmental predictor variables are used to model suitable areas where a species is likely to occur (Fig. 1e).

With the increasing availability and use of species-level distribution data for basic and applied issues in ecology, biogeography and conservation, it becomes necessary to scrutinise uncertainties arising from different methods to construct species ranges. A central prerequisite for identifying these uncertainties is the consideration of the basic problems inherent to all large-scale distribution datasets: incompleteness and data quality. In particular for tropical regions and speciose, cryptic or otherwise inconspicuous taxa, distribution datasets often are incomplete (Prance *et al.*, 2000; Crisp *et al.*, 2001; Linder, 2001; Graham *et al.*, 2004; Schmidt *et al.*, 2005; Cayuela *et al.*, 2009). For instance, continental-scale analyses of plant richness were based on, e.g. *c.* 5% of Neotropical (Morawetz & Raedig, 2007) and *c.* 10–15% of African angiosperm species (Küper *et al.*, 2004). And even in regions with a longstanding floristic tradition like Europe, only about 20% of the flora has

yet been mapped (Jalas & Suominen, 1972–1994; Jalas *et al.*, 1996, 1999; Kurtto *et al.*, 2004, 2007). This difference in incompleteness is reflected by the number of tree specimens contained in the Missouri Botanical Garden VAST (VAScular Tropicos) database and others: in Great Britain and Ireland each tree species is documented with on average 4311 records, and in the Netherlands with 6687 respectively; for selected tropical American countries with up to 16-fold higher tree species richness only between six (Guatemala) and 21 (Nicaragua) documented records per tree species exist on average (Cayuela *et al.*, 2009).

Probably the most critical issue contributing to incompleteness of large-scale distribution datasets is geographically heterogeneous sampling effort (Nelson *et al.*, 1990; Gaston, 2003; Graham & Hijmans, 2006; Morawetz & Raedig, 2007; Raedig *et al.*, 2010). In general, sampling effort decreases with increasing difficulty of sampling, thus collecting activity in remote regions is low. Hence, Beck *et al.* (2007) found that the mean number of locality records per species was low in general, and lowest for amphibians and vascular plants (six and seven records per species respectively) along the eastern flank of the Andes in Peru and Bolivia. Schulman *et al.* (2007) showed that sampling effort in Amazonia is highly heterogeneously distributed and classified 43% of Amazonia as botanically not explored, a further 28% as poorly and only 2% as comparably well-collected (at 0.5° grid resolution).

In addition to the spatial incompleteness, the quality of data in terms of taxonomic correctness varies strongly between distribution datasets. In a growing number of studies partly unrevised data sources are exploited. Verifying unrevised botanical data, Hopkins (2007) found error rates of up to more than 40% for Amazonian plants. Unfortunately, revised distribution data covering large areas are hardly available. The few available large-scale datasets share a common characteristic which makes their statistical analysis difficult: the number of occurrences for the species are often low and datasets contain high portions of singletons (e.g. Condit *et al.*, 2002, Panama: 22%; Beck *et al.*, 2007, Bolivian Andes: 28%; Morawetz & Raedig, 2007, Neotropics: 21%). In distribution modelling, species with few records are typically excluded due to statistical constraints. This exclusion will aggravate the effect of species richness patterns tending to be dominated by widespread species (Jetz & Rahbek, 2002; Lennon *et al.*, 2004; Kreft *et al.*, 2006). However, the excluded species probably have the smallest ranges and are therefore of greatest conservation concern (Graham & Hijmans, 2006; Cayuela *et al.*, 2009). Thus, for presence-only datasets, a high proportion of species with few occurrences (e.g. Morawetz & Raedig, 2007: 42.5% of species with three or less occurrences) advises against distribution modelling. If the scale of analysis is fine-grained and sufficient data are available, ranges have been successfully modelled (e.g. Thuiller *et al.*, 2003; Graham & Hijmans, 2006; Hernandez *et al.*, 2006), and are

useful for conservation purposes such as new population discovery (Raxworthy *et al.*, 2003; Williams *et al.*, 2009).

The construction of species ranges is hampered by both incompleteness of distribution data and varying data quality. Against this background, understanding the effects of the different methods to construct species ranges is relevant for biodiversity mapping, macroecology, climate change research and conservation. Effects of different data types have been assessed for species richness at large geographic extents deduced from survey data and museum specimens (Guralnick & Van Cleve, 2005) as well as expert ranges and survey data (Hurlbert & White, 2005; Hurlbert & Jetz, 2007; McPherson & Jetz, 2007). These studies were restricted to well-known taxa and regions and already provide evidence for the relevance of the underlying data for scale issues of species richness, or for the deduction of drivers of species richness, but did not explore species ranges as such. At the regional scale, different methods for constructing species ranges have been explored (Graham & Hijmans, 2006), yet an examination of species range construction for the detection of large-scale distribution patterns is lacking. Furthermore, the influence of data types and species range construction remains unknown for diverse groups such as plants, for which distribution data are disproportionately scarce, and for humid tropical regions which are particularly diverse but under-collected (Prance *et al.*, 2000; Cayuela *et al.*, 2009).

The goal of this study was to analyse the effects of different methods to construct species ranges on large-scale spatial patterns of species richness for a diverse tropical plant taxon. We thus selected the palm genus *Bactris* which represents one of the few cases where both revised point-to-grid ranges and expert ranges are available for a tropical plant genus (Henderson *et al.*, 1995; Henderson, 2000). We chose common methods used to construct species ranges from these two sources and constructed point-to-grid ranges, expert ranges, convex polygons and modelled ranges. We then compared resulting species ranges and aggregated maps of species richness (e.g. Graham & Hijmans, 2006; Diniz-Filho *et al.*, 2008; Gove *et al.*, 2008). More specifically, we addressed the following questions: (1) What are the differences between the constructed species ranges in terms of frequency distribution and further range attributes? (2) What are the differences between the derived maps of species richness regarding spatial structure, spatial congruence, in particular of centres of species richness, and across different spatial scales? (3) How can these differences be explained and accounted for?

Methods

Construction of species ranges

Point-to-grid maps based on revised specimens of 73 species (87 species and subspecific taxa) of *Bactris* are

provided in Henderson (2000) at a $1° \times 1°$ resolution. Expert ranges for 64 species were extracted from the *Field Guide to the Palms of the Americas* (Henderson *et al.*, 1995). In this work, ranges of 550 Neotropical palms are depicted as expert range maps. In order to compare datasets, taxonomies used in Henderson (2000) and Henderson *et al.* (1995) had to be adjusted (64 species, Appendix S1, see supplementary material which is available on the Supplementary Content tab of the article's online page at http://dx.doi.org/10.1080/14772000.2011.588726). Polygons of expert ranges were first digitised and then overlaid with the $1° \times 1°$ grid in order to extract occurrences per grid cell. The centres of the grid cells from point-to-grid maps were interpolated to convex hulls using Hawth's Analysis Tools extension for ArcGIS, version 2.10. The resulting polygons built our convex polygon dataset and for reasons of comparability were recalculated to the same $1°$ grid cells.

