

# Coherence and discontinuity in the scaling of species' distribution patterns

Stephen Hartley\*, William E. Kunin, Jack J. Lennon†  
and Michael J. O. Pocock‡

*Centre for Biodiversity and Conservation, School of Biology, University of Leeds, Leeds LS2 9JT, UK*

The spatial distribution of a species can be characterized at many different spatial scales, from fine-scale measures of local population density to coarse-scale geographical-range structure. Previous studies have shown a degree of correlation in species' distribution patterns across narrow ranges of scales, making it possible to predict fine-scale properties from coarser-scale distributions. To test the limits of such extrapolation, we have compiled distributional information on 16 species of British plants, at scales ranging across six orders of magnitude in linear resolution (1 m to 100 km). As expected, the correlation between patterns at different spatial scales tends to degrade as the scales become more widely separated. There is, however, an abrupt breakdown in cross-scale correlations across intermediate (*ca.* 0.5 km) scales, suggesting that local and regional patterns are influenced by essentially non-overlapping sets of processes. The scaling discontinuity may also reflect characteristic scales of human land use in Britain, suggesting a novel method for analysing the 'footprint' of humanity on a landscape.

**Keywords:** area of occupancy; fractal; hierarchy; range size; scale–area plot; spatial aggregation

## 1. INTRODUCTION

Understanding the abundances and spatial distributions of species are key goals in ecology (Levin 1992) with implications for the genesis, maintenance and loss of species diversity (Bolker & Pacala 1999; Rosenzweig & Ziv 1999; Hartley & Kunin 2003). However, even to describe these patterns remains a challenge, as different aspects of a species' distribution are manifest at different spatial scales (Pearman 1997; Menendez & Thomas 2000). Coarse-resolution maps reveal a species' geographical range, while progressively finer resolutions reveal regional ubiquity, habitat preferences, population structure, and even population size or cover. A great deal of national and global information on species' distributions is available at relatively coarse scales, but conservation and other management decisions often require relatively fine-scale (e.g. population) information. It would thus be helpful to be able to translate information from one scale to another.

It is a simple matter to generate a coarse-scale map from finer-scale information, but the opposite transition (fine from coarse) has generally been deemed impossible. Nonetheless, Kunin (1998) and He & Gaston (2000) demonstrated that, where species display consistent scaling properties across scales, it is possible to predict fine-scale abundance from information gathered at coarser scales. This is because the number of grid cells occupied at a coarse scale is a function of the number of cells occupied at a finer scale and the degree to which those fine-scale cells are clustered together (the more highly clustered they are, the fewer coarse-scale cells they will occupy).

Turning the relationship around, we can predict the number of cells occupied at a fine scale from the number occupied at a coarse scale, so long as we know the degree of clustering to expect at the fine scale.

Given only a coarse-scale distribution map (e.g. a national atlas), the simplest assumption is that the unknown fine-scale clustering is the same as the observable coarse-scale clustering. Depending upon how clustering is measured, this assumption can be developed as a fractal model of species' distributions (Kunin 1998; Harte *et al.* 1999) or as a negative binomial model with a constant cross-scale aggregation parameter (He & Gaston 2000; Kunin *et al.* 2000). Other statistical models may also be possible. Alternatively, empirical relationships between coarse and fine-scale clustering can be determined and used elsewhere for prediction (Kunin 1998). For example, even if species' distributions were in general twice as aggregated at fine scales as at coarse scales, good fine-scale predictions could be made so long as the relative behaviour of species remained constant: the most aggregated species at one scale being the most aggregated at other scales as well. We may describe such scaling as 'coherent'. Thus, to make fine-scale predictions, we are interested not only in the *absolute* constancy of patterns across scales, but also in the degree of cross-scale *correlation* among species' properties.

To date, such ideas have been examined across only a relatively narrow range of scales. Kunin's (1998) analysis, for example, used 50 km and 10 km linear-resolution data to predict 2 km scale information. But the potential value of such extrapolations (e.g. for conservation decision making) would depend on their utility over much wider ranges of scales. Here, we extend the investigations of Kunin (1998) to see just how far such projections can meaningfully be performed. Specifically we ask: do species' distributions display the same level of clustering at

\* Author and address for correspondence: School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington, New Zealand (stephen.hartley@vuw.ac.nz).

† Present address: Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK.

‡ Present address: School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK.

Table 1. National grid-cell occupancy of the 16 study species observed at two different grid resolutions.

