

Article Type: Article

Upscaling biodiversity: estimating the Species-Area Relationship from small samples

Authors:

William E. Kunin^{1,2,*}, John Harte³, Fangliang He⁴, Cang Hui⁵, R. Todd Jobe^{6,7}, Annette Ostling⁸, Chiara Polce^{1, 18}, Arnošt Šizling⁹, Adam B. Smith^{3,10}, Krister Smith¹¹, Simon M. Smart¹², David Storch^{9,13}, Even Tjørve¹⁴, Karl-Inne Ugland¹⁵, Werner Ulrich¹⁶, Varun Varma^{1,17}

* Corresponding author. E-mail: w.e.kunin@leeds.ac.uk

Addresses:

1. Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK, and
2. Stellenbosch Institute for Advanced Studies (STIAS), Wallenberg Research Centre at Stellenbosch University, Stellenbosch 7600, South Africa
3. Energy and Resources Group and Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA
4. Department of Renewable Resources, University of Alberta, Edmonton, Alberta T6G 2H1, Canada
5. Centre for Invasion Biology, Department of Mathematical Sciences, Stellenbosch University, and African Institute for Mathematical Sciences, Stellenbosch 7600, South Africa
6. Department of Geography, University of North Carolina, Chapel Hill, NC 27599-3220 USA
7. Current address: Signal Innovations Group, Inc., 4721 Emperor Blvd., Suite 330 Durham, NC 27703, USA
8. Ecology & Evolutionary Biology, University of Michigan, Ann Arbor MI, USA
9. Center for Theoretical Study, Charles University and the Academy of Sciences of the Czech Republic, Jilská 1, 110 00 Praha 1, Czech Republic
10. Centre for Conservation and Sustainable Development, Missouri Botanical Garden, 4344 Shaw Boulevard, St. Louis, MO 63110, USA

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/ecm.1284](https://doi.org/10.1002/ecm.1284)

This article is protected by copyright. All rights reserved

11. Senckenberg Research Institute and Natural History Museum, Frankfurt am Main, Germany
12. NERC Centre for Ecology and Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP UK
13. Department of Ecology, Faculty of Science, Charles University Viničná 7, 128 44 Praha 2, Czech Republic
14. Lillehammer University College, PO Box 952, NO-2604, Lillehammer, Norway
15. Department of Biology, University of Oslo, PB 1064 Blindern, 0316 Oslo, Norway
16. Faculty of Biology and Environmental Protection, Nicolaus Copernicus University, Lwowska 1, 87-100 Toruń, Poland
17. Current address: Department of Biosciences, University of Exeter, Exeter EX4 4QD, UK
18. Current affiliation: European Commission, Joint Research Centre, 21027 Ispra (VA), Italy

Running head: Upscaling biodiversity

1 **Abstract**

2

3 The challenge of biodiversity upscaling, estimating the species richness of a large area from scattered
4 local surveys within it, has attracted increasing interest in recent years, producing a wide range of
5 competing approaches. Such methods, if successful, could have important applications to multi-scale
6 biodiversity estimation and monitoring. Here we test 19 techniques using a high quality plant dataset:
7 the GB Countryside Survey 1999, detailed surveys of a stratified random sample of British landscapes.
8 In addition to the full dataset, a set of geographical and statistical subsets was created, allowing each
9 method to be tested on multiple datasets with different characteristics. The predictions of the models
10 were tested against the “true” species-area relationship for British plants, derived from
11 contemporaneously surveyed national atlas data. This represents a far more ambitious test than is
This article is protected by copyright. All rights reserved

12 usually employed, requiring 5 – 10 orders of magnitude in upscaling. The methods differed greatly in
13 their performance; while there are 2326 focal plant taxa recorded in the focal region, up-scaled species
14 richness estimates ranged from 62 to 11,593. Several models provided reasonably reliable results across
15 the 16 test datasets: the Shen and He and the Ulrich and Ollik models provided the most robust
16 estimates of total species richness, with the former generally providing estimates within 10% of the true
17 value. The methods tested proved less accurate at estimating the shape of the Species-Area Relationship
18 as a whole; the best single method was Hui's Occupancy Rank Curve approach, which erred on average
19 by <20%. A hybrid method combining a total species richness estimate (from the Shen and He model)
20 with a downscaling approach (the Šizling model) proved more accurate in predicting the SAR (mean
21 relative error 15.5%) than any of the pure upscaling approaches tested. There remains substantial room
22 for improvement in upscaling methods, but our results suggest that several existing methods have a high
23 potential for practical application to estimating species richness at coarse spatial scales. The methods
24 should greatly facilitate biodiversity estimation in poorly-studied taxa and regions, and the monitoring
25 of biodiversity change at multiple spatial scales.

26
27 **Keywords:** spatial scale; Species-Area Relationship; species richness; biodiversity estimation;
28 monitoring; upscaling; methods comparison

31 **Introduction**

32
33 Biological diversity is intrinsically scale-dependent. While the issue of spatial scaling has only recently
34 become prominent in many other areas of scientific research, the appreciation of scale issues in
35 biodiversity research dates back to the foundations of the discipline. The most widely-used tool for
36 describing biodiversity scaling remains the Species-Area Relationship (SAR), first devised more than a
37 century ago (Watson 1835, Arrhenius 1921, Gleason 1922). The SAR represents species richness
38 explicitly as a function of sample area, which is to say, as a function of spatial scale. The scale-
39 dependence of biodiversity as reflected in the SAR represents the combined effects of statistical
40 sampling and ecological processes. As one examines communities across ever wider expanses, the
41 number of species inevitably rises for a number of reasons: larger samples incorporate more individuals
42 (allowing more species to be sampled), they encompass a wider range of habitats and environmental
43 conditions, and bridge barriers to dispersal (Shmida and Wilson 1985, Drakare et al. 2006), The wide
44 interest in SARs over many decades (e.g. Preston 1960, Connor and McCoy 1979, Rosenzweig 1995,
45 Scheiner et al. 2011, He and Hubbell 2011, Storch 2016) testifies to the long-standing appreciation by
46 ecologists of the centrality of scaling issues.

47

48 Classically, SARs have been drawn by conducting intensive biological surveys of different sized areas,
49 which may be nested (e.g. a quadrat within a field, within a county, within a nation) or non-overlapping
50 samples (e.g. a series of islands or political entities of different sizes), and may be ecological isolates
51 (e.g. islands or discrete forest patches) or arbitrarily defined samples from a larger whole (e.g. quadrats
52 or political entities); a great deal of discussion has focussed on the properties of SARs composed in
53 these different ways (e.g. Rosenzweig 1995, Scheiner 2003, Tjørve and Turner 2009, Scheiner et al.
54 2011). The shape of SARs has also been hotly contested, and after decades of debate about the relative
55 merits of power law and logarithmic models (e.g. Connor and McCoy 1979), in recent years a wide
56 range of other functional forms have been explored (reviewed by Tjørve 2003, 2009, see also Scheiner
57 et al. 2011). More than 180 years after its birth, the SAR remains an active topic of ecological research.

