



Uncertainty and the measurement of terrestrial biodiversity gradients

...what we observe is not nature itself, but nature exposed to our method of questioning

Werner K. Heisenberg (1958, p. 46)

The geographical distribution of biodiversity results from many complex phenomena acting dynamically at multiple scales. Species, populations and individuals are continuously moving, occupying (and evolving to occupy) different parts of the changing geographical and ecological space according to their requirements, limitations and interactions with other species that are also on the move. The aggregation of these dynamic processes is a moving kaleidoscope of patterns that could only be accurately described by the omniscient GIS imagined by Colwell & Coddington (1994), whereby the exact positions of all individuals from all species are always known. Obviously, this cannot exist, so uncertainty becomes an intrinsic part of the study of biodiversity.

In biogeography and macroecology, characteristics of terrestrial assemblages (e.g. species richness) are measured from the list of species present in a number of territorial units, i.e. areas of known location, shape and size (grain and definition of boundaries). Data on species presence can be obtained by either aggregating survey records or overlaying species range maps onto these units (McPherson & Jetz, 2007). Although multiple combinations of presence data and territorial units are possible, studies of biodiversity patterns usually rely on one of only a few data types: (1) *field plots* of limited dimensions sampled with standardized survey techniques (e.g. Schmidt *et al.*, 2008); (2) *local checklists* from sites of known limits and varying areas (e.g. Hortal *et al.*, 2008); (3) *atlas data* obtained from natural history collections and the literature (e.g. Carnicer *et al.*, 2007); and (4) *range map data*, for which polygons of species distribution maps are drawn from field guides, taxonomic reviews or expert opinion, usually using a coarser resolution than atlas data (e.g. Svenning *et al.*, 2008).

Whereas field plots and local checklists are usually built as 'community counts' describing the assemblages present at specific locations, atlas and range map data are usually thought of as estimates of species distributions. Species distribution models are excluded because they are not commonly used to study biodiversity patterns and they present additional problems.

Range map data are currently the most frequent measure of species richness, replacement and diversification at large scales, followed by atlas data from regions with reliable inventories. Range map data provide a wide geographical coverage of information, often at a 1° cell resolution (roughly 100 × 100 km), which is fine enough to provide detail about diversity variations, and coarse enough not to compromise the reliability of the derived biodiversity measures. Hurlbert & Jetz (2007) described a mismatch between richness, measured by range maps, and atlas data. Aggregating these two sources of data at different grains, they showed that, at resolutions < 2°, range maps overestimate the known area of occupancy of most species, giving flawed estimates of species richness patterns (see also Hurlbert & White, 2005), and possibly compromising macroecological and large-scale conservation assessments. They conclude that analysis of range map data might need scales of 2° or higher for less well-known taxa such as amphibians or insects (Hurlbert & Jetz, 2007). Although I share their opinion on the limited value of range map data for conservation planning, I would argue that range map data at finer resolutions, such as 1°, can provide reasonable and reliable measures of important aspects of assemblage diversity if used cautiously, because no data type is inherently more subject to uncertainty than the others, and different measures provide complementary views of biodiversity.

Measures are merely representations of reality, so error is inevitable and each data type incorporates particular sources of error. For example, uncertainty in field plot

data might arise from the limited temporal resolution of the surveys or from the lack of representation of species that are unevenly distributed within the studied land patches. Local checklists, on the other hand, usually come from non-standardized surveys carried out over long periods, so errors can arise from sporadic recording of vagrant individuals of species from outside the locality, and from eventual spurious effects caused by the different areas and shapes of localities and/or any significant bias in geographical location.

Atlas and range map data represent two opposing ways of measuring species distributions: whereas range maps attempt to measure their whole extent, atlas data attempt to measure actual occupancy. Therefore, the sources of uncertainty for these two measuring strategies show a contrast. Range maps exaggerate the distribution of species, giving rise to 'false positives': most species do not inhabit all the grid cells within the polygon of its range map (Hurlbert & White, 2005). Conversely, atlas data involve problems with temporal resolution and representativeness as a result of sampling error and incompleteness, giving incomplete species distributions (false negatives). Most species have not been recorded in some of the grid cells that they actually occupy, and many grid cells have been insufficiently sampled, so atlas data for almost all regions and taxa present broad geographical gaps in knowledge (see Fig. 3 in Hurlbert & Jetz, 2007). This can be ameliorated by assessment of sampling coverage and reliability (Hortal *et al.*, 2007), but such quality control is rarely used.