(Fractal dimensions (national and local) are calculated over a 10-fold range of scales, see § 2c for details. The two fractal dimensions are uncorrelated (Pearson's  $r = -0.14$ , d.f. = 14,  $p > 0.5$ ).)

family	species	national grid-cell occupancy		fractal dimensions	
		100 × 100 km	10 × 10 km	(national) $D_{10-100}$	(local) $D_{0.001-0.01}$
(Pteridophyte) <sup>a</sup>	<i>Dryopteris submontana</i>	2	13	0.83	0.58
	<i>Thelypteris palustris</i>	20	57	0.44	1.53
Brassicaceae	<i>Iberis amara</i>	4	18	0.68	1.32
	<i>Arabis glabra</i>	7	17	0.42	0.66
Campanulaceae	<i>Phyteuma orbiculare</i>	4	27	0.81	1.42
	<i>Lobelia urens</i>	4	6	0.15	1.19
Caryophyllaceae	<i>Silene otites</i>	1	5	0.77	1.13
	<i>Dianthus armeria</i>	12	20	0.20	0.61
Fabaceae	<i>Genista pilosa</i>	4	11	0.43	1.31
	<i>Lathyrus palustris</i>	9	14	0.17	1.53
Lamiaceae	<i>Clinopodium calamintha</i>	7	37	0.72	0.94
	<i>Mentha pulegium</i>	7	14	0.26	1.01
Orchidaceae	<i>Orchis purpurea</i>	3	15	0.65	0.94
	<i>Himantoglossum hircinum</i>	7	16	0.39	0.63
Orobanchaceae	<i>Orobanche reticulata</i>	1	7	0.83	0.42
	<i>Orobanche purpurea</i>	8	13	0.23	0.81

<sup>a</sup> Order: Pteridophyte; families: Dryopteraceae and Thelypteraceae, respectively.

all scales, or are there gradual shifts in scaling properties? Furthermore, if scaling properties alter, do they change in predictable (coherent) ways, or do species shift relative to one another? Finally, if species do shift in relative scaling properties, do they do so gradually, or are there particular scales of abrupt transition?

To answer these questions we require detailed distributional information for a range of contrasting species across a wide range of spatial scales. In this study we present the results of the first (to our knowledge) multi-species investigation to collect data over a sufficiently wide range of ecologically relevant scales, i.e. from the scale of individuals to the scale of national range maps.

## 2. METHODS

### (a) National atlas data

The distribution patterns of 16 British plant species (table 1) were assessed at 16 spatial resolutions spanning six orders of magnitude, from 1 m to 100 km linear resolution. The species chosen are all 'rare' or 'scarce' in Britain (defined as species occupying 15 or fewer, or 100 or fewer cells of a national 10 km resolution grid, respectively), as these have the best documented distributions (Stewart *et al.* 1994; Wigginton 1999; Preston *et al.* 2002). The species consist of eight taxonomically linked pairs, representing a variety of life-history strategies and habitat requirements, with one species of each pair having a tightly clumped national distribution and the other having a nationally scattered distribution (as observed at a 10–100 km resolution). This design allows a phylogenetically controlled test of whether national-scale patterns are correlated with local-scale patterns.

National distributions were determined at 1 km resolution using data supplied by the Biological Records Centre (BRC) at Monks Wood, UK. These data were then 'blocked-up' to create distribution maps at 2, 5, 10, 20, 50 and 100 km linear resolutions. In a few cases (16% of records), no 1 km resolution

information had been recorded for some occupied 2 km cells ('tetrads'). In such cases the total number of 1 km cells was estimated by assuming that the tetrads lacking 1 km occupancy data were similar to the tetrads of known occupancy. Thus for a species with, for example, nine sites recorded to 1 km precision within six tetrads and an additional occupied tetrad lacked 1 km precision data, the additional tetrad was assumed to contain 1.5 (nine out of six) occupied 1 km sub-cells. The same logic was applied to 10 km cells with missing 5 km precision records (as not all tetrad records can be mapped precisely into a 5 km cell).

### (b) Field survey

To determine patterns of distribution at scales finer than 1 km, four occupied 1 km cells were selected for each species in a stratified random manner. The stratification was achieved by dividing the set of occupied 1 km cells into a 'core' half and a 'marginal' half according to the number of other occupied cells in their neighbourhood. The radius of their neighbourhoods was different for each species, and was chosen to be the radius that gave the maximum amount of information for creating a ranking (information being measured by the Shannon index,  $-\sum p_i \ln(p_i)$ , where  $p_i$  is the proportion of sites with 0, 1, 2, 3... other sites within the specified neighbourhood radius). Two 'core' and two 'marginal' cells were selected at random for each species. Each square kilometre was surveyed in the field on a hectare by hectare basis (i.e. presence and absence was recorded on a 100 m resolution grid) for a period of four person days. When an occupied hectare was discovered, a subjective estimate of population size within the grid cell was made, which allowed the set of occupied hectares to be ranked by approximate species abundance. Two occupied hectares within each surveyed kilometre were then randomly selected as a stratified sub-sample, one each from the high- and the low-density halves of the list, and surveyed at 10 m resolution. Finally, two occupied 10 m grid cells from each of the surveyed hectares were sub-sampled (again in a stratified manner based on estimated populations)