58

59 The reason for the continued popularity of the SAR is obvious: it provides a clear language for
60 expressing species-richness information across the full range of ecologically relevant scales. As such, it
61 has great potential as a tool for describing and monitoring multi-scale aspects of biodiversity. Policy is
62 often concerned with the preservation of biodiversity at national, continental (e.g. Gothenburg targets,
63 2001) or global (e.g. CBD, 2002) scales, whereas most biodiversity monitoring is conducted at very fine
64 spatial scales (sometimes $<1 \text{ m}^2$). This mismatch between the scales of our policies and of our data
65 creates serious challenges, especially when assessing biodiversity change. It has recently become
66 apparent, for example, that environmental changes may affect biotic diversity differently at different
67 scales (Smart et al. 2006a, Keith et al. 2009, Keil et al. 2011); biotic homogenization for example may
68 increase local (α) diversity while decreasing diversity at coarser (β and γ) scales (Socolar et al. 2016);
69 conversely some invasive species may decrease α while increasing γ -scale richness (Rosenzweig 2001,
70 Powell et al. 2013). SARs reflect biodiversity across a wide range of scales (incorporating α , β , γ and
71 coarser scales) and so should provide an efficient tool for examining and communicating such
72 complexities. Global biodiversity monitoring needs have further increased the interest in SARs and
73 biodiversity scaling, due to the need to infer biodiversity patterns from growing global databases of
74 point locations to the regional scale; that is, biodiversity upscaling. Coordinated local sampling
75 schemes, together with reliable/robust upscaling methods, are critical for the integration and
76 generalisation of biodiversity information at large scales. Efficient tools for building reliable and
77 accurate SARs may prove increasingly useful for predicting the response of biodiversity to
78 environmental changes across scales, and to assess global conservation policy options (Pereira et al.
79 2013, Geijzendorffer et al. 2016).

80

81 However, one serious problem prevents the widespread application of SARs to multi-scale biodiversity
82 monitoring. The requirement for exhaustive surveys over large areas makes it impractical to survey

83 SARs repeatedly over a short period of time. Indeed, for many poorly studied taxa and regions, it would
84 be difficult to amass sufficient information to provide even a single coarse scale biodiversity estimate
85 with confidence (e.g. Erwin 1982, May 1990). If the SAR is to fulfil its promise, we need to develop
86 new approaches to parameterising it with finite investments of surveying effort.

87
88 Harte and Kinzig (1997) were the first to explore a method for upscaling biodiversity from local
89 samples. Their approach was based on the idea that the SAR should rise faster with area if dissimilarity
90 in species occurrences in small plots (species turnover or β diversity) increases more rapidly with
91 distance between plots (Harte et al., 1999; Krishnamani et al. 2004). Unfortunately the method
92 involved strong implicit assumptions that limited its applicability. More recently Harte and colleagues
93 have proposed more sophisticated and general approaches based on the maximum entropy inferential
94 method (Harte et al. 2008, 2009; Harte and Kitze, 2015). The past 15 years have seen a proliferation of
95 other new methods to address this problem, based on approaches ranging from relative abundance
96 distributions (Ulrich and Ollik 2005), species accumulation curves (Shen and He 2008), least distance
97 spanning paths (Smith 2009), multi-site zeta diversity of compositional turnover (Hui and McGeoch
98 2014) and 3-dimensional manifolds (Polce 2009). This sudden flowering of alternative approaches
99 brings with it a new challenge: how do we best choose a method for a particular application? Many of
100 the models have been tested against data, of course, but each against a different dataset, and in many
101 cases the tests have been relatively modest: attempting to up-scale by only 1 or 2 orders of magnitude,
102 or even less. This paper addresses this issue by testing a wide range of biodiversity upscaling
103 approaches on a single high quality dataset across a substantial range of scales, within a well studied
104 system. By working in an area with a “known” SAR, we can judge the effectiveness of the various
105 methods in estimating coarse-scale biodiversity.

106

107 **Methods**

108

109 **The CS dataset**

110 We make use of the GB Countryside Survey (CS), a periodic botanical survey programme organised by
111 the NERC Centre for Ecology and Hydrology (CEH). The CS focusses on a stratified random sample of
112 1 km cells within Britain, chosen to represent the full range of British landscapes (for further details on
113 CS methods see Firbank et al. 2003). Specifically, we will rely on the CS survey of 1998/1999
114 (hereafter “CS1999”), which coincides with the survey period for the *New Atlas of the British and Irish*
115 *Flora* (Preston et al. 2002), which we can use to generate our “true” SAR (see below). A total of 569 1
116 km² cells were examined in CS1999, scattered over the whole of Britain and its inshore islands (but
117 excluding Northern Ireland and more distant island groups). Within each 1 km cell, a wide range of
118 surveys was conducted, which can be roughly divided into areal surveys (various sized surveys of

119 habitat blocks) and linear surveys (1 x 10 m surveys of linear features such as roadsides, hedgerows and
120 banks of waterways). For our purposes, the most statistically “representative” surveys were the so-
121 called “X” plots, five of which are sited at random (one in each of 5 equally-sized subsections) within
122 each surveyed 1 km cell. The only departure from truly random placement is that X plots were not
123 allowed to overlap with linear features (but see below). X plots have the added advantage (for this
124 work) in being multi-scaled: each consists of a nested series of quadrats at 4, 25, 50, 100 and 200 m²
125 scales. Species presence/absence is measured at all five scales, and estimates of cover for each species
126 are recorded at the finest (2 x 2 m = 4 m²) and coarsest (14.14 x 14.14 m = 200 m²) scales. We made
127 data from all 5 scales available to researchers (in most cases: the authors of upscaling methods),
128 although most used only the coarsest scale (200 m²) data in fitting their models.

129
130 The fact that X plots were not allowed to overlap linear features arguably makes them less diverse in
131 species composition than truly random quadrats would be, as the inclusion of (potentially dissimilar)
132 vegetation from such strips would likely enhance diversity (Smart et al. 2006b). Consequently, we
133 developed a synthetic second set of samples, which we termed “X+Linear” samples (for clarity, the
134 original surveys are hereafter referred to as “X-only” samples). These composite samples were created
135 by choosing the linear feature closest in space to each X plot, and merging its species with those in the
136 coarsest (200 m²) X plot sample to produce an aggregate sample representing 210 m² (see Figure 1).
137 Where the same linear sample was the nearest neighbour of more than one X plot, it was assigned to the
138 X plot in closest proximity, and others were paired with their 2nd nearest linear surveys. If the X-only
139 analyses arguably underestimate local richness, these X+Linear composite plots are likely to
140 overestimate it, as they tacitly assume that all X plots would have included linear features had they been
141 placed truly at random. We feel confident that a truly representative sample would fall somewhere
142 between these two.

143 144 **Subsamples**

145 To provide a richer test of the various methods available, we developed a total of 16 test data sets. The
146 largest of these is the “Full” sample, which covers all 569 CS survey cells within the surveyed area, and
147 all five X plots within each. We also developed five regional subsamples, covering the “North”,
148 “Centre”, “East”, “West” and “South” of Britain (Figure 1). These were non-overlapping regions,
149 chosen to roughly correspond to natural divisions of the area, and as such they were not equal in area.
150 More importantly, they were also not equal in biodiversity, with pronounced regional differences in both
151 α and β diversity between regions (encompassing e.g. a > 2-fold range in mean species richness at 100
152 km² scale c.f. Lennon et al. 2001). We also developed two sets of five statistical subsamples from the
153 full dataset. “Wide-shallow” (WS) samples covered the full set of sampling locations, but included only
154 one X plot (or X+linear sample) of the five generally available at each site. By contrast, “Narrow-deep”

155 (ND) samples included all five X plots at each site, but included only 1/5 of the survey sites, chosen as a
156 stratified random sample following the original CEH landscape stratification. Both WS and ND sample
157 sets were non-overlapping, so that the sum of all five subsamples in either set constituted the Full
158 British CS sample.

