What Do Range Maps and Surveys Tell Us About Diversity Patterns?

**Bradford A. Hawkins · Marta Rueda ·
Miguel Á. Rodríguez**

Palmer et al. (2008) begin their article by pointing out that diversity numbers "*should mean something*". Clearly, it is futile to try to understand diversity patterns if we do not know what the patterns really are. Thus, the first steps in studying geographic diversity patterns are to decide how to generate quantitative estimates, followed by an evaluation of their robustness, if possible. Unfortunately, as the other papers in this forum clearly demonstrate, this is more easily said than done.

There are currently three methods to generate diversity patterns over large geographical areas. The most widely used to date is to overlay distribution (range) maps and sum the maps that overlap a predefined point or grid system. The resultant maps are referred to as either richness maps or, more properly, range overlap maps (LaSorte and Hawkins 2007). Alternatively, richness maps are generated using broadly based systematic surveys that actually record the species observed in sites, transects or grids. Such surveys have historically been very limited over large scales, but the recent interest in climate change and biological conservation has led to the

B. A. Hawkins (✉)
Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA
e-mail: bhawkins@uci.edu

M. Rueda · M. Á. Rodríguez
Department of Ecology, University of Alcalá, 28871 Alcalá de Henares, Madrid, Spain

M. Rueda
e-mail: marta.rueda@uah.es

M. Á. Rodríguez
e-mail: miguela.rodriguez@uah.es

development of continental surveys for a wide range of plant and animal groups encompassing Europe, North America and Australia (e.g., Jalas and Suominen 1972–1996; Kudrna 2002; Barrett et al. 2003; USGS Patuxent Wildlife Research Center 2006). Regionally based surveys are also common (e.g., Lewis and Pomeroy 1989; Bons and Geniez 1996; Preston et al. 2002; Martí and del Moral 2003; Loffler and Loffler 2005; El Din 2006). Thus, we can expect more large-scale diversity analyses to be based on survey data in the future. The final method used to generate diversity data is that employed by Palmer et al. (2008) in their first data set, the vascular floras of North America. In this case, multiple, independently generated data sources are combined to generate an overall picture of diversity. Palmer et al. (2008) discuss issues associated with this method, and we do not pursue it further here. Rather, we focus on the two most widely used methods and attempt an evaluation of what range maps and surveys tell us about broad-scale patterns of diversity.

The advantages and disadvantages of survey-based and map-based data have been repeatedly discussed in the diversity literature (Freitag and Van Jaarsveld 1995; Hurlbert and White 2005; Hurlbert and Jetz 2007), but what is clear is that neither method is perfect, as surveys will invariably differ in sample intensity both spatially and temporally and may miss rare species (generating false negatives), whereas range maps, by filling in gaps between records, can record species being present in places where they are not (generating false positives). These contrasting properties have led to the conclusion that map-based data are inherently more coarse grained, whereas survey data are relatively more fine grained (Hurlbert and White 2005). This is almost certainly true. It has also been argued that map-based data are less reliable than survey data and contain less information, especially at finer scales (Hurlbert and Jetz 2007). But underlying this latter conclusion is the assumption that survey data are less error prone and thus are 'truer' than distribution maps, which must also be true at some scale (e.g., range maps would undoubtedly overestimate richness in hectare sized plots). On the other hand, given that undersampling is known to be important in even intense sample regimes (e.g. Soria-Auza and Kessler 2008) and in the best studied parts of the world (e.g. Prendergast and Eversham 1995), it may be that neither method should be relied on uncritically.

In this paper we explore this issue, focusing on three questions: (1) how do geographic richness patterns differ between survey- and map-based data, (2) at what grain, if any, do any differing patterns converge, and (3) are map-based diversity estimates necessarily inflated? Our approach generally follows that used by Hurlbert and Jetz (2007) in their evaluation of these questions for South African and Australian birds, except our geographical focus is on western and central Europe, where sampling intensity and distribution data are very good relative to the rest of the world. We also agree with Freitag and Van Jaarsveld (1995) that survey data suffer from undersampling problems that may also generate misleading diversity patterns. We note that our data are for animals rather than plants, but the issues involved are general and should be considered in all studies of diversity patterns.