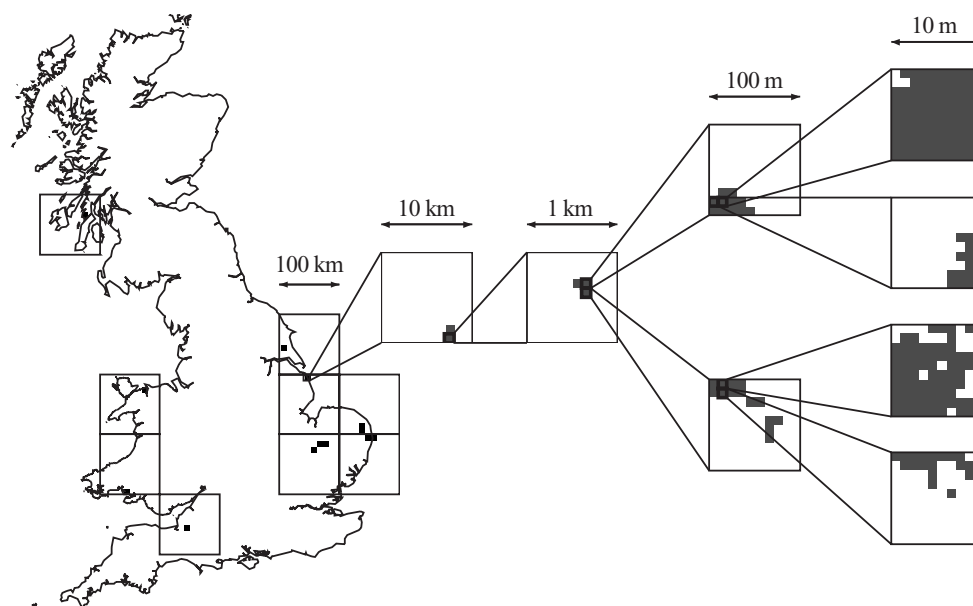


Figure 1. Distribution of the marsh pea, *Lathyrus palustris* (one of the 16 species surveyed), viewed at six different resolutions. The national distribution was constructed from all records held by the BRC, made between 1987 and 1999. This was supplemented by a field survey of four sample 1 km  $\times$  1 km squares mapped at 100 m resolution (one of which is illustrated above), eight hectares (100 m  $\times$  100 m) mapped at 10 m resolution (two of which are illustrated) and 15 10 m  $\times$  10 m quadrats mapped at 1 m resolution (four of which are illustrated).

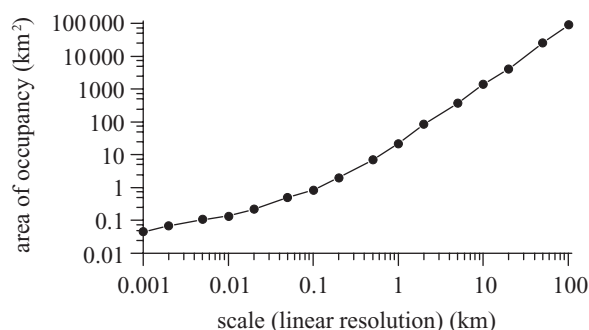


Figure 2. Scale-area plot for the British distribution of *Lathyrus palustris*.

and mapped at 1 m resolution or higher (figure 1). Of course, in cases where only one hectare or one 10 m square was occupied within the larger 10  $\times$  10 grid, only this one square could be sub-sampled at a finer resolution.

### (c) Single-species scaling relationships

The field survey data were also blocked-up to intermediate resolutions (e.g. 2 m, 5 m) and combined with the BRC data to estimate the total number of grid cells occupied nationally at each of 16 different scales ranging from 1 m to 100 km linear resolution (i.e. using square grid cells from 1 m<sup>2</sup> to 10 000 km<sup>2</sup> in area). Multiplying the number of grid cells occupied at each scale by the size of the cell produces an estimate for the 'area of occupancy' (range size) observable at each scale (Gaston 1991). For each species, we then plotted area of occupancy as a function of linear resolution (scale) to produce a so-called scale-area curve (e.g. figure 2, see also Kunin (1998)). (When resolution is measured in units of area, the term 'range-area relationship' (Green *et al.* 2003; Ostling *et al.* 2003) may be preferable, because it emphasizes the fundamental connection with the species-area relationship (SAR).)

From the changing slopes of each scale-area plot we computed 15 scale-specific box-counting fractal dimensions ( $D_{ij}$ ) for each species' distribution (Kunin 1998). This local fractal dimension characterizes the spatial distribution of a species between scales  $i$  and  $j$  (the fine and coarse scales, respectively). A maximum value of  $D_{ij} = 2$  occurs when the occupied fine-scale cells are clustered together so as to fill each of the occupied coarse-scale cells completely, while a minimum value of  $D_{ij} = 0$  indicates that each fine-scale cell falls in a separate coarse-scale cell. For some analyses we calculated a more generalized fractal dimension applicable to a 10-fold range of scales (e.g. 1–10 km) by linear regression of  $\log(\text{occupancy})$  versus  $\log(\text{scale})$ . In either case,  $D_{ij} = 2 - b_{ij}$ , where  $b_{ij}$  is the regression coefficient or slope of the regression between scales  $i$  and  $j$  inclusive.

### (d) Correlation and predictability

The matrix of 15 scale-specific fractal dimensions (rows) by 16 species (columns) was analysed using cluster analysis with the Pearson correlation coefficient as the measure of similarity between rows (Everitt 1993). Consistent groupings of correlated scales were identified by comparing the dendrograms generated by (i) average linkage, (ii) median and (iii) centroid-clustering algorithms (SPSS Inc. 1999).