159

160 Each of the 16 samples (full set + 5 regions + 5 WS + 5 ND) were assessed for both X-only and
161 X+linear sample strategies, making a total of 32 potential tests for each method employed. However,
162 the stratified nature of the statistical samples tended to make their multiple runs quite similar to each
163 other, and thus treating them as five separate estimates would both overstate their independence and
164 give them undue weight in the overall analysis. Consequently, to simplify reporting, each set of
165 statistical subsamples (WS and ND) were summarised by a single (mean) performance score, thus
166 leaving 16 tests (full set + 5 regions + WS mean + ND mean = 8, for each X-only and X+linear
167 datasets).

168

169 **The challenge**

170 The task we set ourselves was to estimate the SAR for scales ranging from 100 km² (10 x 10 km, the
171 minimum mapping unit of Preston et al. 2002) to the whole of Britain (or of a specific subregion) using
172 only the CS survey data. Even the finest of these scales was 500,000 times coarser than the 200 m²
173 scale of an X-plot survey (or 476,190 times larger than the 210 m² of an X+linear sample). For the
174 purpose of this exercise, we will treat the area of Britain as the summed area of all the 100 km² cells
175 covering Britain itself and the major outlying islands of the Shetland, Orkney and Hebridean Islands, a
176 total of 278,500 km². This is almost 14 billion times larger than scale of a single X plot, and
177 approximately 500,000 times larger than the full set of survey sites combined (more precisely: 503,799
178 times the area of the full set of X plots, or 479,808 times the area of the full X+linear sample). Levels of
179 upscaling in statistical subsamples (with only 1/5 as many samples used) were five times greater still
180 (2,518,995-fold for X-only analyses; 2,399,040-fold for X+Linear). The regional subsamples cover
181 areas between 46,100 and 77,200 km², with correspondingly smaller numbers of samples, giving
182 upscaling levels comparable to those for the full national dataset. Several of the methods considered
183 here have been tested before, in particular using tropical forest survey data from relatively small (e.g 50
184 ha, Shen and He 2008) plots. Such applications involve only relatively modest upscaling; the challenge
185 presented here is substantially more ambitious and more typical of the sort of tasks a practical upscaling
186 approach would be asked to perform in e.g. regional or national biodiversity estimation. To our
187 knowledge, only a few past papers (Ugland et al. 2003, Krishnamani et al. 2004, Harte et al. 2009) have
188 attempted comparable levels of upscaling, each for only a single model.

189

190 **Upscaling methods**

This article is protected by copyright. All rights reserved

191 As noted in the introduction, there has been a proliferation of novel methods for upscaling biodiversity
192 in recent years. We have brought together most of the global community of researchers addressing this
193 issue, presenting each with the same CS datasets. To ensure high levels of familiarity with the models
194 employed, most methods were fit by their original proponents, with the exception of the three variants
195 of the Ugland model and the Lomolino model, which were prepared by a working group composed of
196 ET, AS, RTJ, KIU and WU, and the power and logarithmic models, fit by VV and WEK. Further
197 details of the models are given in the sections below.

198 199 **Harte MaxEnt method**

200 The maximum entropy theory of ecology (METE) predicts the shape of metrics describing patterns in
201 the spatial distribution, abundance, and energetics of species (Harte, et al., 2008; Harte, 2011; Harte and
202 Newman, 2014). METE is a state variable theory in which the maximum entropy inference procedure
203 (Jaynes, 1982), coupled with constraints arising from knowledge of quantities such as the number of
204 species and the number of individuals at plot scale, determine unique and testable macroecological
205 metrics across all scales. METE predicts a non-power-law but universal form for the SAR; in particular,
206 if the local log-log slope of the SAR at each spatial scale is plotted against the average abundance per
207 species at each scale, then all SARs are predicted to fall on a universal curve (Harte et al., 2009).

208 Upscaling species richness can either be carried out from knowledge of the number of species and the
209 number of individuals at any one spatial scale, or alternatively from knowledge of the number of species
210 at two spatial scales (from which information the abundance at each of those scales can be inferred from
211 METE). The CS dataset provides abundance information in terms of % cover, but not the number of
212 individuals (which is hard to assess in many plant species). For that reason we can upscale using the X-
213 only plot data, which does include measured values of species richness at several plot-sized scales, but
214 we cannot use the X+linear plot data, as only one scale is available.

215 The capacity of METE to upscale has been tested successfully for tree species in the Western Ghats,
216 where species richness was upscaled over a scale range of 24 million, from ¼ ha plots where census
217 data are available to the entire 60,000 km² biome (Harte et al., 2009). Other tests of upscaling with this
218 method have been carried out for arthropods and trees in a Panamanian Preserve and trees in the
219 Amazon (Harte and Kitzes, 2015). An important limitation of the MaxEnt method, however, is that it is
220 designed only for upscaling species richness within contiguous blocks of similar habitat. Moreover,
221 accumulating evidence (Harte, 2011; Harte and Newman, 2014), suggests that due to its reliance on
222 equilibrium statistical outcomes METE's successes are restricted to relatively undisturbed ecosystems,
223 with failures observed in habitats strongly influenced by human activity.

224 225 **Ugland TS loglinear method**

226 If METE is designed for uniform habitat, Ugland and colleagues' (2003) TS model was explicitly
227 designed for surveys covering multiple potentially dissimilar communities. Most assemblages have a
228 complex covariance structure between species and subareas. This leads to a largely unrecognized aspect
229 of predicting the number of species by upscaling: with the addition of new subareas or habitats the
230 observed species accumulation curve (across regions or habitats) will not only extend the previous
231 within-habitat accumulation curve, but also tend to lie above the accumulation curves for smaller
232 subareas. The rate of (vertical) increase of the species-accumulation curves provides the best estimate of
233 total species richness. Ugland et al. (2003) derived an exact analytical expression for the expectance
234 and variance of the species accumulation curve in all random subsets from a given area. In this method
235 the whole area is divided into subareas, and an increasing sequence of accumulation curves is
236 constructed as follows. The first accumulation curve (the bottom curve) is obtained by taking the
237 average of all single subareas. The second accumulation curve is obtained by taking the average of all
238 accumulation curves based on two randomly chosen subareas. For example, if there are five subareas,
239 the total number of subsets of two subareas is the binomial coefficient $5 \cdot 4 / 2 \cdot 1 = 10$, so the second
240 accumulation curve will be the average of 10 curves. In the same way the third accumulation curve is
241 the average of accumulation curves based on all possible subsets of three subareas. This procedure is
242 repeated until we end up with the last accumulation curve which is obtained by randomization of all
243 available samples in the data set.

244 It is the terminal points of this increasing sequence of species accumulation curves that contain the
245 crucial information of the accumulation rate of new species as sampling effort is increased to new
246 subareas. The total species curve (the TS-curve) is then defined as the curve connecting these end
247 points. In a semilogarithmic plot these curves frequently appear linear, and Ugland's estimator is then
248 simply the linear extrapolation of the TS-curve to the whole area in the semilog plot.

249 **Ugland Ten-at-a-time method**

250 We also used a variant of the method presented in Ugland et al. (2003), where the mean number of
251 species in a set of samples with the same number of plots is regressed with a semi-log function against
252 the log of summed plot area. In this case, we used 10 groups of 10 plots, 20 plots, 30 plots, and so on,
253 until the last group contained the entire set of plots (of which there is but one group). We examined
254 groups of 50, 100, 150, etc. plots, but the results were similar to the method using multiples of 10 plots
255 at a time.