Our database comprises all non-marine herptiles (47 amphibian plus 78 reptile species) and mammals (139 species) native to continental western Europe and Great Britain (all other islands were excluded). Herptile range and survey distribution maps were obtained from Arnold (2002) and Gasc et al. (1997), respectively. For mammals, range maps were extracted from the European Mammal Assessment

database (IUCN 2007), and survey maps from Mitchell-Jones et al. (1999). Eastern European areas (i.e., Belarus, Russia and Ukraine) were excluded because of obviously poor sampling in the case of the herptile survey data (Gasc et al. 1997), and because this region was not covered by the mammal survey (Mitchell-Jones et al. 1999). In all cases, we digitized the maps in ArcGIS 9.2 to generate the herptile and mammal species richness data in three grids of differing grain, all comprising nearly square cells. The smallest grain coincided with the survey map sources, consisting of UTM cells of an average area of 2,500 km^2 (~50 km grain) (for details see Gasc et al. 1997 and Mitchell-Jones et al. 1999). The intermediate and large grains comprised cells of average areas of 10,000 km^2 (~100 km grain), and 22,500 km^2 (~150 km grain), generated by joining four adjacent cells of the original UTM grid in the first case, and nine adjacent cells in the second case. However, we increased the numbers of joined cells around coasts and major lakes for both grids so that these cells had similar areas to inland cells when possible. Once the grids were generated, any coastal cells containing less than 50% of the land mass of inland cells were excluded from the analysis, leaving 1,920, 513 and 263 cells, respectively. Evaluation of the data first comprised mapping the cell richness values for each group and grain size to picture the resulting geographical patterns. We then correlated the cell map- vs. survey-based richness values for each group and grain to determine how closely coupled the values were for the alternative methods used to estimate diversity. Finally, we mapped the differences in the richness values for herptiles and mammals obtained using maps vs. surveys to examine the spatial structure of any lack of congruence in the data sets. This final method was used only for the ~50 km grain data due to strong congruence of patterns at the larger grains.

We found very similar patterns for both herptiles and mammals. Most striking is that five of the six maps for both groups reveal very similar spatial patterns (Fig. 1b–f for herptiles, and Fig. 2b–f for mammals); central Europe is characterized by smooth gradients of richness increasing to the south, whereas in the west richness increases to the south and then reverses in Iberia and to a lesser extent in Italy. The only non-congruent patterns are found in the survey data at the ~50 km grain (Figs. 1a and 2a), in which gradients are generally more erratic and with strong reversals in the southeast. Because the spatial patterns are very similar across all data at larger grains, it is not surprising that survey estimates and map estimates are also very strongly correlated at both ~100 and ~150 km grains (Fig. 3b–c for herptiles and 3e–f for mammals), despite the existence of outliers in a few cells at the eastern edges of the grid systems where undersampling during the surveys is likely to be an issue. It is also notable that the absolute richness estimates derived from surveys and maps are very similar, and cell richness estimated by maps is usually close to those found during the surveys. In contrast, as expected from Figs. 1a,d and 2a,d, the correlations between map and survey estimates are much noisier at the smallest grain (Fig. 3a,d), and estimates based on the maps are very often much higher than the survey data. We also found strong spatial structure in those cells where survey estimates were substantially lower than map estimates, with the strongest underestimates located in the Mediterranean region and Balkans (Fig. 4).