To test the limits of predictability explicitly we applied a fractal model of species' distributions (which assumes a constant  $D_{ij}$  for all scales) by performing a linear extrapolation of  $\log(\text{occupancy})$  versus  $\log(\text{scale})$  using ordinary least-squares regression (Kunin 1998; Kunin *et al.* 2000). The slope and intercept of the regression were parameterized using data from a 10-fold range of scales (as described in § 2c) to predict occupancy at all finer scales for which test data were available. In comparing 'observed' and 'predicted'  $\log(\text{occupancy})$ , the 16 species were treated as independent replicates for each possible pairing of the predictor and predicted scales.

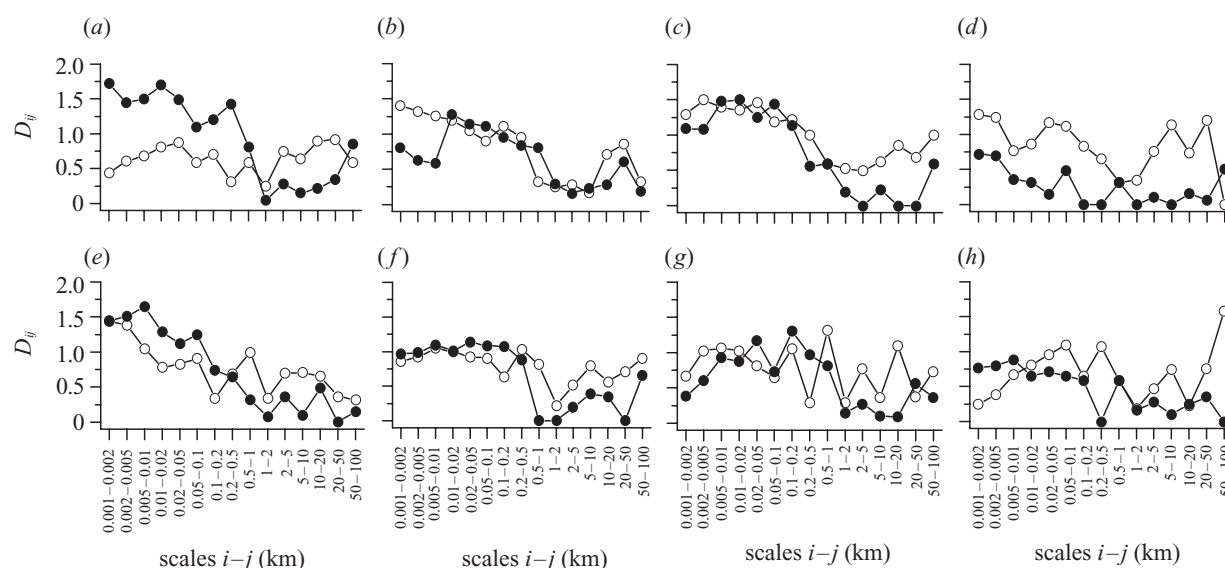


Figure 3. Scale-specific box-counting fractal dimensions ( $D_{ij}$ ) of the 16 species' distributions, paired by family. (a) Open circles, *Dryopteris submontana*; closed circles, *Thelypteris palustris*. (b) Open circles, *Iberis amara*; closed circles, *Arabis glabra*. (c) Open circles, *Phyteuma orbiculare*; closed circles, *Lobelia urens*. (d) Open circles, *Silene otites*; closed circles, *Dianthus armeria*. (e) Open circles, *Genista pilosa*; closed circles, *Lathyrus palustris*. (f) Open circles, *Clinopodium calamintha*; closed circles, *Mentha pulegium*. (g) Open circles, *Orchis purpurea*; closed circles, *Himantoglossum hircinum*. (h) Open circles, *Orobancha reticulata*; closed circles, *Orobancha purpurea*.

### 3. RESULTS

#### (a) Absolute values of $D_{ij}$ across scales

Observed  $D$ -values were not constant across the full range of scales, as predicted for truly fractal distributions, but rather showed consistent trends (figure 3). In general, the lowest  $D$ -values (that is, the most diffuse distributions) were found at intermediate resolutions, especially *ca.* 1–2 km scales (figure 3; table 2c).

Overall, fine-scale (1–100 m) species' distributions showed consistently higher  $D$ -values (more clustered and space-filling distributions) than those found at coarse scales (2–100 km). This may be a general phenomenon, or may simply reflect the fact that our sample was drawn from rare and scarce species, which by virtue of their low occupancy at the 10 km scale will tend to have relatively low national-scale  $D$ -values (for related discussion see Haskell *et al.* (2002)).