256 **Ugland PAM method**

257 A third method of applying the Ugland approach was pioneered by Jobe (2008), using the non-
258 hierarchical clustering method algorithm known as partitioning around medoids (PAM) to determine the
259 subclasses of sites for computing species accumulation curves. The original Ugland estimation method
260 requires an *a priori* grouping of observations, so the introduction of PAM clustering allows such group
261 assignments to be done on an objective basis in cases where no such classification is available. There

262 are no hard and fast rules for selecting these groups, but the goal is for groups to contain ecologically
263 distinct observations (e.g. communities, assemblages, etc). PAM makes the grouping process more
264 objective by using compositional similarity among sites as reflected in the clustering algorithm to select
265 both the optimal number of groups and the membership of each group.

266

267 **Shen and He method**

268 There is a growing literature of methods devoted to estimating species richness in an area from random
269 samples taken from within it (e.g. Palmer 1990, Chao 2005, Magnussen et al. 2006), often using
270 resampling techniques with replacement. While these methods are not designed to estimate the full
271 SAR, they can be used to upscale from a set of point data to estimate the overall species richness of the
272 area from which they were drawn, and thus to estimate at least one point (the top) of the SAR. Many of
273 these methods, however, have been shown to overestimate richness (e.g. Xu et al. 2012). Shen and He
274 (2008) developed a novel approach based on sampling without replacement, using information on
275 presence/absence data on species incidence, based on a modified Beta distribution. The method is not
276 spatially explicit, and provides a single estimate for the species richness of the full sampled area. To
277 derive finer scale estimates, the area to be estimated was shifted downwards (but see Discussion). In the
278 X-only datasets, the Shen and He model was fitted both to data from the full 200 m² survey plots, but
279 also to the finest scale (4 m²) survey data, allowing the model's sensitivity to sample plot size to be
280 assessed.

281

282 **Šizling method**

283 Arnošt Šizling and David Storch (Appendix S1) have developed a method using the frequency
284 distribution of species' occupancies to estimate the shape of the SAR between two fixed scales, based
285 on their "finite area model" of the SAR (Šizling and Storch 2004); different species-occupancy
286 distributions produce SARs with different degrees of curvature, with the standard deviation of
287 occupancy playing a key role (see Appendix S1). This approach is a 'scaling between method', rather
288 than an upscaling method *per se*; that is, it estimates the increase in species richness as one moves from
289 a unit survey plot (here a 200 or 210 m² CS sample) up to a predetermined maximum value. Thus it
290 requires an estimate of "known" global species richness for the area in question and information from
291 local samples to estimate species richness at scales in between these two known points on the curve. It
292 would have been unfair to provide this model with more information than its competitors, and so the
293 modeler had to make an arbitrary global richness estimate (1000) to implement his model; but in
294 practice, the method might best be combined with other methods that make effective global richness
295 estimates in order to estimate the SAR as a whole (see Discussion). The method is based on the fact that
296 if we assume aggregated distributions, the proportional occupancy constrains the size of the maximum
297 gap in a species' distribution (the "area of saturation," Šizling and Storch 2004), which in turn

298 determines the number of species sampled within given size window, i.e. in a specific area. As that and
299 occupancy of the unit area together determine the slope of log-SAR (z), one could compose the SAR for
300 any given number of species randomly chosen from the observed frequency distribution of occupancies,
301 and thus estimate species richness of any area between the unit and total areas.

302

303 **Hui models**

304 Cang Hui developed three additional new approaches for this paper; each will be described briefly here,
305 with full details and computer codes given in Appendix S2.

306 **Hui 1: Occupancy Rank Curve (ORC)**

307 This approach proportionally scales up a sampling occupancy rank curve (ORC) by assuming that the
308 sampling is sufficient and representative of the wider area from which the samples were drawn.

309 Specifically, if one plots the number of sites occupied by species in order of ubiquity, the resulting ORC
310 for samples follows closely a truncated power law (Hui 2012): $O = c_1 e^{c_2 \cdot R} R^{c_3}$, where O and R represent
311 the occupancy and the ranking of a species. This shape consists of two components: a power-law
312 function depicting the scale-free relationship between species ranks and their occupancies, and an
313 exponential cut-off depicting a Poisson random process of species occupancy. The power-law
314 component is largely applicable to widespread/common species, with their distributions reflecting the
315 spatial partitioning (or sharing) of heterogeneous, often approximately fractal, habitat, whilst the
316 exponential cut-off reflects the chance events of flickering presence/absence of rare species. This
317 method then scales up the sampled ORC to estimate the true ORC proportionally according to the
318 sampling effort (replacing c_1 from the sampling ORC with $C_1 = c_1/s$, where $0 \leq s \leq 1$ represents
319 sampling effort) and the maximum ranking for the enlarged ORC (i.e. solving $1 = C_1 e^{c_2 \cdot R} R^{c_3}$ for R)
320 then represents the true number of species in the community.

321 **Hui 2: Hypergeometric Discovery Curve (HDC)**

322 Sampling patterns do not necessarily follow the same shape as the true biodiversity patterns, because the
323 probability of discovering a species in a sample does not correlate linearly with the species' true
324 occupancy; the probability of encountering very rare species in a moderately-sized sample is near zero,
325 with probability rising with occupancy in a sigmoid fashion and asymptoting near 1 for very common
326 species. The sampling theory of species abundances has been extensively studied (Dewdney 1998;
327 Green and Plotkin 2007), and Hui has developed an equivalent sampling theory of species occupancies,
328 together with its continuous approximation for random sampling (Appendix S2). In particular, we need
329 the sampling probability ($prob(i|j)$) of discovering a species in i samples given a specific true
330 occupancy of j . For random sampling without replacement, this follows a hypergeometric distribution.
331 Importantly, sampling can affect the shape of observed occupancy frequency distribution (OFD),
332 $f(i) = \sum_{j=1}^m prob(i|j)F(j)$, where f is observed OFD, F true albeit unknown OFD, and m the sample
333 extent divided by the grain. This formulation follows the discrete Fredholm equation (also Volterral
This article is protected by copyright. All rights reserved

334 integral equation) of the first kind (Arfken 1985), with $prob(i|j)$ the kernel function and F a solvable
 335 positive vector. Despite the diverse parameteric forms of OFDs (Hui and McGeoch 2007), we reduce
 336 the computational demand for parameter optimisation by using a lognormal distribution ($F(j) = S \cdot$
 337 $LN(j|\mu', \sigma')$) centred at the middle of the possible logarithmic occupancy ($\mu' = \ln(m) / 2$) such that its
 338 95% confidence interval encompasses the entire range of occupancy at logarithmic scale ($\sigma' =$
 339 $\ln(m) / 3.92$), making species richness the sole variable to be estimated from the parameter
 340 optimisation.

341 **Hui 3: Zeta diversity**

342 Zeta diversity represents the overlap in species across multiple samples (Hui and McGeoch 2014).
 343 Unlike pairwise beta diversity which lacks the ability to express the full set of diversity partitions among
 344 multiple (≥ 3) samples, zeta diversity can express and potentially explain the full spectrum of
 345 compositional turnover and similarity (Latombe et al., 2017), with power law and negative exponential
 346 the most common forms of zeta diversity declines (with increasing number of included samples). We
 347 use a truncated power law to ensure a good fit to zeta diversity decline and then estimate the number of
 348 new species that are expected to occur when adding extra samples (i.e. the level of completeness) based
 349 on fitted zeta diversity decline. The expected number of species in an area can then be estimated
 350 according to the generic estimator developed in Hui and McGeoch (2014); note, the Chao II estimator is
 351 only a special case for exponentially declining zeta diversity. As the formulation is based on
 352 combinatorial probabilities, to reduce the overflow error (a combination of floating-point inaccuracy in
 353 any numerical computation platforms and combinatorial explosion [of formulation complexity] with
 354 increasing number of samples), we first estimate the number of new species encountered when adding
 355 one extra sample and then calculate the expected number of species using integral approximation.