We used MAXENT (Maximum Entropy Modeling) to model ranges based on point-to-grid occurrences and environmental data (Phillips *et al.*, 2006). MAXENT is a frequently used software and has been shown to be capable of modelling presence-only data when dealing with scarce occurrences (Hernandez *et al.*, 2006; Pearson *et al.*, 2007; Wisz *et al.*, 2008; Williams *et al.*, 2009). Using randomly assigned pixels from a background area, MAXENT distinguishes between the two instances 'presence' and 'random' which is comparable to using pseudo-absences instead of absences (Phillips *et al.*, 2006). For detailed information on our application of MAXENT (see Appendix S3, available online, as above).

Environmental predictors were derived from the World-Clim Global Climate GIS database (20 predictors; Hijmans, 2005). Additionally, two categorical data layers, ecoregions (Olson *et al.*, 2001) and vegetation zones (Bletter *et al.*, 2004), were included. To avoid over fitting (Guisan & Thuiller, 2005), the number of occurrences should be at least ten times the number of predictors (Harrell, 2001; Gibson *et al.*, 2007), and multicollinearity among predictors should be avoided (Graham, 2003; Guisan & Thuiller, 2005; Dormann *et al.*, 2007). Therefore, we excluded the most-correlated predictors based on a Spearman rank correlation matrix to minimise both collinearity and number of predictors (see Appendix S2, which is available online, as above).

Predictors that finally entered the modelling were altitude, maximum temperature of the warmest month, precipitation seasonality, precipitation of the warmest quarter, ecoregion and vegetation zone. Robust models could be generated for 48 species (see Appendix S3).

According to the general purpose of our study, we decided to use a fixed threshold to determine area predicted as suitable for a species ($P > 0.5$ for all *Bactris* species).

Comparing datasets

Range size was measured as the sum of occupied grid cells. Since coastal grid cells can affect the calculation of range size, the area of these cells was adjusted based on the fractional land surface in five classes from 0.2 to 1 degree. Range size frequency distributions (RSFDs) as well as mean, median and maximum range size were compared for all range types and tested for similarity (Mann–Whitney Rank Sum Test).

Maps of species richness were generated for each dataset by summing up incidences per $1°$ grid cell. Species richness generally shows a strong pattern of spatial autocorrelation, i.e. values of nearby grid cells are more similar values than distant ones due to the spatially and temporally induced causal interaction among nearby locations (Fortin *et al.*, 2002). The inherent spatial autocorrelation of richness maps was assessed with Moran's I correlograms. Moran's I values generally vary between -1 and 1, with values greater than zero indicating positive autocorrelation and values close to zero indicating absence of spatial autocorrelation. In order to quantify differences between resulting species richness patterns, pairwise values of species richness were plotted against each other to determine their degree of similarity (Pearson's r). Moreover, to investigate how patterns of similarity change with spatial scale, we determined the correlation between richness pairs at increasingly coarser grain sizes (up to $5°$). We further examined the spatial congruence of richness centres. First, richness centres were defined as cut-off percentages of species richness for very small (1%) to fairly large (25%) richness centre size. Then, the spatial overlap of richness centres at a certain cut-off level was determined as the proportional overlap between two richness centres.

Results

Range sizes of the *Bactris* species studied vary from small-ranging species such as *B. glassmannii* and *B. horridispatha* with four point occurrences to wide-ranging species with more than 100 point occurrences (*B. maraja*, *B. simplicifrons* (see Appendix S3, available online). Range sizes calculated for the four different range types differed considerably (Table 1) but generally followed a strongly right-skewed distribution which reached or approximated normality after $\log_{10}$-transformation (Fig. 2). The median range size was smallest for point-to-grid data (15.9), and the RSFD of this data type differed significantly from all other RSFDs ($P < 0.001$). The median range size was about four to seven times larger for expert ranges (66.3), modelled ranges (111.6) and convex polygons (70.0). RSFDs of the latter three were statistically indistinguishable from each other ($P = 0.079$–0.714). However, the maximum range size of expert ranges (918.8) was larger than both modelled ranges (437.2) and convex polygons (770.6).

Table 1. Statistical descriptors of geographic range size, measured as the number of grid cells, for four different range types.

Range type	Median	Mean	SD	Min	Max	Skewness	Kurtosis
Point-to-grid range	15.9	23.7	26.6	3.2	137.2	2.6	7.5
Expert range	66.3	150.7	211.2	4.0	918.8	2.1	3.8
Convex polygons	70.0	180.2	222.0	5.4	770.6	1.5	1.2
Modelled range	111.6	143.2	114.7	11.6	437.2	1.2	0.7

[*]SD, standard deviation; Min, minimum; Max, maximum.

The much smaller size of point-to-grid ranges was also reflected in overall lower species numbers per grid cell and in a high porosity of the species richness map (Fig. 3a). The location of maximum richness differed for each richness map (Fig. 3). Maximum point-to-grid richness (n = 15 species) was located in the Peruvian Amazonia near Iquitos ($-73.5°$ W, $-3.5°$ S) (Fig. 3a). Grid cells with highest species richness (n = 18) for expert ranges were located in the eastern Guyanas and north of Manaus in Brazil (Fig. 3b). These cells were embedded in a coherent belt of high species richness extending from French Guiana and Suriname southward to the Amazon, and along the Amazon up to the Iquitos region and southern Colombia. Maximum convex polygon richness (n = 21) was located in two grid cells at the Japurá river in western Brazil (66.5° W, 1.5° S; 67.5° W, 1.5° S) (Fig. 3c). Aside of these grid cells, a larger centre of convex polygon richness with grid cells containing 20 species was located in the border region of Colombia, Peru and Brazil. Both centres were part of a belt with continuously high species richness stretching from the Guyanas to the Amazon and almost covering the entire western Amazonian region. The maximum grid cell count of modelled *Bactris* ranges (n = 23) was located in northern Guyana (60.5° W, 6.5° N; Fig. 3d). Here, the entire Guyanas constituted a contiguous richness centre; further richness centres with more than 20 species lay along the Amazon. In parallel to the different location of richness centres, the prominent distribution gap of *Bactris* in the Cerrado and the Caatinga, two savannah formations in Brazil, was more or less detectable for point-to-grid, expert and modelled richness, and entirely lacking for convex polygon richness.