#### (b) Correlations between scales

The results of our correlation analysis indicate both predictable and unpredictable cross-scale behaviour. Within two broad scaling domains (1–500 m and 1–50 km), the fractal dimensions of species' distributions, though not necessarily constant, vary in a coherent manner, as shown by the large number of significant cross-scale correlations (table 2a) and the two large groupings revealed by a cluster analysis of the correlation matrix (table 2b). (Out of the 105 correlations presented in table 2a, 26 are significant at  $p \leq 0.05$ , 16 at  $p \leq 0.01$  and nine at  $p \leq 0.001$ .) However, between these two domains there is an abrupt discontinuity in scaling behaviour, with near-zero or even negative correlations. In other words, the pattern of a species' distribution (i.e. whether it is aggregated or diffuse at a particular scale relative to other species) remains fairly consistent within each of the two scaling domains, but the rank order of species' patterns changes considerably as

one crosses from one domain to the other. Hence, the spatial pattern of a population at a 1 km or coarser resolution has little bearing on the species' relative distributional properties at 0.2 km or finer scales. A phylogenetically controlled analysis of contrasts between species pairs reveals the same qualitative correlation structure (see electronic Appendix A, available on The Royal Society's Publications Web site).

The breakdown in cross-scale correlations creates a similar split in predictive power (figure 4). Coarse-scale (e.g. 10–100 km) information can be used to predict occupancy at other fairly coarse scales (e.g. 1 km resolution) with reasonable accuracy, but the quality of the prediction degrades rapidly when multi-kilometre data are used to predict sub-kilometre occupancy. However, starting at scales finer than 200 m, long-range extrapolations to yet finer scales regain their predictive power. Qualitatively similar results (available upon request) were obtained by extrapolating a negative binomial model of species' distributions (for details of method see He & Gaston (2000) and Kunin *et al.* (2000)).

The sudden drop in mid-range predictability is unlikely to be caused by the differences in data-collection methods between local and regional scales, because correlations across species will be unaffected by any systematic scale-dependent biases in the observed  $D$ . Moreover, the lowest correlation in the leading diagonal of table 2a, which occurs between  $D_{0.2-0.5}$  and  $D_{0.5-1}$ , uses data that were all collected by the same method (the 100 m resolution field survey).

### 4. DISCUSSION

The box-counting fractal dimension ignores information concerning the density of points (number of records) within a grid cell; hence, as a metric of spatial



Table 2. (a) Pearson correlation coefficients between the  $D_{ij}$  values of the 16 species. Single-test significance values are as follows: \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ . There are numerous strong intercorrelations between local-scale patterns ( $D_{0.001-0.002}$  to  $D_{0.2-0.5}$ ) and between regional-scale patterns ( $D_{1-2}$  to  $D_{20-50}$ ), but  $D_{0.5-1}$  does not correlate strongly with any other  $D$ -value. A phylogenetically controlled test based upon the eight species-pair comparisons yields the same qualitative result (see electronic Appendix A). (b) Four-group membership obtained from cluster analysis (CA), using correlation coefficients as a measure of similarity. (c) Mean and coefficient of variation (CV) of  $D_{ij}$  ( $n = 16$ ).

(a)		0.001–0.002	0.002–0.005	0.005–0.01	0.01–0.02	0.02–0.05	0.05–0.1	0.1–0.2	0.2–0.5	0.5–1	1–2	2–5	5–10	10–20	20–50	50–100
scale $i-j$ (km)																
0.002–0.005	0.9***															
0.005–0.01	0.7***	0.8***														
0.01–0.02	0.6*	0.5*	0.8***													
0.02–0.05	0.4	0.5	0.7**	0.9***												
0.05–0.1	0.5	0.4	0.6*	0.7**	0.8***											
0.1–0.2	0.2	0.3	0.6*	0.7***	0.8***	0.5										
0.2–0.5	0.3	0.3	0.4	0.6**	0.8***	0.6*	0.6*									
0.5–1	–0.2	–0.1	0.0	0.1	0.0	–0.2	0.1	0.0								
1–2	0.1	0.2	0.0	0.1	0.3	0.1	0.2	0.1	0.0	0.4						
2–5	0.0	0.1	–0.1	–0.2	0.0	–0.2	–0.1	–0.1	–0.1	0.4	0.6*					
5–10	0.0	0.0	–0.2	–0.1	0.2	0.2	–0.1	0.2	0.2	0.1	0.6*	0.7**	regional domain			
10–20	0.1	0.3	0.1	0.0	0.0	–0.2	0.1	–0.1	–0.1	0.3	0.6**	0.8***	0.5			
20–50	–0.1	–0.1	–0.3	–0.1	–0.1	–0.1	0.1	0.1	0.2	0.1	0.6*	0.6*	0.7**	0.5		
50–100	–0.3	–0.2	0.0	0.1	0.1	0.1	0.1	0.1	0.4	0.2	0.0	0.1	0.2	0.0	0.0	
(b)																
CA grouping	$a$	$a$	$a$	$a$	$a$	$a$	$a$	$a$	$a$	$b$	$c$	$c$	$c$	$c$	$c$	$d$
(c)																
mean $D_{ij}$	0.97	1.01	1.02	1.03	1.01	0.95	0.85	0.71	0.71	0.61	0.21	0.4	0.4	0.47	0.49	0.55
CV	0.45	0.36	0.36	0.33	0.32	0.28	0.42	0.56	0.56	0.52	0.67	0.62	0.81	0.69	0.75	0.76