357 **Ulrich and Ollik method**

358 Ulrich and Ollik (2005) made use of a different method based on Relative Abundance Distributions
 359 (RADs), which was originally designed to estimate the upper and lower limits of species richness in a
 360 focal region. Under the assumption that the occupancy - species rank order distribution is either a
 361 lognormal or a logseries and that the least abundant species has an occupancy of one cell (200 m²), they
 362 estimated upper species richness boundaries from the logseries by

$$363 \quad E_s = \frac{\ln Int + \ln N_{A1} - \ln N_{S1}}{slope} \quad (1)$$

364 and lower species richness boundaries from the lognormal distribution by

$$365 \quad E_s = \frac{2 \ln Int + \ln N_{A1} - 2 \ln N_{S1}}{slope} \quad (2)$$

366 where $\ln Int$ and $\ln slope$ are natural logarithm of the intercept (Int) and the slope of an exponential
367 regression through the middle 50 percentile of the respective abundance distributions and $\ln N_{SI}$ and \ln
368 N_{AI} are the natural logarithms of the numbers of individuals of the most abundant species of the whole
369 community within the area A_{total} and of the sample of area A_I , respectively. N_{AI} comes from proportional
370 upscaling of the sample area to total area: $N_{AI} = N_{SI}A_{total} / A_I$.

371

372 **Smith method**

373 A species–distance relationship (SDR) was explored by Smith (2008) as a method for estimating the
374 SAR from point survey data. The SDR slope was found to be highly correlated with the slope of the
375 SAR for the US Breeding Bird Survey data at large geographic scales. The SDR is calculated by
376 estimating the path of shortest length connecting a set of localities, then estimating cumulative distance
377 and cumulative diversity along the path. In the present analysis, data for all X- or X+linear plots were
378 lumped within a given 1 km² sampling cell (except for the Wide-Shallow subsamples, as these only
379 contained one X plot per cell). This is because locality size *per se* was found not to have a significant
380 influence on the slope of the SDR, whereas sample size (which affects number of individuals surveyed)
381 per locality did.

382 SDRs were calculated for all subsets of the Countryside Survey data using 1 km² cells as
383 localities. No correction was made for sample size. Distance was calculated as Cartesian distance
384 between the midpoints of the cells. Mean slopes of the SDR are based on 200 values (100 paths, each
385 containing 10 cells and measured in forward and reverse directions). To estimate the slope of the SDR,
386 linear regression and standardized major-axis regression were performed. Setting then the slope of the
387 SDR to equal the slope of the SAR, diversity estimates were made for the relevant portions of Britain by
388 assuming two different values for alpha diversity. First, average alpha diversity was calculated for the
389 plots (200 m² or 210 m² for X- and X+linear plots, respectively). Second, average alpha diversity per
390 cell (1 km²) was calculated by combining all plots in a sampling cell; this will underestimate diversity
391 for a 1-km² area.

392

393 **Polce and Kunin method**

394 The SAR rises for two reasons (see e.g. Scheiner et al. 2011): a larger area both encompasses more
395 environmental and spatial diversity than a small area, and it includes more total individuals (and thus
396 constitutes a larger sample). These two component processes, increased sample size and increased
397 spatial differentiation, may be expected to behave rather differently with increasing area. In order to
398 factor out these two component processes, we randomly sampled (a) different numbers of quadrat
399 surveys from constant sized “windows” of focal area (to estimate the pure sample size effect), and (b)
400 constant numbers of quadrat samples chosen from different sized windows (to estimate the pure spatial
401 scale effect), and tested the fit of a range of convex and sigmoid curves (from Tjørve 2003) to each

402 component process. Note that in these analyses, total sample size for a set of quadrats is expressed in
403 units of area (total m² surveyed), as that is essential for later steps of the analysis. We then constructed
404 a 3-dimensional manifold model as a multiplicative combination of the best-fitting sample-size and
405 scale models (see Polce 2009). Pilot work suggested that the MMF model [$Y = (a \text{ Samplesize}^c)/(b +$
406 $\text{Samplesize}^c)$] provided the best fit to the pure sample size component (sampled within a fixed window
407 size), whereas a power law ($Y = d \text{ Scale}^z$) performed best for pure spatial differences (at constant sample
408 size). These two component models could then be combined multiplicatively, to derive a final model:

$$409 \quad Y = (a * \text{Scale}^z * \text{Samplesize}^c)/(b + \text{Samplesize}^c) \quad (3)$$

410 Fitting this 3-dimensional model to the dataset, the SAR can be estimated as the the value of Y over the
411 diagonal line where $\text{Samplesize} = \text{Scale}$.

412

413 **Lomolino model**

414 We also fit a suite of models commonly fit to SARs and to the plot-based species-accumulation curve
415 (SAC) from each dataset (see Tjørve 2003 for models). Preliminary results here indicated that in most
416 cases the “Lomolino” model worked best ($S = a / (1 + b^{\log_{10}(c/A)})$), where S is number of species, A is
417 area, and a, b, and c are model parameters fit using the Gauss-Newton method for non-linear regression
418 (Myers 1990). In most cases the AIC weight of the Lomolino model was ~1, and where it was not, it
419 was equally tied with other models that were nested within the Lomolino model. Therefore, we used
420 only the Lomolino model to fit each dataset.

421

422 **Power-law and Logarithmic models**

423 To complement the range of recently-derived methods, we have included a few “old-fashioned”
424 approaches to SAR estimation. Arrhenius (1921) proposed a power law ($S = cA^z$) as the best descriptor
425 of the SAR, and Preston (1962) suggested that the “canonical” SAR would have an exponent (z) of 0.25.
426 Subsequent work (e.g. Connor and McCoy 1979, Rosensweig 1995) has suggested somewhat less steep
427 z-values predominate in many continental systems, with a consensus z of approximately 0.2. Thus, we
428 generated SAR estimates by simply computing mean species richness at the 200 m² scale X plot
429 samples (and 210 m² for the X+Linear samples) and scaling up to coarser resolutions using power-law
430 curves with these two slopes. We also took advantage of the multi-scaled nature of the CS X plot
431 surveys, fitting both power and semi-logarithmic (after Gleason 1922) models to the observed species
432 richness of each plot at the 5 scales of measurement (4, 25, 50, 100, and 200 m²), and extrapolating
433 median estimates for each. As the X+Linear data are available only at a single scale, these
434 extrapolations of power-law and semi-logarithmic curves can be done only on the X-only datasets.

435

436 **Model summary**

437 Altogether, we have assembled thirteen different models for upscaling biodiversity, and several of them
438 (the power-law, Shen and He, Ugland's TS and Ulrich and Ollik's methods) have been implemented in
439 multiple forms, for a total of nineteen sets of predictions. These methods may be grouped conceptually,
440 based on the approaches they take to the challenge of estimating coarse scale species richness from fine
441 scale samples (Figure 2). Three of the methods (power law, logarithmic and Lomolino) involve
442 parameterizing and extrapolating a well-studied SAR curve from the observed data. This is an entirely
443 phenomenological approach to upscaling. Two other models (Harte's MaxEnt model and Hui's HDC)
444 also extrapolate functions, but with curves that are built on a strong underlying rationale concerning the
445 patterns expected from random community patterns under constraints. Three models are based on
446 sampling processes from species occupancy (Hui ORC, Šizling) or abundance (Ulrich and Ollik)
447 distributions. Two additional models (Shen and He, Ugland's TS) focus specifically on sampling
448 processes and the resulting accumulation of species. The Polce and Kunin model is similar to Ugland's
449 sampling process approach, but with an explicit emphasis on spatial turnover processes. Such spatial
450 turnover in species is central to Hui's Zeta model, and plays a substantial role in the Smith model as
451 well, which in turn links back to phenomenological curve estimation approaches.