Species richness patterns resulting from all four range types showed a spatial structure (Fig. 3e). Positive spatial autocorrelation occurred up to approximately 2000 km in all four range types indicating that neighbouring grid cells had more similar values of species richness. In concordance with the high porosity illustrated in Fig. 3a, point-to-grid richness exhibited the most distinct pattern and by far the lowest level of spatial autocorrelation (Fig. 3e). The spatial structure of species richness was more similar for modelled ranges, expert ranges and convex polygons.

Grid cell-wise plotting of the species richness demonstrated the degree of congruence between different range types (Fig. 4). Whereas point-to-grid data showed low correlations with all other richness patterns (Pearson *r*-values 0.38 to 0.49, Fig. 4a–c), the remaining richness patterns were more similar. Expert and convex polygon richness (Fig. 4d) showed the highest similarity ($r = 0.90$). Decreasing the grid resolution to 2° yielded higher correlations of richness patterns (*r*-values 0.64 to 0.92, Fig. 5a). From 3° to 5° grid resolution, correlation of richness patterns became consistently more similar to each other with *r*-values ranging from 0.81 to 0.95.

Further examination of the spatial congruence of richness centres confirmed these patterns (Fig. 5b). Point-to-grid data showed the lowest spatial congruence of richness centres compared with all other pairs. For instance, when comparing the 10% most species-rich grid cells, the spatial congruence between richness centres appeared to be relatively low and varied between *c.* 20% and 40% (Fig. 5b, dotted lines). With a less strict definition of richness centres, congruence of pairs increased. Consequently, for 25% of most species-rich grid cells, spatial congruence ranged between 34.9% and 78.3%. Again, the highest similarity was observed between expert range and convex polygon richness.

Discussion

We selected the well-researched but under-collected palm genus *Bactris* to examine uncertainties arising from different range type data and range construction methods. Although palms and particularly Amazonian palms represent a traditionally under-collected group (Balick *et al.*, 1982), *Bactris* was chosen because the distribution data have been revised recently and the collection frequency is probably higher than for other palm genera.

Thus, our analyses were based on two high-quality, recently revised distribution datasets but of an unknown degree of incompleteness. This incompleteness is mainly due to heterogeneous spatial sampling effort: whereas in locations that are comparatively well-collected (e.g. Iquitos, Tefe, Manaus), specimens of *Bactris* have been sampled, for areas in-between, hardly any specimens exist in the herbaria (Henderson, 2000). The degree of incompleteness in distribution data is further aggravated by the low overall number of verified geographic records per species (mean = 25). Robust ranges could only be modelled for 48 *Bactris*
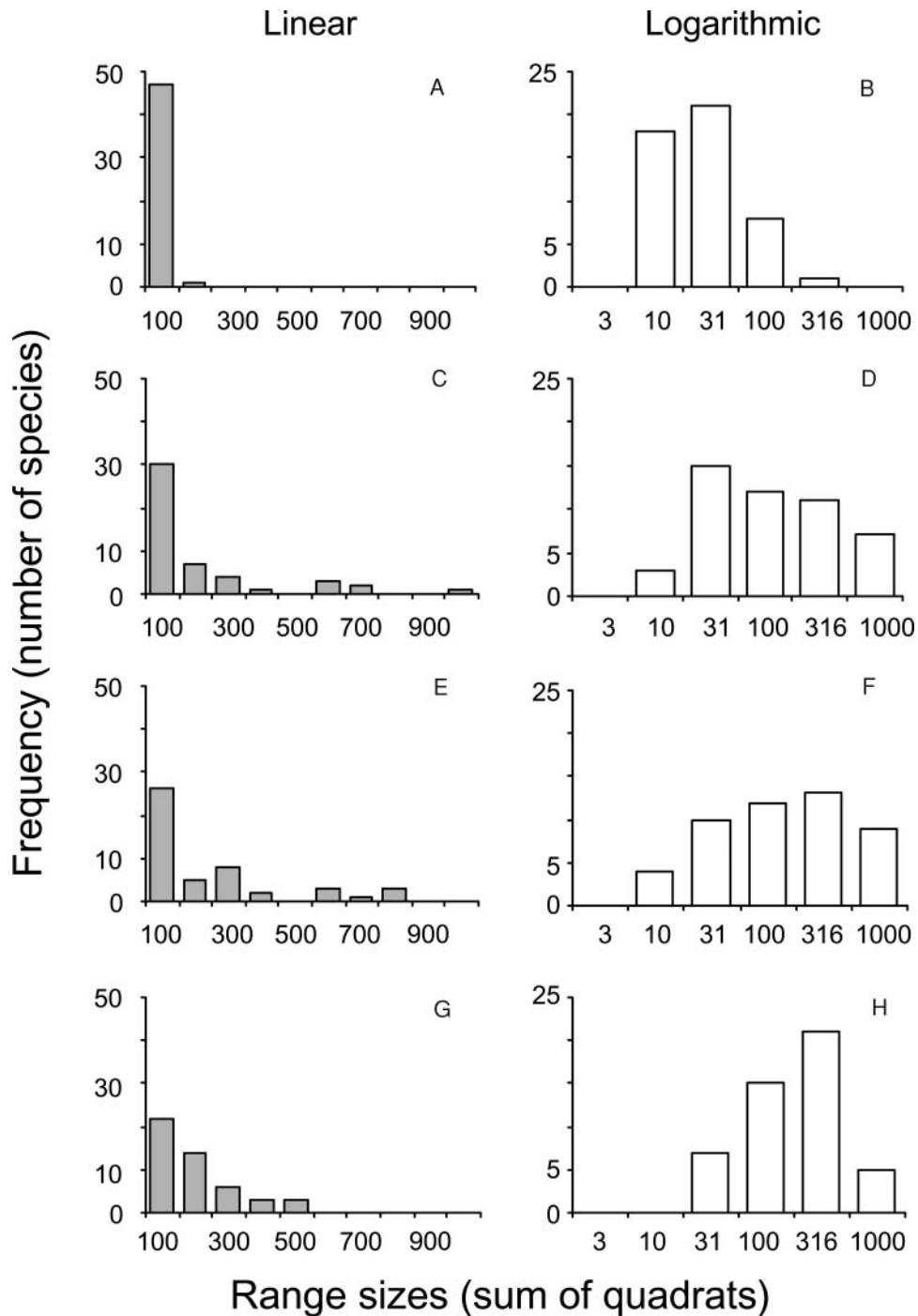
Fig. 2. Range size frequency distributions (RSFDs) on linear and logarithmic scale for four different range types. (A–B) Point-to-grid data, (C–D) expert range data, (E–F) convex polygon data and (G–H) modelled range data.

species. This reduced the present analysis to 75% of the total number of species in the genus useful for analysis.