Our results are close to those recently reported by Hurlbert and Jetz (2007), with the difference being that they found that map and survey data generated similar patterns at a minimum grain size of 220 km, whereas we found almost
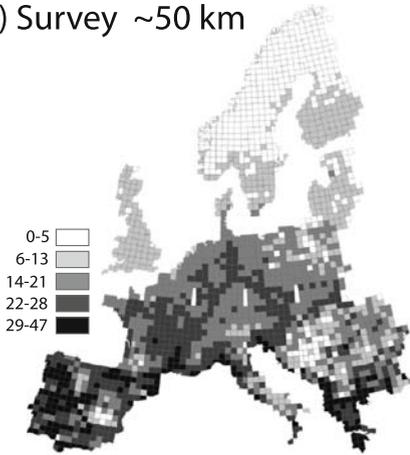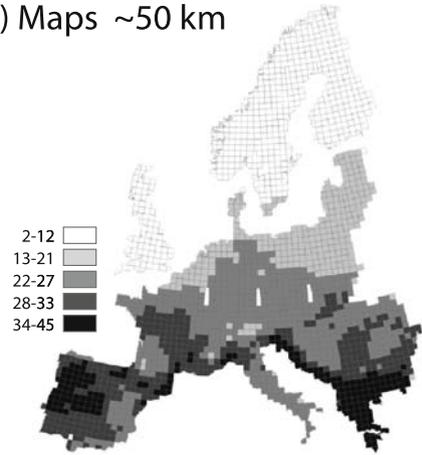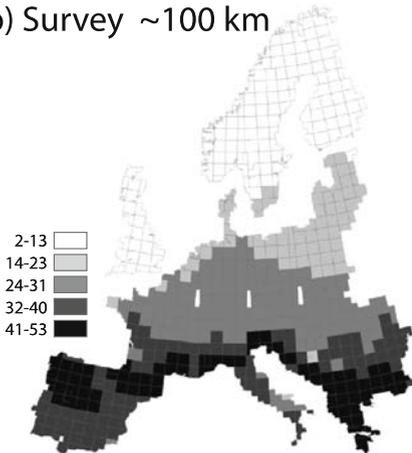
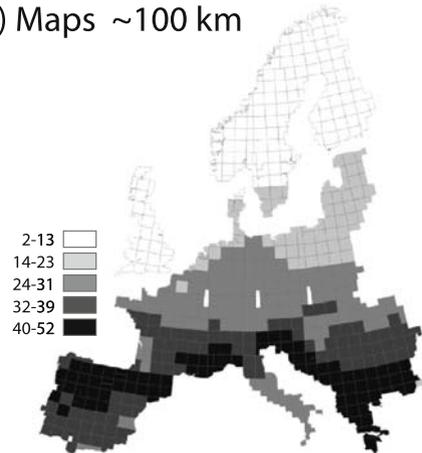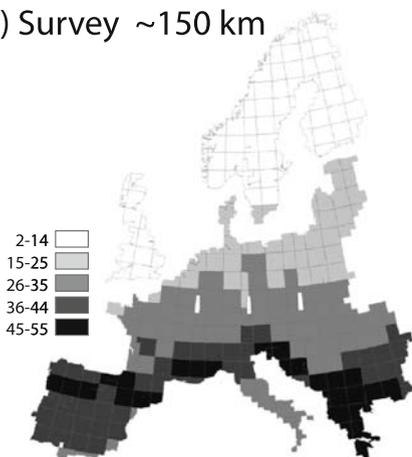