452

453 **Estimating the “True SAR”**

454 The quality of the various SAR predictions can only be tested by comparing them to the “true” SAR for
455 the focal region. This was estimated using data from the *New Atlas of the British and Irish Flora*
456 (Preston et al. 2002; hereafter: “NABIF”), which was compiled based on surveys from the late 1990s,
457 thus approximately at the same time as the CS 1999 sample. In contrast to an earlier attempt at a UK
458 floral atlas (Perring and Walters 1962), the NABIF's compilers made a concerted effort to ensure a
459 relatively even survey effort across the area in a fairly narrow time window, and in particular to avoid
460 false negatives due to the underreporting of common species and the false positives that result from the
461 compilation of records over long periods of time. While no biodiversity survey can be treated as
462 perfect, the NABIF is arguably one of the highest quality biodiversity atlases currently available
463 anywhere. In addition to vascular plants, the CS survey included a predefined set of 160 relatively
464 common and distinctive bryophyte and lichen taxa (species or species groups); consequently distribution
465 maps for these taxa were acquired from the bryophyte and lichen recording schemes, respectively (M.O.
466 Hill, pers. comm., J Simkin pers. comm). The true SAR was composed by superimposing a series of
467 coarser grids (with resolutions from 400 km² to 90000 km²) over the distributional dataset. Only grid
468 cells containing >75% land area were included in our analyses for each scale; at coarse scales, grid cells
469 were shifted somewhat (following Tjørve et al. 2008, Keil et al. 2011) to maximise the area fitting this
470 criterion. Our NABIF SAR calculations are being posted online (Polce and Kunin 2017).

471

472 **Assessing model performance**

This article is protected by copyright. All rights reserved

473 To assess the quality of the predictions of each model, we examined two quality criteria, appropriate to
474 somewhat different applications. One goal of diversity upscaling is to estimate the Total Species
475 Richness (TSR) in a focal region, whilst for other applications, it is valuable to estimate species richness
476 across a range of scales within the region, providing an estimate of the region's Species-Area
477 Relationship (SAR). We assessed model predictions against both of these criteria: SAR and TSR fits.
478 To assess the quality of SAR fits, we examined the mean absolute value of the difference between
479 predicted and true species richness values at a given scale, expressed relative to the true richness value
480 at that scale, which we term the "mean relative error" (or MRE):

$$481 \quad \text{MRE} = (1/n) \sum_i (|S_{\text{predicted},i} - S_{\text{true},i}| / S_{\text{true},i}) \quad (4)$$

482 Where $S_{\text{predicted},i}$ is the number of species predicted at scale i , $S_{\text{true},i}$ is the number observed at that scale
483 in the true SAR, and the summation is across n observed scales (9 scales in the regional analyses, 10 in
484 the full national and statistical subsample analyses). Note that we normalise errors by dividing them by
485 the true SAR value at each scale, so that, e.g., a 100 species error is deemed to be a larger mistake when
486 the true value is 100 than it is when the true value is 1000. This has the additional advantage of
487 allowing model fit to be expressed as a dimensionless fraction: the mean proportional error in
488 estimation. We have also calculated model fits using a number of other popular metrics (e.g. RMSE,
489 Pearson χ^2 ; see Data S1), but there is little qualitative effect on our findings; the same models perform
490 well by any sensible measure, with at most slight rearrangements of the order of the winners.
491 The quality of Total Species Richness (TSR) predictions was assessed using this same metric, but
492 evaluated only at the coarsest scale considered (278,500 km² in national analyses, and the area of each
493 region in regional analyses). In addition, we examined the correlation between true TSR and estimated
494 values across datasets, using the non-parametric Spearman's rank correlation, to test how consistently
495 high richness estimates were provided in highly species-rich regions. A similar correlation test was
496 performed for the full SAR fit, comparing the overall slopes of the estimated SARs (on logarithmic
497 axes) over the range of scales examined (100 – 278,500 km²) with the slopes of the true SARs over
498 those scales.

500 **Results**

501
502 The models tested differed greatly in their predictions for British plant richness; while the true TSR
503 value was 2326, the model estimates based on the X-only dataset ranged from only 62 (Median semi-
504 logarithmic curve extrapolation) up to 11,593 (Smith model) species. A somewhat narrower range of
505 predictions for the X+Linear dataset (1136 to 8647) was largely due to the fact that some of the more
506 extreme value models could not be applied to this dataset (e.g. the fitted semi-logarithmic and power-
507 law models, which needed multiple scales of diversity surveys). Examples of the true and estimated
508 SARs for the full British datasets are shown in Figure 3 (full data are provided in Data S1).

510 Fit scores for Total Species Richness predictions are given in Figure 4. Three families of models stand
511 out as the most reliable predictors of TSR: the two applications of Shen and He's method (2008;
512 hereafter "S&H"), the paired upper and lower estimates of Ulrich and Ollik (2005; hereafter: "U&O"),
513 and the Hui ORC models. The best predictive accuracy came from the S&H model, with estimates
514 generally within 10% of the correct TSR value (mean relative error = 0.097 ± 0.085) when
515 parameterised with 200 (or 210 for X+Linear samples) m^2 data; interestingly, the model performed
516 almost as well (mean relative error = 0.110 ± 0.091) when parameterised from much smaller ($4 m^2$)
517 vegetation samples. The U&O method and Hui's ORC model were the next best approaches: the upper
518 (log-series) U&O model had a mean relative error of $0.155 (\pm 0.083)$, whereas the lower (log-normal)
519 U&O model had a mean relative error of $0.211 (\pm 0.080)$. While these two methods are meant to serve
520 as upper and lower estimates, even the upper estimate was usually less than the true TSR. Hui's ORC
521 model performed nearly as well as the best U&O model in accuracy (mean relative error = $0.156 \pm$
522 0.089). The Ugland model, applied using the 10-at-a-time algorithm, performed reasonably well (MRE
523 = 0.210 ± 0.162), as did Hui's HDC model (MRE = 0.272 ± 0.173); no other approach came close (the
524 next best was the Polce & Kunin (P&K) model, MRE = 0.375 ± 0.158). Judging by the (Spearman's
525 rank) correlation coefficients between true and predicted species richness across sample sets, a similar
526 picture emerges, with the S&H methods ($\rho=0.825$ and 0.805 , when parameterised with 200 and $4 m^2$
527 data, respectively) and the Hui HDC, Zeta and ORC models ($\rho=0.800$, 0.752 and 0.697 respectively)
528 showing the highest correlation with true TSR, along with the Ugland (in particular, the 10-at-a-time
529 version with $\rho=0.788$), P&K ($\rho=0.728$) and U&O (both $\rho=0.655$) models.