The unknown degree of incompleteness of the datasets has further implications. Whereas it would be interesting to use the datasets to derive biologically relevant characteristics of species ranges such as the robustness of the

species distribution to threatening processes (Gaston & Fuller, 2009), the AOO and EOO parameters are difficult to distinguish in large-scale datasets. 'Large-scale' or 'coarse grained' refers to grid resolutions which are underlying analyses of distribution pattern at global or continental scale, usually at 0.5° grid resolution (with a grid
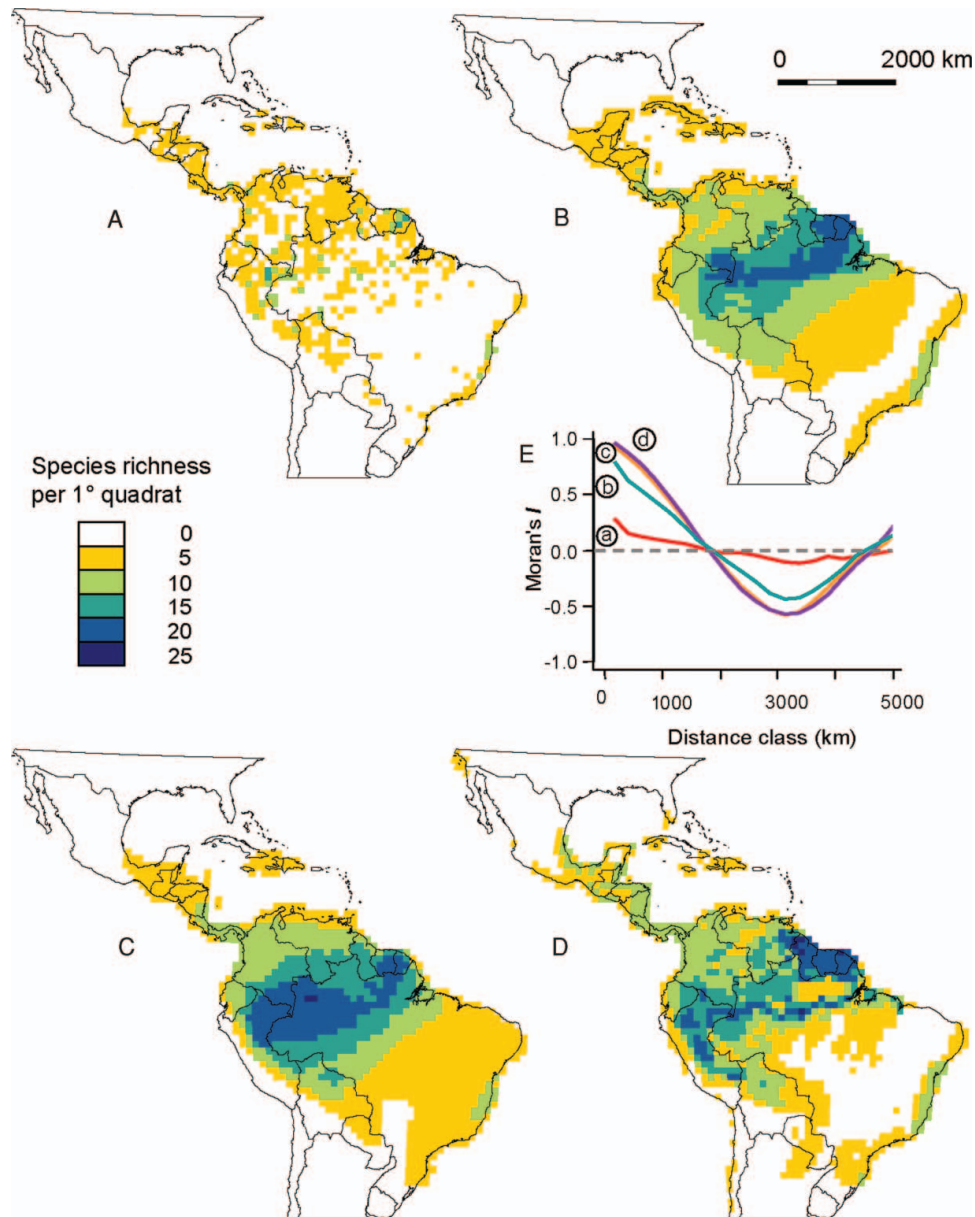
Fig. 3. Species richness and spatial autocorrelation patterns of *Bactris* at 1° × 1° resolution for four different range types. (A) Point-to-grid richness, (B) expert richness, (C) convex polygon richness and (D) modelled richness. Projection: Aitoff, Central Meridian 60° W. (e) Moran's *I* correlograms illustrate patterns of spatial autocorrelation for the richness maps in (A–D).

cell size of *c.* 2500 km^2) and above. The recommendation of the IUCN for an appropriate spatial scale for the usage of AOO is given by a 4 km^2 grid cell size (IUCN Standards and Petitions Subcommittee, 2010). However, as grain size gets larger, AOO and EOO converge. A given 10 000 km^2 grid cell cannot be considered as actual AOO when only one or few point occurrences exist for that grid cell or when point occurrences are concentrated in a small area in the grid cell. Since such detailed information is not available for the majority of large-scale datasets, the us-

age of the AOO parameter is not recommended at large scales.