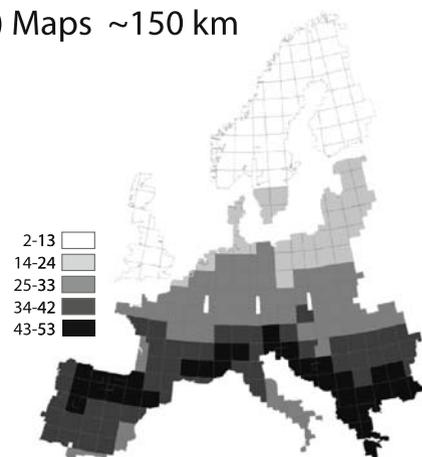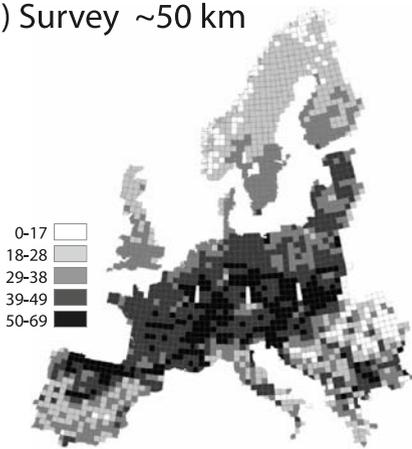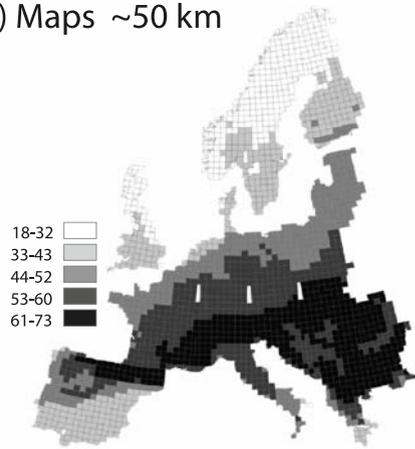**Fig. 1** Species richness maps for European herptiles (reptiles and amphibians) based on survey data (**a**, **b**, **c**) and range maps (**d**, **e**, **f**), at three grain sizes
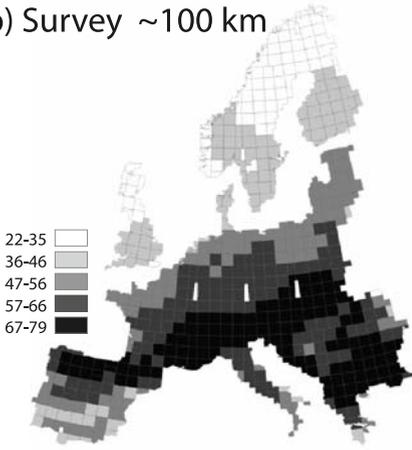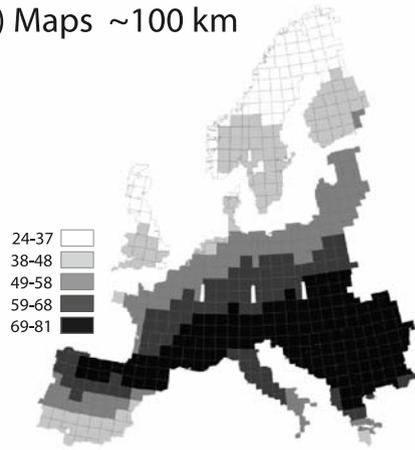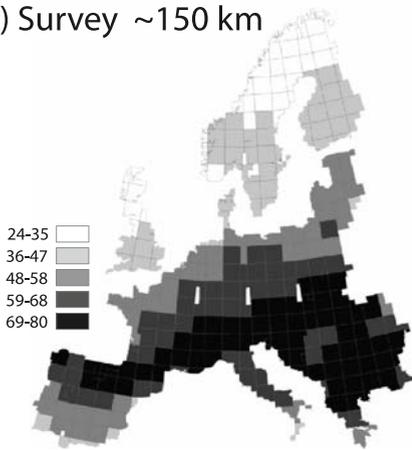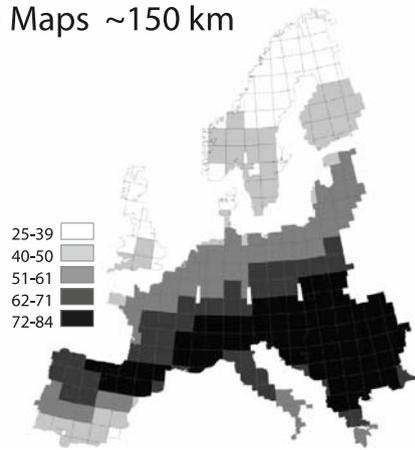