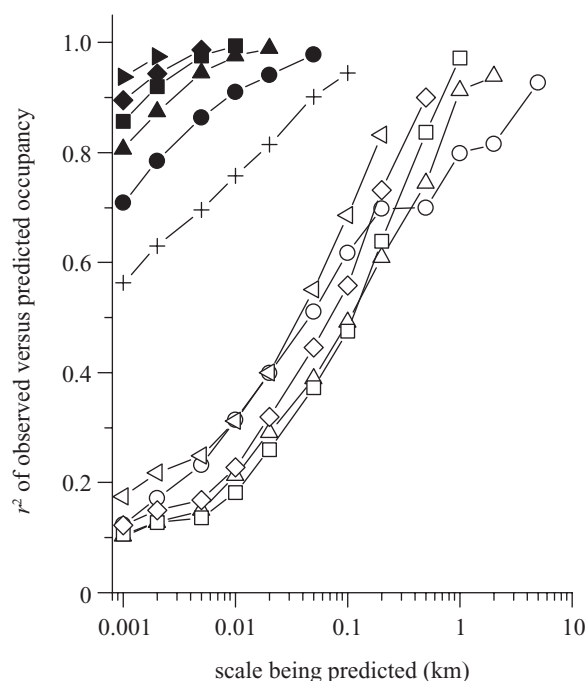


Figure 4. Predictive value of extrapolations (assuming a constant  $D_{ij}$ ) using data from four coarser scales to predict occupancy at a range of finer scales. The quality of each prediction is represented by the  $r^2$  value of a linear regression between  $\log(\text{predicted})$  and  $\log(\text{observed})$  cells occupied at each scale. Predictive accuracy invariably degrades as the range of the prediction increases (further left along each line). Extrapolations from multi-kilometre data (open symbols) produce fairly accurate predictions for the first few jumps in scale, but have little predictive value when applied to sub-500 m scales. However, predictions made using data from 200 m and below (solid symbols) retain high predictive value over wide ranges of finer scales. Range of predictor scales: open circles, 10–100 km; open upright triangles, 5–50 km; open squares, 2–20 km; open diamonds, 1–10 km; open side triangles, 0.5–5 km; crosses, 0.2–2 km; closed circles, 0.1–1 km; closed upright triangles, 0.05–0.5 km; closed squares, 0.02–0.2 km; closed diamonds, 0.01–0.1 km; closed side triangles, 0.005–0.05 km.

pattern it is most appropriate for the analysis of presence-absence data. Had we been analysing surfaces (e.g. density or abundance) then multifractal or spectral analysis would generally have been more suitable (Keitt 2000; Borda-de-Água *et al.* 2002). Application of our method did not require that species' distributions were true mathematical fractals (statistically self-similar across infinite scales) because we fitted only a localized straight line across two, or at the most four, adjacent scales. Indeed, over the full range of scales examined, most species' distributions required at least two straight lines or a polynomial to describe their scale-area relationship adequately.

Most species displayed very diffuse distributions (low fractal dimensions) around the 1–2 km scale, which is also the scale at which Crawley & Harral (2001) found the steepest SAR in their analysis of British plant distributions. As the SAR is essentially the sum of each species' scale-specific incidence probabilities (Harte *et al.* 2001; Lennon *et al.* 2002; Green *et al.* 2003), their results suggest that the patterns documented here may apply to a much wider set of species. An important caveat, however,

is that both Crawley & Harral's (2001) and our own data are liable to an increased risk of false negatives at the 1 km scale. If these false negatives are randomly distributed then they will depress the fractal dimension (or raise the  $z$ -value of an SAR) at that particular scale (Lennon *et al.* 2002; W. R. Turner, unpublished data).

If systematic biases cast doubt on the absolute  $D$ -values recorded, we may nonetheless be able to learn from overall patterns of relative values exhibited by this diverse set of plant species across a wide range of scales. One should always be cautious when trying to infer process from pattern (Molofsky *et al.* 2002); nonetheless, if a single process (e.g. dispersal) governed pattern across all scales, one would expect certain species (e.g. the poorest dispersers) to be consistently the most clumped at all scales. If, as is likely, the set of important processes were to shift gradually across scales, one should expect the gradual diminution in cross-scale correlations that we observed within the two scaling domains. An abrupt discontinuity, though, suggests a rapid change in the dominant process across that scale for all (or most) of the species studied (Turner *et al.* 1989; Wiens 1989).

Following the above logic, the two broad domains of coherent scaling documented here imply that there are coherent sets of processes influencing species' distributions across wide ranges of spatial scales, but the discontinuity at intermediate scales suggests that local and national-scale distribution patterns are dominated by different and essentially non-overlapping sets of processes. Within domains, models of pattern and process based on continuous scaling functions may prove useful (Kunin 1998; He & Gaston 2000), but to move between domains, a two-tier hierarchical model may be more appropriate (cf. Milne 1988; Kolasa 1989; O'Neill *et al.* 1989; Collins & Glenn 1990, 1991; Whittaker *et al.* 2001).