Comparing the four sets of species ranges and species richness, we showed that the method chosen to construct species ranges has a great impact on identifying and interpreting large-scale patterns of plant species richness. This finding is in agreement with Graham & Hijmans (2006) who showed diverging richness patterns at a regional scale for Californian amphibians and reptiles subject to the underlying range construction. McPherson & Jetz (2007) in
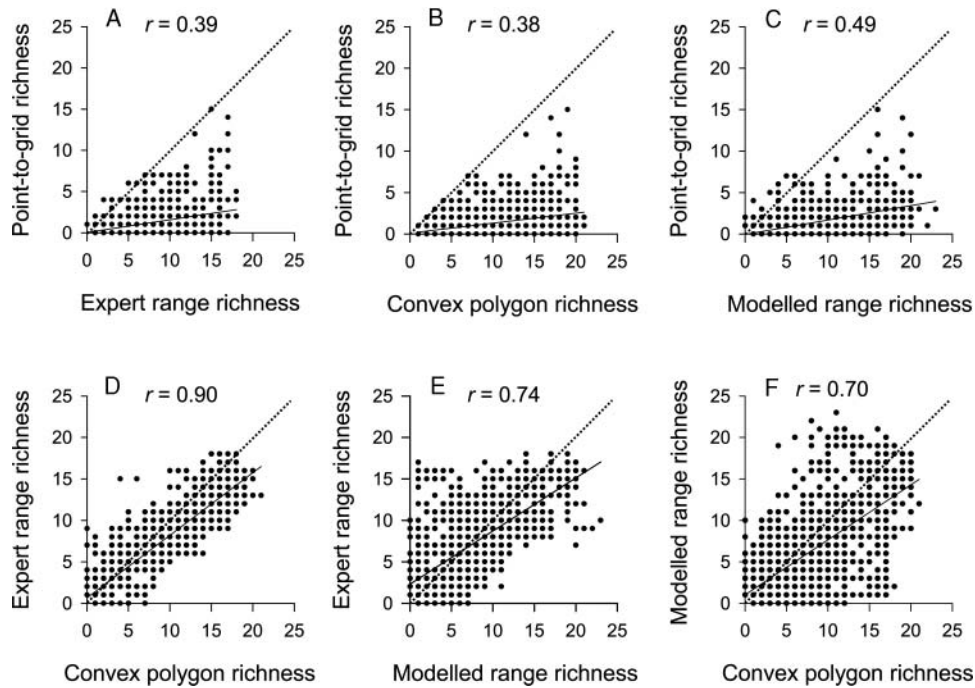
Fig. 4. Pair-wise comparisons of the four richness variables per 1° grid cell and correlation coefficients (Pearson's r) of the plotted relationships. Dotted lines are the equity lines indicating a 1:1 match between two range types. The solid lines indicate linear fits.

contrast found more similar richness patterns when comparing richness patterns for South African birds derived from expert range maps, survey data and species distribution models.

Species ranges

In our point-to-grid dataset, most species ranges were of small and medium size (Fig. 2a–b). While point-to-grid ranges imply a high risk for error of omission, meaning
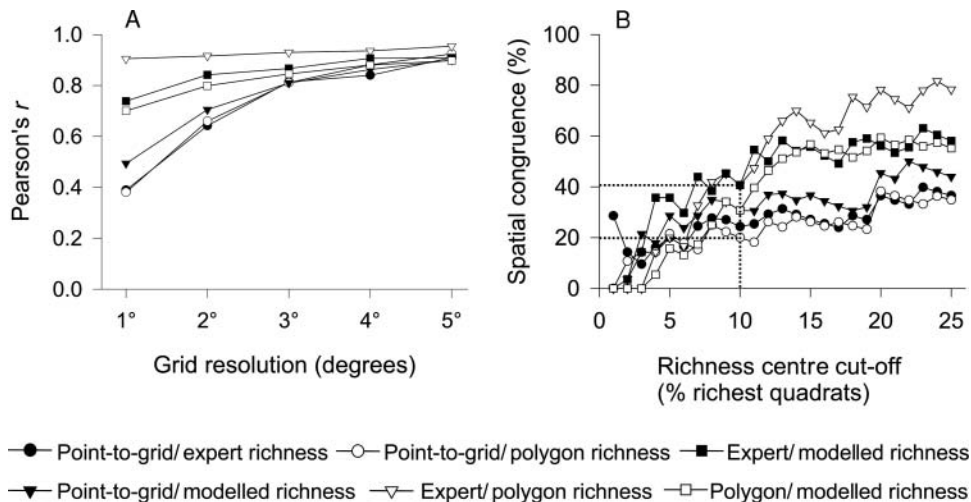

Fig. 5. Pair-wise comparison of richness variables at increasing spatial resolution and for varying definitions of 'richness centres'. (A) Correlation between richness variables at grain sizes from 1° to 5°. Convex polygons are abbreviated as 'polygons'. Total number of grid cells are: 2117 (1°), 608 (2°), 295 (3°), 178 (4°) and 118 (5°). (B) Spatial congruence is shown as proportion of similar grid cells from the total number of grid cells per richness cut-off. Definitions of richness centres are given in 1% steps. The dotted line indicates the spatial congruence of pairs of richness for the upper 10% of most species-rich grid cells, with 10% representing a threshold that is commonly employed to identify richness centres.

that absences of the species are falsely indicated (Fielding & Bell, 1997; Graham & Hijmans, 2006; Rondinini *et al.*, 2006), they minimise the risk of estimating false presences, i.e. errors of commission. In contrast, construction of species ranges using the three other methods implies a higher error of commission. For conservation planning, these errors are considered to be more profound because they potentially lead to the protection of areas where an endangered species is thought to be present but is not (Fielding & Bell, 1997; Rondinini *et al.*, 2006). However, due to omission error, the high point-to-grid richness found in some well-collected grid cells is likely to be similarly high in adjacent but poorly collected grid cells. Consequently, point-to-grid ranges are helpful in directing attention towards regions where high numbers of species may occur.

Expert ranges are based not only on exhaustive knowledge of the respective taxon, but also of the survey area, thus reducing the risk of errors of commission compared to convex polygons. In addition, expert ranges will include grid cells which have not been sampled but presumably contain the species in question and therefore decrease the effects of heterogeneous sampling effort. Thus, expert ranges should represent a valuable source for large-scale assessments of species richness and are probably much more reliable than convex polygons. However, the quality of expert range data is highly dependent on the expertise of the respective specialist(s), and different taxonomists will inevitably generate different range maps for one and the same species (Schulman *et al.*, 2007).

Compared with expert ranges, convex polygons rely only on point occurrences and thus, error of commission is considerably higher. Overestimation of ranges is particularly high when disjunct species occurrences are fitted into one contiguous polygon. As a workaround, a more complex approach to fit better convex polygons to underlying occurrences has been suggested (Burgman & Fox, 2003; IUCN Standards and Petitions Subcommittee, 2010): an alpha hull is constructed using only those occurrences which are connected by a line being a multiple *a* of the average line length. For a large *a*, constructed ranges are coarse (at maximum the convex hull), for a smaller *a*, constructed ranges become finer (at minimum the occurrence points). In the absence of a standard solution to select an adequate *a* for species with occurrences varying from four to 139, we used the simple convex polygon approach.

In spite of using a small set of predictors to avoid overfitting of the model, the use of six predictors would ideally require at least 60 occurrences per species (Harrell, 2001). Our modelled ranges rely on fewer than 60 occurrences for most species (90%). Moreover, the potential identification of different predictors and the potential use of different modelling algorithms would have resulted in different ranges and derived richness maps, but for comparison of different modelling approaches see for example Elith *et al.* (2006) and Phillips *et al.* (2006). Subject to these uncer-

tainties, range sizes for most species were predicted to be rather large, as reflected by the highest median (Table 1). However, the number of largest range sizes (greater than 316 grid cells) was comparatively small (Fig. 2g–h, Table 1). Furthermore, the maximum size of modelled ranges was far lower than of convex polygons and less than half the size of expert ranges. Thus, ranges modelled for widespread species have a lower risk of error of commission.