Fig. 2 Species richness maps for European mammals based on survey data (**a**, **b**, **c**) and range maps (**d**, **e**, **f**), at three grain sizes
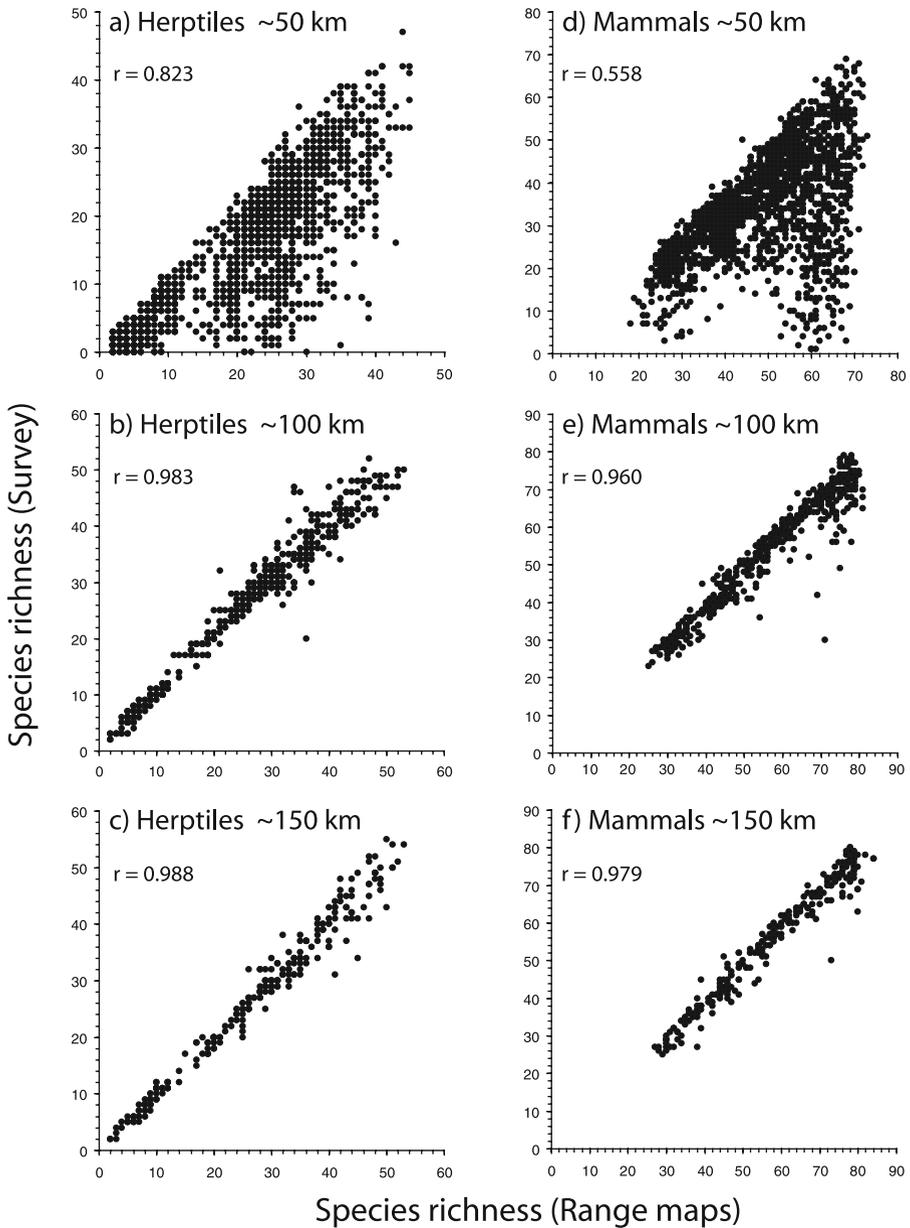
Fig. 3 Scatterplots of cell richness values for survey data versus range map data for herptiles (**a**, **b**, **c**) and mammals (**d**, **e**, **f**) at three grain sizes