The perception of biodiversity gradients and their relationship with the environment vary with the scale and type of data used (e.g. McPherson & Jetz, 2007; Hortal *et al.*, 2008). At large scales, biotic interactions, habitat selection and community processes are less important, and richness gradients and species replacement are more the outcome of the environmental limitations and dispersal ability of the group of species

studied. Therefore, range map data can provide a sharp picture of the large-scale dynamics of species ranges, which may be difficult to measure using other data types, for they depict the maximum extent and limits of species distribution without the noise produced by processes acting at finer scales. By contrast, survey-based measurements (field plots, local checklists and atlas data) are useful for studying the effect of ecological processes, identifying the absence of some species from areas where they could have populations but do not as a result of competitive exclusion and/or lack of adequate habitat or populations of their host or prey species.

The choice of data type should be based on the particular facet of diversity or the phenomenon being studied, the spatial scale under study (or the scale at which the particular phenomenon is most important), and the quality of the data. Field plots, local checklists and high-quality fine-grain atlas data (e.g. 10 × 10 km grid cells) show the effects on richness of landscape structure and habitat availability (Schmidt *et al.*, 2008) and of community processes and neutral dynamics (Carnicer *et al.*, 2007). Here, data incompleteness can be compensated for by using species richness estimators (Colwell & Coddington, 1994). Such scales are also appropriate for studying the functional and phylogenetic structure of communities; species replacement occurs across ecological gradients driven by changes in habitat and biotic interactions, resulting in either phylogenetic over-dispersion or clustering (Webb *et al.*, 2002), depending on the strength of competitive interactions and habitat filtering, respectively.

On the other hand, range map and atlas data of relatively coarse grain are most appropriate for studying the effect of climatic gradients and large-scale historical events, although climate, energy and regional constraints have relevance at finer scales (Carnicer *et al.*, 2007; Hortal *et al.*, 2008). In my opinion, grids of 0.5°–1° (roughly 50 × 50 to 100 × 100 km) could be adequate for most studies, provided that the quality of the data is known and the associated errors are taken into account. These grids may provide a neat description

of large-scale biodiversity gradients without including much of the noise produced by local processes. In my opinion, the effects of glaciations (Araújo *et al.*, 2008) or of the diversification at different phylogenetic levels (Svenning *et al.*, 2008) on the current distribution of biodiversity will be easier to depict using homogeneous descriptions of species distribution ranges. Therefore, range map data would be preferable for analysis of the footprint of large-scale historical processes, avoiding the noise produced by local ecological effects and survey biases that is embedded within atlas data.

To summarize, no kind of assemblage data is *a priori* better than another. Rather, different kinds of data and/or scales measure different aspects of biodiversity, and the choice of one or another depends on the purpose of the study in question. In the absence of Colwell & Coddington's omniscient GIS, the key issue for improving the quality of current research in biogeography and macroecology is that, whatever the measure of biodiversity used, associated uncertainty should be taken into account when analysing data and interpreting the results obtained.

ACKNOWLEDGEMENTS

I thank Shai Meiri, Pilar Rodríguez, John Lamshead, David Nogués, Rich Grenyer, Jorge Lobo and Walter Jetz for their reviews and discussion.

JOAQUÍN HORTAL

NERC Centre for Population Biology,
Imperial College London at Silwood Park,
Ascot SL5 7PY, UK.

E-mail: j.hortal@imperial.ac.uk

REFERENCES

Araújo, M.B., Nogués-Bravo, D., Diniz-Filho, J.A.F., Haywood, A.M., Valdes, P.J. & Rahbek, C. (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography*, **31**, 8–15.
Carnicer, J., Brotons, L., Sol, D. & Jordano, P. (2007) Community-based processes behind species richness gradients: contrasting abundance–extinction dynamics

and sampling effects in areas of low and high productivity. *Global Ecology and Biogeography*, **16**, 709–719.

Colwell, R.K. & Coddington, J.A. (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **345**, 101–118.

Heisenberg, W.K. (1958) *Physics and philosophy: the revolution in modern science*. Harper & Row, New York.

Hortal, J., Lobo, J.M. & Jiménez-Valverde, A. (2007) Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife (Canary Islands). *Conservation Biology*, **21**, 853–863.

Hortal, J., Rodríguez, J., Nieto, M. & Lobo, J.M. (2008) Regional and environmental effects on the species richness of mammal assemblages. *Journal of Biogeography*, **35**, 1202–1214.

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.

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.

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.

Schmidt, M.H., Thies, C., Nentwig, W. & Tschamtker, T. (2008) Contrasting responses of arable spiders to the landscape matrix at different spatial scales. *Journal of Biogeography*, **35**, 157–166.

Svenning, J.-C., Borchsenius, F., Bjorholm, S. & Balslev, H. (2008) High tropical net diversification drives the New World latitudinal gradient in palm (Arecaceae) species richness. *Journal of Biogeography*, **35**, 394–406.

Webb, C.O., Ackerly, D.D., McPeck, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Editor: John Lamshead