Spatial structure and congruence of species richness

The disparities in range size are also clearly reflected by disparities in resulting species richness patterns (Fig. 3a–d). The most conspicuous disparity between point-to-grid richness and all other types of species richness also becomes evident in the low spatial autocorrelation of point-to-grid richness (Fig. 3e) corroborating the strong porosity of this distribution dataset. These differences are in parallel to findings by McPherson & Jetz (2007), but compared to their study, the differences we found were even more pronounced. Furthermore, similar to their study, our modelled richness showed a lower level of spatial autocorrelation than expert richness (Fig. 3e). In contrast to McPherson & Jetz (2007) and Graham & Hijmans (2006), we found that modelled richness was slightly lower than corresponding expert richness (Fig. 4e). Pair-wise comparisons of richness patterns (Fig. 4) otherwise confirmed differences between point-to-grid richness and all other richness variables as well as the relatively high spatial congruence of expert range and convex polygon richness.

One possible solution to the problems associated with incompleteness of distribution data and heterogeneous data quality would be to coarsen the resolution as patterns of richness generated with different data types should converge with coarser grain (Hurlbert & Jetz, 2007). Up to a resolution of 2°, pair-wise correlations of point-to-grid richness with all other variables were relatively low in our study (Fig. 5a). Only at 5° grid resolution did all combinations show high correlation ($r \geq 0.9$). While such low-resolution maps can be helpful to obtain a broad overview on distribution patterns, their usefulness is rather limited in an applied context (Graham & Hijmans, 2006). For instance, Grenyer *et al.* (2006) found that most protected areas are smaller than 1.53 km². In a hypothetical 5° grid cell (equalling roughly 250 000 km²), the associated species can occur anywhere within the grid cell, probably outside of protected areas. Therefore, in particular for analyses with a conservation background, the best scale will be the finest.

While a less strict definition of a richness centre led to an increased spatial congruence of richness pairs at 1° (Fig. 5b), graphs of discrete pairs do not converge. This indicates pronounced differences of richness centres irrespective of the richness threshold used. Again the high spatial congruence of expert range and convex polygon richness is

noticeable. One explanation for their consistently high correlation is that the degree of filling of gaps between documented occurrences is comparable for both methods. However, since expert ranges tend to distinguish disjunct occurrences and to exclude unfavourable areas, the commission error will be lower than for convex polygons. This commission error, however, is not reflected in the correlation coefficients, but becomes apparent in the continuous distribution of convex polygon richness in the Cerrado and the Caatinga (Fig. 3c). Nevertheless, both types of species ranges will rather over-estimate species richness.

Conclusions

The construction of species ranges even for comparatively well-researched tropical plants to analyse biogeographical patterns remains a challenge. Although it is most desirable to draw on standardised methods to deduce sound distribution patterns, there is clearly no silver bullet for the construction of species ranges and richness maps from distribution data at large spatial scales. With regard to the differences contrasted here, the matching of range construction to the available data sources and the due consideration of errors associated with range construction are essential to reduce the bias in our knowledge of distribution patterns.

In the lack of a validation dataset, a typical feature of distribution datasets from tropical regions, the error associated with the construction of ranges cannot be directly quantified. However, the general quality of errors inherent to the method for constructing species ranges as illustrated here has to be considered and has further to be communicated to the users of resulting distribution maps. When species ranges are used as a basis for further analysis and recalculated to a too-fine grain size erroneous presence will be inferred. This error of commission will influence derived models, and thus increase the rates of error compared with the use of original occurrence data (Hurlbert & Jetz, 2007).

Narrow-ranging species often represent a substantial portion of the investigated species and are in the focus of conservation efforts. One option to better integrate these species into large-scale research is the use of geometric approaches which need not meet the statistical requirements modelled ranges are bound to. Such approaches represent a repeatable version of the construction of expert ranges and can also be used when working with large datasets (Hopkins, 2007; Raedig et al., 2010). Another alternative is represented by expert-based range maps which directly integrate both geometric and environmental modelling approaches (Graham & Hijmans, 2006). The best approach for most large-scale distribution datasets probably will be a mixture of existing methods adapted to the special characteristics of the underlying dataset.

One important finding of this study is that the construction of each range type at one point or another requires expert knowledge on the species' ecology. This is evident

for expert ranges, but also necessary for modelled ranges (parameter, model and threshold selection) as well as for geometric approaches such as the alpha hull (determination of factor a) or the buffering of occurrence points (selection of buffer width). Thus, parallel to analysing available distribution data, taxonomic base-line work in order to assemble reliable distribution and ecological data must not be neglected but rather intensified (e.g. Prance et al., 2000; Gropp, 2003; Morawetz & Raedig, 2007; Cayuela et al., 2009), in particular for megadiverse tropical clades.

References

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

Araújo, M.B. & New, M. 2007. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* **22**, 42–47.

Balick, M., Anderson, A. & Freitas Da Silva, M. 1982. Palm taxonomy in Brazilian Amazônia: the state of systematic collections in regional herbaria. *Brittonia* **34**, 463–477.

Beck, S.G., Hernandez, P.A., Jørgensen, P.M., Paniagua, L., Timaná, M.E. & Young, B.E. 2007. Vascular plants. In: Young, B.E., Ed., *Endemic Species Distributions on the East Slope of the Andes in Peru and Bolivia*. NatureServe, Arlington, pp. 18–34.

Bletter, N., Janovec, J., Brosi, B. & Daly, D.C. 2004. A digital base map for studying the Neotropical flora. *Taxon* **53**, 469–477.

Botkin, D.B., Saxe, H., Araújo, M.B., Betts, R., Bradshaw, R.H.W., Cedhagen, T., Chesson, P., Dawson, T.P., Etterson, J.R., Faith, D.P., Ferrier, S., Guisan, A., Hansen, A.S., Hilbert, D.W., Loehle, C., Margules, C., New, M., Sobel, M.J. & Stockwell, D.R.B. 2007. Forecasting the effects of global warming on biodiversity. *BioScience* **57**, 227–236.

Buckley, L.B. & Jetz, W. 2007. Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B: Biological Sciences* **274**, 1167–1173.

Burgman, M.A. & Fox, J.C. 2003. Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Animal Conservation* **6**, 19–28.

Cayuela, L., Golicher, D.J., Newton, A., Kolb, M., de Alburquerque, F.S., Arets, E.J.M.M., Alkemade, J.R.M. & Pérez, A.M. 2009. Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for

effective species conservation. *Tropical Conservation Science* **2**, 319–352.