Recently, there has been considerable debate as to whether the SAR is best described by a single function (usually a power law) or as a set of three distinct phases, each of a different slope and each dominated by different processes (Rosenzweig & Ziv 1999; Lomolino 2000). At fine scales sampling effects and competition for limited space are presumed to be the dominant processes shaping the relationship (at least for plants); at intermediate scales habitat turnover is the proximate explanation, while at continental scales processes of speciation and extinction dominate (Crawley & Harral 2001). It is likely that our analysis of British plant distributions spanned only the first two domains.

It is interesting to note that the sudden drop in predictable scaling occurs at scales (0.2–1 km) that roughly correspond to the typical 'grain' of human land use in Britain. (The median linear dimension of an arable field in England is *ca.* 0.16 km (interquartile range of 0.10–0.23 km) and the median size of an entire farm holding is 0.49 km (interquartile range of 0.25–0.87 km) (Department for Environment, Farming and Rural Affairs, unpublished data).) This probably represents the range of scales at which the footprint of humanity is strongest, disrupting whatever processes and patterns might once have naturally occurred. Human land use is not yet complete enough to remove pattern at coarser scales, as small fragments of natural habitat generally remain within most multi-kilometre cells; whereas at finer

scales, within the remaining habitat fragments, species are free to exhibit their inherent spacing patterns. Hence, the disruption of natural scaling may provide a novel method of quantifying the human impact on natural landscapes.

Further work is necessary to discover whether different taxa, or different landscapes, reveal markedly different scaling of species' distribution patterns. One potential comparison is with the work of Storch *et al.* (2002), who examined the power spectra of avian community composition and environmental heterogeneity along a 50 km transect in Bohemia. After factoring out the effect of habitat heterogeneity, they observed a peak in the power spectrum of community composition at around the 2–7 km scale, which they suggested may be related to the typical distance of breeding-related dispersal in birds.

As an example of how land use may have influenced distribution patterns in our study, consider the two Lamiaceae. Much of the *Clinopodium calamintha* we surveyed was found growing in stretches of anthropogenic habitat, such as roadside verges and field boundaries, while *Mentha pulegium* was often found growing along the sides of seasonal streams and the margins of water bodies; these types of habitat are essentially linear over short distances, and both species exhibited fine-scale fractal dimensions close to one (table 1, figure 3). By way of contrast, *Thelypteris palustris* and *Lathyrus palustris* are both restricted to nationally sparse marshland habitat (often nature reserves surrounded by drained land) yet within this habitat their fine-scale distributions are typically space filling, with a fractal dimension close to 1.5 (table 1, figure 3). It would be relatively easy to go through all of the species in our study and provide *post hoc* explanations for the fractal dimensions observed at different scales; however, to be rigorous such an analysis would require a larger sample set coupled with systematic life-history-trait data. We are currently working on such an analysis of British plant species for the scales 1–100 km.

Whatever their origin, empirical scaling functions are a potentially valuable tool for conservation biologists, as they allow useful predictions of species distributions to be made for scales at which it would be impractical or too costly to survey in the field (Kunin 1998). However, the diminution of correlations with increasing lag and the presence of a scaling discontinuity warn that there will be limitations to this form of predictive spatial ecology. Ultimately, a more process-orientated understanding of the factors that determine species' distributions may be required to extrapolate across the full range of scales. Identifying the relevant processes and their importance at different scales remain important challenges for the future.

This work was supported by a grant from the Natural Environment Research Council of the UK (GR3/11916). Field surveys were conducted with the assistance of Hazel Aitken, Rebecca Aston, James Bunyan, Georges Chekuimo, Ruth Clark, Kara Clarkson, Lucia Galvez, Duncan Gray, Alison McCrea, Sarah Mellor, Karen Shelly, Samantha Staddon and Jonathon Willot. Sadly, Hazel did not live to see this work published. We express our thanks to the Botanical Society of the British Isles, Plantlife and the Biological Records Centre at the Centre for Ecology and Hydrology for national-scale distribution data, to Steve Langton and the Department for Environment, Farming and Rural Affairs for statistics on land use, to Pete Carey for fine-scale maps of *Himantoglossum hircinum* and to numerous private and public land owners for access to field sites. Comments

from Chris D. Thomas, Phil J. Lester and anonymous referees were helpful in the preparation of this manuscript.