530
531 The full SAR fits of the models are given in Figure 5. Accuracy was not as good as for SDR overall,
532 but one of Hui's models is the clear favourite in predicting the curve as a whole: the Hui ORC model
533 was well within 20% of correct SAR values on average (MRE = 0.177 ± 0.059). The lower (log-normal)
534 U&O model performed reasonably well (MRE = 0.272 ± 0.094), as did the Hui HDC model (MRE =
535 0.304 ± 0.202). The upper (log-series) U&O approach and the P&K method competed for fifth place
536 (P&K: MRE = 0.358 ± 0.118 ; U&O2: MRE = 0.369 ± 0.217). The only other models that averaged
537 within 50% of the correct SAR were the Hui Zeta model (MRE = 0.408 ± 0.134), the S&H model (MRE
538 = 0.418 ± 0.212), the Lomolino model (MRE = 0.442 ± 0.110) and the power law model with $z = 0.2$
539 (MRE = 0.451 ± 0.179) or $z = 0.25$ (MRE = 0.496 ± 0.444). As noted above, several other models were
540 tested only on X-only data, but none of them performed well enough to challenge the leading methods.
541 The slopes of the estimated SARs were generally uncorrelated with the true SAR slopes over the scales
542 considered here; only the median logarithmic model showed a significant positive correlation ($\rho =$
543 0.756 , $n = 8$, $p = 0.015$).

544

545 Sometimes consensus models can be constructed that perform more reliably than any one approach by
546 itself, especially when different models have contrasting weaknesses (e.g. Gritti et al. 2013). The P&K
547 and U&O methods tended to make contrasting errors, with the P&K model predicting a lower and
548 steeper SAR than was found in many cases, whilst the U&O method predicted a higher and flatter SAR
549 than that observed over the relevant range of scales, so that there was an inverse correlation between the
550 performance of the two models (Pearson $r = -0.470$). Consequently, the mean of these two estimates
551 often provided a better (and more reliable) SAR estimate than either model by itself (MRE = $0.222 \pm$
552 0.081). An even more successful combined SAR model could be constructed by using the S&H
553 estimate of TSR and then downscaling to finer scales using the Šizling method (MRE = 0.156 ± 0.062),
554 combining the strengths of both models. This combination provides our best SAR predictions.

555

556 The replicate runs of statistically subsampled datasets allow estimates of the variance in index values
557 holding sample effort constant (at 1/5 of the total sample). Figure 6 shows the coefficients of variation
558 in these replicated analyses. Most models showed acceptable levels of variation in estimates, although
559 the Smith (2009) model, Hui's Zeta model and approaches based on median fits of classical SAR
560 models (Power law and semi-logarithmic) showed much higher variation than the others tested. For
561 many of the models (most strikingly in the two Ulrich and Ollik models), variation between runs was
562 substantially higher in the "Narrow-deep" analyses than in the "Wide-shallow" runs, presumably
563 because the latter allowed higher levels of statistical independence between samples. For some of the
564 models (most notably the Lomolino, Uglund PAM and Ulrich and Ollik models) these statistical sub-
565 samples also tended to produce systematically lower up-scaled biodiversity predictions than resulted
566 from the full dataset, even though each set of five (non-overlapping) subsamples comprised the full
567 sample set, and all were being used to estimate the same full British SAR.

568

569 **Discussion**

570

571 The challenge of upscaling biodiversity from plot to regional or national scale is an important goal of
572 spatial ecology, one with the potential for important practical value. If we could reliably estimate coarse
573 scale species richness from fine scale samples, it would allow biodiversity estimation in poorly studied
574 regions and taxa, and facilitate the monitoring of multi-scale biodiversity change and the scaling up of
575 experimental results. A range of methods have been proposed to address this issue, but there has to date
576 been no clear consensus as to their relative strengths and weaknesses. To test these methods, we set a
577 much more ambitious test than has usually been applied, requiring species richness to be estimated at
578 scales some 500,000 times larger than the full dataset used – and 14 billion times larger than a single
579 sample plot (the scale of resolution from which richness was extrapolated by most of the methods). The
This article is protected by copyright. All rights reserved

580 models considered varied greatly in their performance in this test, but the best of them did well enough
581 to suggest that they have the potential for useful application in the near term. Nonetheless, further tests
582 of these methods should be attempted on datasets covering other taxa and regions, so that the generality
583 of our conclusions can be ascertained. Many of the models (especially those with relatively inflexible
584 shapes) may be expected to fit much better in some areas than in others; differences in species richness,
585 evenness, habitat diversity and spatial patchiness may all affect the form of SARs (Tjørve et al. 2008),
586 and thus may improve the relative success of some models over others. Similarly, different models may
587 be differentially sensitive to differences in the structure and intensity of sampling (CS is perhaps a best-
588 case scenario), which may again affect relative performance. Only by examining a wide range of
589 datasets with differently diversity patterns can we be certain of the generality of our results.

590

591 **Specific model performance**

592 Harte and colleagues (Harte et al. 1999, 2005, 2007) pioneered the study of biodiversity upscaling, and
593 their MaxEnt approach (Harte et al. 2008, 2009) is an important conceptual advance. As expected in the
594 fragmented and human-influenced habitats of the UK, the METE model performed poorly in our trials,
595 greatly underestimating coarse-scale species richness despite its record of success in upscaling within
596 relatively undisturbed and contiguous habitat (Harte et al. 2009; Harte and Kitzes, 2015). Harte's
597 MaxEnt approach can be estimated using surprisingly little information (see Methods), which makes it a
598 strikingly efficient tool, but also a very inflexible one. That property is a virtue when applying the
599 model to the sort of homogeneous natural community for which it was designed, but it may create
600 difficulties in applying the model to more anthropogenic landscapes. METE relies on natural
601 communities displaying statistical patterns that maximize entropy within ecological constraints, patterns
602 which may be slow to stabilise (Harte 2011). It would be useful to conduct future tests of the METE
603 upscaling method within contiguous extents of UK biomes that are relatively undisturbed by human
604 activity, such as within large areas of heathland.

605

606 After Harte et al.'s original (1999) paper, the TS method proposed by Ugland et al. (2003) is arguably
607 one of the longest established and best supported methods in the literature. For example, Jobe (2008)
608 found it to have a reasonable predictive accuracy when applied to tree diversity in the southeastern
609 USA. Extrapolation of the semilogarithmic curve fitted to the terminal points of the species
610 accumulation curves is a robust approach that is designed for heterogeneous environments and it is
611 insensitive to shifts in species abundance, as only presence/absence information is taken into account.
612 This is a great advantage in most applications as there is often substantial variability in the assessment
613 of numbers of individuals, and in many datasets (as here) data on population sizes are not available at
614 all. The TS-curve estimates the accumulation rate of new species as more subareas are covered; thus
615 only species' spatial distributions affect the curve.

This article is protected by copyright. All rights reserved

616

617 We tested three different implementations of Ugland's approach, but none of them predicted the SAR
618 very well. The approaches showed more than two-fold differences between the highest (PAM) and
619 lowest (10-at-a-time) estimates, but all three curves were substantially higher and flatter than the true
620 SAR over the scales considered here. The discrepancy is probably the result of the large number of
621 species that occur in just a few plots (e.g. 24.6% of all species were found in just one plot in the X-only
622 dataset), which causes the TS curve to rise very steeply initially, and then overshoot. This steepness
623 occurs at relatively fine scales (between the 200 m² scale of the survey plots and the scale of the finest
624 Atlas grid, 100 km²), but when extrapolated to the scales investigated here the curves flatten out and
625 have lower slopes than the actual SAR. The differences in performance between the 3 implementations
626 of Ugland's TS approach were instructive. While the PAM approach formed groups of similar plots, the
627 10-at-a-time approach assembled sets at random, and predicted fewer species at every scale. This
628 occurred because PAM groups were more divergent in composition between groups, resulting in faster
629 species accumulation curve as groups are combined.