CONDIT, R., PITMAN, N., LEIGH, E.G., JR., CHAVE, J., TERBORGH, J., FOSTER, R.B., NUNEZ, P., AGUILAR, S., VALENCIA, R., VILLA, G., MULLER-LANDAU, H.C., LOSOS, E. & HUBBELL, S.P. 2002. Beta-diversity in tropical forest trees. *Science* **295**, 666–669.

CRISP, M.D., LAFFAN, S., LINDER, H.P. & MONRO, A. 2001. Endemism in the Australian flora. *Journal of Biogeography* **28**, 183–198.

DINIZ-FILHO, J.A.F., BINI, L.M., VIEIRA, C.M., BLAMIRES, D., TERRIBILE, L.C., BASTOS, R.P., DE OLIVEIRA, G. & DE SOUZA BARRETO, B. 2008. Spatial patterns of terrestrial vertebrate species richness in the Brazilian Cerrado. *Zoological Studies* **47**, 146–157.

DORMANN, C.F., MCPHERSON, J.M., ARAÚJO, M.B., BIVAND, R., BOLLIGER, J., CARL, G., DAVIES, R.G., HIRZEL, A., JETZ, W., KISSLING, D.W., KÜHN, I., OHLEMÜLLER, R., PERES-NETO, P.R., REINEKING, B., SCHRÖDER, B., SCHURR, F.M. & WILSON, R. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628.

ELITH, J., GRAHAM, C.H., ANDERSON, R.P., DUDÍK, M., FERRIER, S., GUISAN, A., HIJMANS, R.J., HUETTMANN, F., LEATHWICK, J.R., LEHMANN, A., LI, J., LOHMANN, L.G., LOISELLE, B.A., MANION, G., MORITZ, C., NAKAMURA, M., NAKAZAWA, Y., OVERTON, J., PETERSON A.T., PHILLIPS, S.J., RICHARDSON, K., SCACHETTI-PEREIRA, R., SCHAPIRE, R.E., SOBERÓN, J., WILLIAMS, S., WISZ, M.S. & ZIMMERMANN, N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151.

FIELDING, A.H. & BELL, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24**, 38–49.

FORTIN, M.J., DALE, M.R.T. & VER HOEF, J. 2002. Spatial analysis in ecology. In: EL-SHAARAWI, A.H. & PIEGORSCH, W.W., Eds., *Encyclopedia of Environmetrics.* Wiley, Chichester, pp. 2051–2058.

GASTON, K.J. 1991. How large is a species' geographic range? *Oikos* **61**, 434–438.

GASTON, K.J. 2003. *The structures and dynamics of geographic ranges*, 2nd edition. University Press, Oxford.

GASTON, K.J. & FULLER, R.A. 2009. The sizes of species' geographic ranges. *Journal of Applied Ecology* **46**, 1–9.

GIBSON, L., BARRETT, B. & BURBIDGE, A. 2007. Dealing with uncertain absences in habitat modelling: a case study of a rare ground-dwelling parrot. *Diversity and Distribution* **13**, 704–713.

GOVE, A.D., DUNN, R.R. & MAJER, J.D. 2008. The importance of species range attributes and reserve configuration for the conservation of angiosperm diversity in Western Australia. *Biodiversity and Conservation* **17**, 817–831.

GRAHAM, C.H., FERRIER, S., HUETTMAN, F., MORITZ, C. & PETERSON, A.T. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution* **19**, 497–503.

GRAHAM, C.H. & HIJMANS, R.J. 2006. A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography* **15**, 578–587.

GRAHAM, M.H. 2003. Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809–2815.

GRENYER, R., ORME, C.D.L., JACKSON, S.F., THOMAS, G.H., DAVIES, R.G., DAVIES, T.J., JONES, K.E., OLSON, V.A., RIDGELY, R.S., RASMUSSEN, P.C., DING, T., BENNETT, P.M., BLACKBURN, T.M., GASTON, K.J., GITTLEMAN, J.L. & OWENS, I.P.F. 2006.

Global distribution and conservation of rare and threatened vertebrates. *Nature* **444**, 93–96.

GROPP, R.E. 2003. Threatened species: university natural science collections in the United States. *Systematics and Biodiversity* **1**, 285–286.

GUISAN, A. & ZIMMERMANN, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**, 147–186.

GUISAN, A. & THUILLER, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**, 993–1009.

GURALNICK, R. & VAN CLEVE, J. 2005. Strengths and weaknesses of museum and national survey data sets for predicting regional species richness: comparative and combined approaches. *Diversity and Distributions* **11**, 349–359.

HARRELL, F.E.J. 2001. *Regression Modeling Strategies – With Applications to Linear Models, Logistic Regression, and Survival Analysis.* Springer, New York.

HENDERSON, A. 2000. *Bactris* (*Palmae*)*,* Flora Neotropica Monograph 79. The New York Botanical Garden, Bronx, New York.

HENDERSON, A., GALEANO, G. & BERNAL, R. 1995. *Field Guide to the Palms of the Americas*. Princeton University Press, Princeton.

HERNÁNDEZ, H.M. & NAVARRO, M. 2007. A new method to estimate areas of occupancy using herbarium data. *Biodiversity and Conservation* **16**, 2457–2470.

HERNANDEZ, P.A., GRAHAM, C.H., MASTER, L.L. & ALBERT, D.L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* **29**, 773–785.

HIJMANS, R.J., CAMERON, S.E., PARRA, J.L., JONES, P.G. & JARVIS, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965–1978.

HOPKINS, M.J.G. 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. *Journal of Biogeography* **34**, 1400–1411.

HURLBERT, A.H. & WHITE, E.P. 2005. Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters* **8**, 319–327.

HURLBERT, A.H. & JETZ, W. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences USA* **104**, 13384–13389.

IUCN STANDARDS AND PETITIONS SUBCOMMITTEE. 2010. *Guidelines for Using the IUCN Red List Categories and Criteria.* Version 8.0. Prepared by the Standards and Petitions Subcommittee in March 2010. http://intranet.iucn.org/ webfiles/doc/SSC/RedList/RedListGuidelines.pdf (accessed on 1 May 2010).

JALAS, J. & SUOMINEN, J., Eds. 1972–1994. *Atlas Florae Europaea*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

JALAS, J., SUOMINEN, J. & LAMPINEN, R., Eds. 1996. *Atlas Florae Europaeae*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

JALAS, J., SUOMINEN, J., LAMPINEN, R. & KURTTO, A., Eds. 1999. *Atlas Florae Europaeae*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

JETZ, W. & RAHBEK, C. 2002. Geographic range size and determinants of avian species richness. *Science* **297**, 1548–1551.

JETZ, W., SEKERCIOGLU, C.H. & WATSON, J.E.M. 2008. Ecological correlates and conservation implications of overestimating species geographic ranges. *Conservation Biology* **22**, 110–119.