identical patterns at 100 km. Although this slight difference may not appear to be serious, we believe it is key to understanding the relationship between survey and map data.

The obvious challenge to all grid-based surveys is to sample at least one site in every cell, and to sample all cells intensively enough to record the presence of

Fig. 4 Geographic distribution of differences between richness estimates derived from range maps and estimates derived from the continental surveys at the ~50 km grain size for (a) herptiles and (b) mammals. Negative values indicate that range maps underestimate the number of species found in cells during the surveys, whereas positive values indicate that the maps predict more species than actually observed
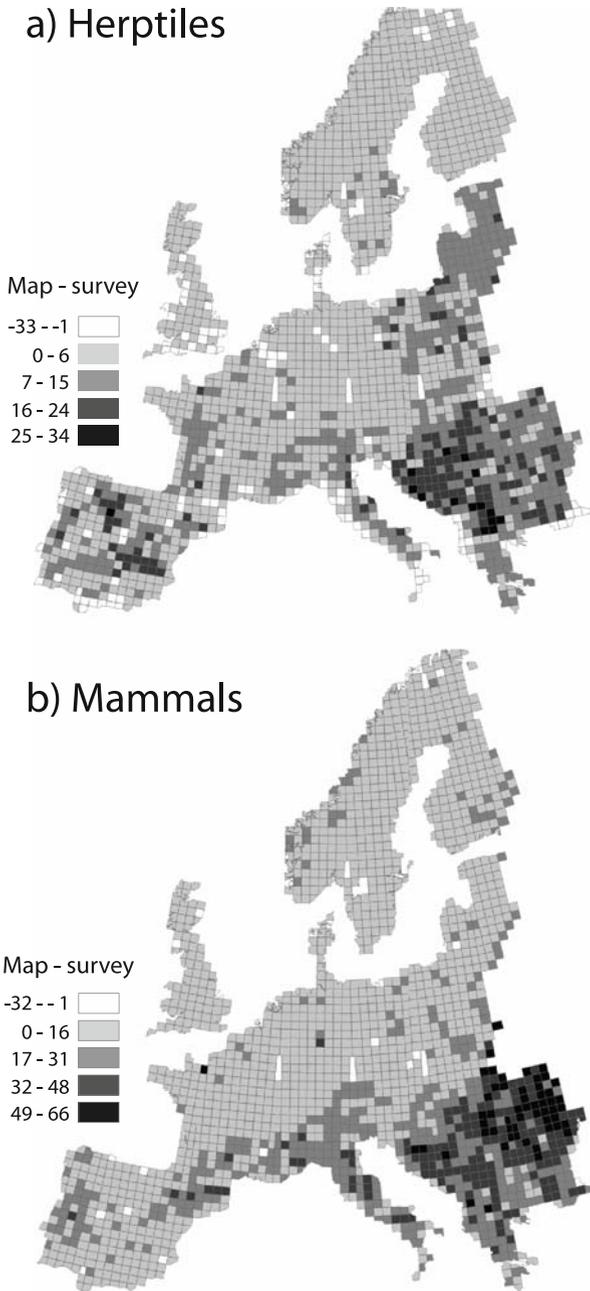
(ideally) all resident species. Unfortunately, this is essentially impossible to do in practice; even the surveys of Great Britain, which has the best documented fauna and flora in the world, have not adequately sampled all survey cells covering the island, complicating attempts to identify diversity hotspots and evaluate responses of

plants and animals to climate change and human disturbance (Prendergast and Eversham 1995). This is also true for Europe in general. For example, the survey sources we use here report undersampling in both the Mediterranean region and the Balkans, which probably accounts for the apparent drop in both herptile and mammal richness to the southeast in the original UTM data (see Figs. 1 and 2) and the strong differences between the map-based and survey-based data in the south (see Fig. 4). Clearly, if undersampling is a problem in parts of the world with very dense road networks and a large pool of workers to conduct surveys over many years, it is even more serious an issue in regions with sparse roads, relatively few survey teams or surveys conducted for only a few years. Thus, given the inevitability of sampling problems in geographically extensive surveys, it would be unwise to accept the validity of patterns derived from surveys without careful evaluation. On the other hand, map-based data also clearly have problems as well, so the key question becomes, is there a way to resolve the dilemma of which diversity patterns to believe? The results of Hurlbert and Jetz (2007) and ours provide a tentative answer to this question.

Both analyses find that at larger grains map-based and survey-based geographic patterns are very similar, the difference being the grain at which patterns become congruent. Hurlbert and Jetz (2007) found that grains of at least 220 km were necessary, whereas we find strong similarity at 100 km. One possible reason for this difference is that they were concerned with birds, whereas we analyze the less mobile herptiles and mammals, and perhaps the various groups have different occupancy patterns, with birds being more patchily distributed. However, an alternative explanation is related to relative sampling problems when comparing Europe with South Africa and Australia. Western and central Europe is without doubt much better sampled than Africa and Australia, and the survey data are relatively more complete, with fewer gaps in coverage. Thus, with the increase in the quality of the survey data, the grain at which map and survey data converge decreases. This also means that not only are richness estimates based on range maps very robust at large grains (200+ km), but they are also accurate at moderate grains (100 km), and apparent discrepancies in patterns at moderate scales could be due to survey error rather than map error. Thus, based on the results currently available, it appears that at coarse grains map and survey data can be used interchangeably in most parts of the world, but in regions where surveys are likely to be incomplete, maps can generate more accurate patterns at moderate grains than do surveys.