## REFERENCES

- Bolker, B. M. & Pacala, S. 1999 Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal. *Am. Nat.* **153**, 575–602.
- Borda-de-Água, L., Hubbell, S. P. & McAllister, M. 2002 Species–area curves, diversity indices, and species abundance distributions: a multifractal analysis. *Am. Nat.* **159**, 138–155.
- Collins, S. L. & Glenn, S. M. 1990 A hierarchical analysis of species' abundance patterns in grassland vegetation. *Am. Nat.* **135**, 633–648.
- Collins, S. L. & Glenn, S. M. 1991 Importance of spatial and temporal dynamics in species regional abundance and distribution. *Ecology* **72**, 654–664.
- Crawley, M. J. & Harral, J. E. 2001 Scale dependence in plant biodiversity. *Science* **291**, 864–868.
- Everitt, B. S. 1993 *Cluster analysis*, 3rd edn. London: Edward Arnold.
- Gaston, K. J. 1991 How large is a species' geographic range? *Oikos* **61**, 434–438.
- Green, J. L., Harte, J. & Ostling, A. 2003 Species richness, endemism and abundance patterns: tests of two fractal models in a serpentine grassland. *Ecol. Lett.* **6**, 919–928. (DOI 10.1046/j.1461-0248.2003.00519.x.)
- Harte, J., Kinzig, A. & Green, J. 1999 Self-similarity in the distribution and abundance of species. *Science* **284**, 334–336.
- Harte, J., Blackburn, T. & Ostling, A. 2001 Self-similarity and the relationship between abundance and range size. *Am. Nat.* **157**, 374–386.
- Hartley, S. & Kunin, W. E. 2003 Scale-dependency of rarity, extinction risk, and conservation priority. *Conserv. Biol.* **17**, 1559–1570.
- Haskell, J. P., Ritchie, M. E. & Olff, H. 2002 Fractal geometry predicts varying body size scaling relationships for mammal and bird home ranges. *Nature* **418**, 527–530.
- He, F. & Gaston, K. J. 2000 Estimating species abundance from occurrence. *Am. Nat.* **156**, 553–559.
- Keitt, T. H. 2000 Spectral representation of neutral landscapes. *Landscape Ecol.* **15**, 479–493.
- Kolasa, J. 1989 Ecological systems in hierarchical perspective: breaks in community structure and other consequences. *Ecology* **70**, 36–47.
- Kunin, W. E. 1998 Extrapolating species abundance across spatial scales. *Science* **281**, 1513–1515.
- Kunin, W. E., Hartley, S. & Lennon, J. J. 2000 Scaling down: on the challenge of estimating abundance from occurrence patterns. *Am. Nat.* **156**, 560–566.
- Lennon, J. J., Kunin, W. E. & Hartley, S. 2002 Fractal species distributions do not produce power-law species–area relationships. *Oikos* **97**, 378–386. (DOI 10.1034/j.1600-0706.2002.970308.x.)
- Levin, S. A. 1992 The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967.
- Lomolino, M. V. 2000 Ecology's most general, yet protean pattern: the species–area relationship. *J. Biogeogr.* **27**, 17–26. (DOI 10.1046/j.1365-2699.2000.00377.x.)
- Menendez, R. & Thomas, C. D. 2000 Metapopulation structure depends on spatial scale in the host-specific moth *Wheeleria spilodactylus* (Lepidoptera: Pterophoridae). *J. Anim. Ecol.* **69**, 935–951. (DOI 10.1046/j.1365-2656.2000.00449.x.)

- Milne, B. T. 1988 Measuring the fractal geometry of landscapes. *Appl. Math. Comput.* **27**, 67–79.
- Molofsky, J., Bever, J. D., Antonovics, J. & Newman, T. J. 2002 Negative frequency dependence and the importance of spatial scale. *Ecology* **83**, 21–27.
- O'Neill, R. V., Johnson, A. R. & King, A. W. 1989 A hierarchical framework for the analysis of scale. *Landscape Ecol.* **3**, 193–205.
- Ostling, A., Harte, J., Green, J. L. & Kinzig, A. P. 2003 A community-level fractal property produces power-law species–area relationships. *Oikos* **17**, 218–224. (DOI 10.1034/j.1600-0706.2003.12600.x.)
- Pearman, D. 1997 Towards a new definition of rare and scarce plants. *Watsonia* **21**, 231–251.
- Preston, C. D., Pearman, D. A. & Dines, T. D. 2002 *New atlas of the British and Irish flora*. Oxford University Press.
- Rosenzweig, M. L. & Ziv, Y. 1999 The echo pattern of species diversity: pattern and processes. *Ecography* **22**, 614–628.
- SPSS Inc. 1999 *SPSS base 10.0 applications guide*. Chicago, IL: SPSS Inc.
- Stewart, A., Pearman, D. A. & Preston, C. D. (eds) 1994 *Scarce plants in Britain*. Peterborough, UK: Joint Nature Conservation Committee.
- Storch, D., Gaston, K. J. & Cepák, J. 2002 Pink landscapes: 1/f spectra of spatial environmental variability and bird community composition. *Proc. R. Soc. Lond. B* **269**, 1791–1796. (DOI 10.1098/rspb.2002.2076.)
- Turner, M. G., Dale, V. H. & Gardner, R. H. 1989 Predicting across scales: theory development and testing. *Landscape Ecol.* **3**, 245–252.
- Whittaker, R. J., Willis, K. J. & Field, R. 2001 Scale and species richness: towards a general, hierarchical theory of species diversity. *J. Biogeogr.* **28**, 453–470. (DOI 10.1046/j.1365-2699.2001.00563.x.)
- Wiens, J. A. 1989 Spatial scaling in ecology. *Funct. Ecol.* **3**, 383–397.
- Wigginton, M. J. (ed.) 1999 *British red data books: vol. 1. Vascular plants*. Peterborough, UK: Joint Nature Conservation Committee.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

Visit <http://www.pubs.royalsoc.ac.uk> to see an electronic appendix to this paper.