KREFT, H., SOMMER, J.H. & BARTHLOTT, W. 2006. The significance of geographic range size for spatial diversity patterns in Neotropical palms. *Ecography* **29**, 21–30.

KÜPER, W., SOMMER, J.H., LOVETT, J.C., MUTKE, J., LINDER, H.P., BEENTJE, H.J., ANGELE, R.S., VAN ROMPAEY, R.S.A.R., CHATELAIN, C., SOSEF, M. & BARTHLOTT, W. 2004. Africa's hotspots of biodiversity redefined. *Annals of the Missouri Botanical Garden* **91**, 525–535.

KURTTO, A., LAMPINEN, R. & JUNIKKA, L., Eds. 2004. *Atlas Florae Europaeae*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

KURTTO, A., FRÖHNER, S.E. & LAMPINEN, R., Eds. 2007. *Atlas Florae Europaeae*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

LENNON, J.J., KOLEFF, P., GREENWOOD J.J.D. & GASTON, K.J. 2004. Contribution of rarity and commonness to patterns of species richness. *Ecology Letters* **7**, 81–87.

LINDER, H.P. 2001. Plant diversity and endemism in sub-Saharan tropical Africa. *Journal of Biogeography* **28**, 169–182.

LUOTO, M., PÖYRY, J., HEIKKINEN, R.K. & SAARINEN, K. 2005. Uncertainty of bioclimate envelope models based on the geographical distribution of species. *Global Ecology and Biogeography* **14**, 575–584.

MCPHERSON, J.M. & JETZ, W. 2007. Type and spatial structure of distribution data and the perceived determinants of geographical gradients in ecology: the species richness of African birds. *Global Ecology and Biogeography* **16**, 657–667.

MOAT, J. 2007. Conservation assessment tools extension for ArcView 3.x. http://www.kew.org/gis/projects/cats/catsdoc.pdf (accessed 1 May 2010). Royal Botanic Gardens, Kew, London.

MORAWETZ, W. & RAEDIG, C. 2007. Angiosperm biodiversity, endemism and conservation in the Neotropics. *Taxon* **56**, 1245–1254.

NELSON, B.W., CARLOS, A.C.F., DA SILVA, M.F. & KAWASAKI, M.L. 1990. Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature* **345**, 714–716.

OLSON, D.M., DINERSTEIN, E., WIKRAMANAYAKE, E.D., BURGESS, N.D., POWELL, G.V.N., UNDERWOOD, E.C., D'AMICO, J.A., ITOUA, I., STRAND, H.E., MORRISON, J.C., LOUCKS, C.J., ALLNUTT, T.F., RICKETTS, T.H., KURA, Y., LAMOREUX, J.F., WETTENGEL, W.W., HEDAO, P. & KASSEM, K.R. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51**, 933.

ORME, C.D.L., DAVIES, R.G., BURGESS, M., EIGENBROD, F., PICKUP, N., OLSON, V.A., WEBSTER, A.J., DING, T., RASMUSSEN, P.C., RIDGELY, R.S., STATTERSFIELD, A.J., BENNETT, P.M., BLACKBURN, T.M., GASTON, K.J. & OWENS, I.P.F. 2005. Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**, 1016–1019.

PEARSON, R.G., RAXWORTHY, C.J., NAKAMURA, M. & TOWNSEND, P.A. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* **34**, 102–117.

PHILLIPS, S.J., ANDERSON, R.P. & SCHAPIRE, R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259.

PRANCE, G.T., BEENTJE, H., DRANSFIELD, J. & JOHNS, R. 2000. The tropical flora remains undercollected. *Annals of the Missouri Botanical Garden* **87**, 67–71.

RAEDIG, C., DORMANN, C., HILDEBRANDT, A. & LAUTENBACH, S. 2010. Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. *Biodiversity and Conservation* **19**, 1523–1546.

RAPOPORT, E. 1982. *Areography: geographical strategies of species*. Pergamon Press, Oxford.

RAXWORTHY, C.J., MARTINEZ-MEYER, E., HORNING, N., NUSSBAUM, R.A., SCHNEIDER, G.E., ORTEGA-HUERTA, M.A. & PETERSON, A.T. 2003. Predicting distributions of known and unknown reptile species in Madagascar. *Nature* **426**, 837–841.

RONDININI, C., WILSON, K.A., BOITANI, L., GRANTHAM, H. & POSSINGHAM, H.P. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters* **9**, 1136–1145.

SCHIPPER, J., CHANSON, J.S., CHIOZZA, F., COX, N.A., HOFFMANN, M., KATARIYA, V., LAMOREUX, J., RODRIGUES, A.S.L., STUART, S.N., TEMPLE, H.J., BAILLIE, J., BOITANI, L., LACHER, T.E., JR., MITTERMEIER, R.A., SMITH, A.T., ABSOLON, D., AGUIAR, J.M., AMORI, G., BAKKOUR, N., BALDI, R., BERRIDGE, R.J., BIELBY, J., BLACK, P.A., BLANC, J.J., BROOKS, T.M. *ET AL.* 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**, 225–230.

SCHMIDT, M., KREFT, H., THIOMBIANO, A. & ZIZKA, G. 2005. Herbarium collections and field data as a base of phytodiversity maps for Burkina Faso. *Diversity and Distributions* **11**, 509–516.

SCHULMAN, L., TOIVONEN, T. & RUOKOLAINEN, K. 2007. Analysing botanical collecting effort in Amazonia and correcting for it in species range estimation. *Journal of Biogeography* **34**, 1388–1399.

SEGURADO, P. & ARAÚJO, M.B. 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography* **31**, 1555–1568.

THUILLER, W., VAYREDA, J., PINO, J., SABATE, S., LAVOREL, S. & GRACIA, C. 2003. Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). *Global Ecology and Biogeography* **12**, 313–325.

TSOAR, A., ALLOUCHE, O., STEINITZ, O., ROTEM, D. & KADMON, R. 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions* **13**, 397–405.

WILLIAMS, J.N., SEO, C., THORNE, J., NELSON, J.K., ERWIN, S., O'BRIEN, J.M. & SCHWARTZ, M.W. 2009. Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions* **15**, 565–576.

WILLIS, F., MOAT, J. & PATON, A. 2003. Defining a role for herbarium data in Red List assessments: a case study of *Plectranthus* from eastern and southern tropical Africa. *Biodiversity and Conservation* **12**, 1537–1552.

WISZ, M.S., HIJMANS, R.J., LI, J., PETERSON, A.T., GRAHAM, C.H. & GUISAN, A. 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* **14**, 763–773.

Associate Editor: Charlie Jarvis