In the absence of error-free survey data it is not possible to determine the grain at which map-based patterns lose accuracy, but it may be at something less than 50 km, where in the well-sampled parts of northern and western Europe the finest grained map and survey data do not strongly differ (see Fig. 4). On the other hand, we can be certain that the map-derived data are accurate at 100 km, since the survey and map data are very similar even in undersampled regions (see Figs. 1 and 2). It could be argued that the map data are not in fact accurate at 50 km, given the differences in the patterns we find within Europe, but this does not mean that the survey data can be taken at face value at the smallest grain either, and other papers in this forum address the range of issues with finer-scale diversity data and how to deal with sampling problems. Irrespective, if the goal of an analysis is to document broad-scale patterns at moderate to large grains, map data can be as reliable as survey data

and in some cases they may be even more reliable. Simulation analyses have also shown that range-map data can generate very accurate geographical diversity patterns, and it is possible to assess probable error rates under different ecological conditions (LaSorte and Hawkins 2007).

If the comparisons of survey and map data conducted to date are generally applicable, this represents a boon to biodiversity analyses. Given the extreme difficulty for surveys to quantify accurately the diversity of any major taxonomic group even in very well known parts of the world, it is obvious that we must be very cautious about using survey data from less well studied regions. It is of course these latter areas where diversity data are most urgently needed, as they include the hyper-rich tropics under threat from human impacts and climate change. In other words, if we have to wait until complete survey data are available for a continental grid system in Africa or South America to know how diversity hotspots arise and are maintained, it will be too late, despite our best efforts to overcome gaps in our knowledge. But recent advances in converting occurrence records into distribution maps, which can then be used to study broad-scale diversity patterns, provide a shortcut that will allow us to document diversity sooner rather than later. The realization that gaps in knowledge of species' distributions can be filled using 'bioclimatic envelopes' and 'niche modeling' has lead to a very active research community (Guisan and Zimmermann 2000; Pearson and Dawson 2003; Pearson et al. 2006; Araújo and New 2007; Guisan et al. 2007) dedicated to generating range maps in both extra-tropical and tropical regions (e.g., Buckland and Elston 1993; Peterson et al. 2002; Thuiller et al. 2003; Chefaoui et al. 2005). As these maps become available, they can be converted into richness data and subject to the full range of biodiversity analyses (see e.g., Williams et al. 2003). Of course, the maps have to be at least as reliable as maps generated by other currently used methods (poor quality maps cannot be expected to generate high quality richness data), and the grain of the analysis cannot be too small. But, current evidence suggests that broad-scale patterns derived from range maps are not strongly affected by the details of the maps used, even in the tropics (Mathias et al. 2004), so if diversity patterns generated by maps are as robust as they appear to be, regional, continental and global-scale diversity analyses for a wide range of taxonomic groups for which survey data do not exist will soon be possible. We strongly endorse this approach as a solution to the problem of understanding diversity in the face of imperfect knowledge.

References

Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22: 42–47

Arnold, EN (2002) *A field guide to the reptiles and amphibians of Britain and Europe*. Harper Collins, London

Barrett G, Silcocks A, Barry S, Cunningham R, Poulter R (2003) *The new atlas of Australian birds*. CSIRO Publishing, Collingwood

Bons J, Geniez P (1996) *Amphibiens et reptiles du Maroc (Sahara Occidental compris)*. Atlas boigéographique. Asoc Herpetol Española, Barcelona

Buckland ST, Elston DA (1993) Empirical models for the spatial distribution of wildlife. *J Appl Ecol* 30: 478–495

Chefaoui RM, Hortal J, Lobo JM (2005) Potential distribution modelling, niche characterization and conservation status of Iberian *Copris* species in central Spain. *Biol Conserv* 122: 327–338

El Din SB (2006) *A guide to the reptiles and amphibians of Egypt*. The American University in Cairo Press, Cairo & New York

Freitag S, Van Jaarsveld AS (1995) Towards conserving regional mammal species diversity: a case study and data critique. *S Afr J Zool* 30: 136–144

Gasc JP, Cabela A, Crnobrnja-Isailovic J, Dolmen D, Grossenbacher K, Haffner P, Lescure J, Martens H, Martínez Rica JP, Maurin H, Oliveira ME, Sofianidou TS, Veith M, Zuiderwijk A (eds) (1997) *Atlas of amphibians and reptiles in Europe*. Collection Patrimoines Naturels, 29, Soc. Europaea Herpetol., Muséum Nat d'Histoire Naturelle & Service du Petrimone Nat, Paris

Guisan A, Graham CH, Elith J, Huettmann F, the NCEAS Species Distribution Modelling Group (2007) Sensitivity of predictive species distribution models to change in grain size. *Diversity Distrib* 13: 332–340

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135: 147–186

Hurlbert AH, White EP (2005) Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecol Letters* 8: 319–327

Hurlbert AH, Jetz W (2007) Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proc Natl Acad Sci USA* 104: 13384–13389

IUCN (2007) *European mammal assessment*. http://ec.europa.eu/environment/nature/conservation/species/ema/

Jalas J, Suominen J (1972-1996) *Altas Florae Europaeae*. The Committee for Mapping the Flora of Europe and Soc Biolog Fennica Vanamo

Kudrna O (2002) *The distribution atlas of European butterflies*. Oedippus No. 20

LaSorte FA, Hawkins BA (2007) Range maps and species richness patterns: errors of commission and estimates of uncertainty. *Ecography* 30: 649–662

Lewis A, Pomeroy, D (1989) *A bird atlas of Kenya*. Rotterdam: Balkema

Loffler L, Loffler, P (2005) *Swaziland tree atlas - including selected shrubs and climbers*. Southern African Bot Divers Network Rep No. 38, SABONET, Pretoria

Mathias PBVC, Mendonça CV, Range, TFLVB, Diniz-Filho JAF(2004) Sensitivity of macro-ecological patterns of South American parrots to differences in data sources. *Glob Ecol Biogeogr* 13: 193–198

Martí R, del Moral JC(2003) *Atlas de las aves reproductoras de España*. Dir. Gen. de Conserv. de la Naturaleza & Soc Española de Ornitol, Madrid

Mitchell-Jones AJ, Amori G, Bodgdanowicz W, Krystufek B, Reijnders PJH, Spitzenberger F, Stubbe, M, Thissen, JBM, Vohralík V, Zima J (1999) *The atlas of European mammals*. Academic Press, London

Palmer MW, McGlinn DJ, Fridley JD (2008) Artifacts and artifictions in biodiversity research. *Folia Geobot* (this issue)

Pearson RG, Dawson, TE (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecol Biogeogr* 12: 361–371

Pearson R, Thuiller W, Araújo MB, Martinez E, Brotons L, McClean C, Miles L, Segurado P, Dawson T, Lees D (2006) Model-based uncertainty in species' range prediction. *J Biogeogr* 33: 1704–1711

Peterson AT, Ortega-Huerta MA, Bartley J, Sánchez-Cordero V, Soberón J, Buddemeier RH, Stockwell DRB (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature* 416: 626–629

Prendergast JR, Eversham BC (1995) Butterfly diversity in southern Britain: hotspot losses since 1930. *Biol Conserv* 72: 109–114

Preston CD, Pearman DA, Dines TD (2002) *New atlas of the British and Irish flora*. Oxford University Press, Oxford

Soria-Auza RW, 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 Distrib* 14: 123–130

Thuiller W, Araújo MB, Lavorel S (2003) Generalized models versus classification tree analysis: predicting spatial distributions of plant species at different scales. *J Veg Sci* 14: 669–680

USGS Patuxent Wildlife Research Center (2006): *The North American breeding bird survey, results and analysis 1966-2005. Version 6.2.2006*. USGS Patuxent Wildl Res Center, Laurel

Williams SE, Bolitho EE, Fox S (2003) Climate change in Australian tropical rainforest: an impending environmental catastrophe. *Proc R Soc Lond B* 270: 1887–1892