630

631 The TS model's prediction of high, shallow SARs over the scales considered here was shared by several
632 other models without explicit spatial structure (e.g. the Ulrich and Ollik 2005 and Shen and He 2008
633 approaches). Indeed, in the case of S&H, the SAR approached an asymptote at a value close to the true
634 S value (see below). By ignoring spatial structure in species occupancy, these approaches tend to bring
635 in more new species with each added sample initially, but rapidly exhaust the species pool, so that few
636 species remain to be added at coarser scales (Scheiner et al. 2011). The spatial structure of natural biotic
637 communities means that expanding the sample continues to bring in new environments and thus new
638 species even at coarse spatial scales.

639

640 Another time-honoured approach to upscaling is curve extrapolation. We explored a range of options
641 here, including traditional canonical power-laws, but also several methods (median power-law,
642 logarithmic and Lomolino curves) that made use of the multi-scale nature of the field survey data to
643 estimate the slope of species accumulation. None performed particularly well in our comparisons, yet
644 some fared almost as well as some of the more complex approaches. The Lomolino model was the best
645 of a suite of 14 models (Tjørve 2003) commonly fit to species-area relationships, but its accuracy was
646 sensitive to the spatial dispersion and density of plots. When extrapolated from the entire dataset, the
647 Lomolino model sometimes gave accurate estimates of the total number of species, but underestimated
648 species number by several hundred when data subsets were used. The model displayed asymptotic
649 behaviour, rising very little above ca. 100 km². Our results suggest a cautious approach should be used
650 when fitting asymptotic models to SARs, even when the model fits well at the fine scale of survey plots.

651

652 The classical power-law relationship provided a surprisingly good fit to some of the datasets, although
653 different values of the exponent z fit different cases. However, the more variable slopes fit using the
654 median value of z fitted from the multiscale X-plot surveys (from 4 to 200 m² scales) produced
655 generally lower slopes, with very poor predictive power. These low fitted slopes are probably affected by
656 the uniformity of land management at these fine scales, especially in the X-only plots which were
657 constrained not to cross linear features; these resulted in particularly low SAR curves for the fitted
658 logarithmic model, which predicted a total of only 62 species for all of Britain, despite the presence of
659 more than 1000 species in the overall sample set! On the other hand, despite its abysmal performance in
660 estimating total S, the fitted logarithmic model was the only one of all those tested that showed a
661 significant positive correlation with the slope of the true SAR across datasets. Linear extrapolation
662 methods may predict unrealistically high total species richness when the true underlying species
663 accumulation curves reach an asymptote within the extrapolation domain. For example, in an
664 investigation of arthropods in the Azorean Laurisilva forests, Hortal et al. (2006) found very low beta
665 diversity and a rapidly saturating total richness, so that linear extrapolation became heavily biased. In
666 the UK, however, underlying heterogeneity is sufficient that 55% of sampled species were found in 7 or
667 fewer sampling quadrats. This large fraction of species with a narrow geographical distribution prevents
668 the species accumulation curve from flattening out, and thus favours straight line extrapolation.

669
670 Several other models showed relatively poor performance. The Smith (2008) model not only showed a
671 low predictive accuracy for both TSR and SAR shape, it also displayed extreme variability in richness
672 predictions across the multiple replicate subsamples, suggesting that its estimates are unstable. Unless
673 those problems can be addressed, there is little to recommend it for future applications. On the other
674 hand, the poor performance of the Šizling model (see Appendix S1) is not surprising, as it has been used
675 here for a task rather different from the one for which it was designed. The Šizling model is designed to
676 downscale the SAR from a known value of total species richness, based on the species-occupancy
677 distribution observed within a sample of cells. As such, its application here required the choice of an
678 arbitrary estimate of total richness (1000), which was not very accurate. The method is included here,
679 however, as it provides a valuable component of a mixed modelling framework, if used together with a
680 companion model for estimating total richness (see below).

681
682 The best performance in our tests came from a series of relatively recent models: the Shen and He
683 (2008), Ulrich and Ollik (2005) and Polce (2009) models, and the three Hui models and Šizling model
684 introduced here. Each had distinctive strengths and weaknesses. The Shen and He model performed
685 both well and consistently in estimating total S, but proved to be ill-suited to assessing the shape of the
686 SAR, presumably because it ignores the spatial structure of samples. Clearly, the development of a
687 spatially explicit version of this model should be a priority for future research. The Hui ORC and HDC

688 models performed more consistently, providing credible TSR estimates and the best estimates of the
689 SAR as a whole (ORC) of any model considered; they certainly merit further attention. HDC requires
690 reliable numbers of observed rare species in samples, while ORC relies on robust/representative
691 estimates of sampling occupancies for common species. The CS data obviously fulfill the latter of these
692 requirements (sampling common species) very well, but even a survey of this scale (and expense)
693 samples only a tiny fraction of rare species. This may help explain the superior performance of the
694 ORC model in our analyses. The Ulrich and Ollik method proved third-best in total richness estimation,
695 and provided the second best SAR fit of the models tested, suggesting it may be a useful alternative.
696 However, its performance was only moderate in either regard, and the two versions of the model did not
697 consistently bracket the true value, as they were meant to do (in most cases, both estimates were above
698 the true value of species richness). This suggests that the true occupancy - species rank order
699 distribution is not a symmetric lognormal but skewed in the lower part to have more rare than abundant
700 species.

701
702 The S&H and U&O methods are both examples of a broader literature devoted to estimating overall
703 species richness in an area based on representative samples (see also e.g. Palmer 1990, Chao 2005,
704 Magnussen et al. 2006). These methods have been designed to estimate TSR, but they are not explicitly
705 aimed at SAR estimation; thus it is not surprising that they both perform the former task more
706 effectively than the latter. Many of the methods developed for TSR estimation require large proportions
707 of the focal biota to be observed (see Ulrich and Ollik 2005), making them inappropriate for large-scale
708 applications such as the one attempted here. Moreover, systematic biases in most such estimates have
709 been documented in the past (reviewed in Shen and He 2008), further undermining their applicability.
710 The two methods employed here were both explicitly developed with an aim to increasing the accuracy
711 and range of such projections. While these models differ fundamentally in their approaches (with S&H
712 using sampling theory, whereas U&O extrapolate relative abundance distributions), our results here
713 suggest that they have both been quite successful in this respect.

714
715 The Polce & Kunin model was explicitly designed for the more difficult task of SAR estimation. While
716 it performed moderately well in our tests, its finer scale estimates (in particular) were often substantially
717 lower than expected. One potential reason for this is the clustered nature of the CS sample set, with 5
718 samples taken in each focal 1 km² site. The P&K method involved sampling random sets of
719 observations from varying sized sampling windows; when small numbers of samples were drawn from
720 relatively small areas (e.g. 400 km² or less), there was consequently a high probability of drawing
721 multiple samples in close proximity to one another, sampling less diversity than expected of a truly
722 random sample of that size. While the logic of the method (separating pure sample size and pure spatial
723 extent effects) is compelling, there clearly remains considerable scope for improvements.

724
725 Two of most accurate individual methods for SAR estimation were developed for this paper: Hui's ORC
726 and HDC methods. Both made use of the distribution of occupancy values across species in the sample.
727 The models differed in what they did with those values: the ORC method extrapolated the curve of
728 species occurrence frequencies using a truncated power law to assess how many species would be
729 expected to occupy 1 or more 200 m² plot, had all of Britain been surveyed; the HDC method examines
730 the number of species represented by different levels of occupancy in the sample, and estimates from
731 observation probabilities how many other such species were likely to have been missed. The SAR
732 downscaling approach developed by Šizling and Storch, which provided even better SAR estimates
733 when married to the Shen and He (2008) TSR estimate, was also based on species occupancy
734 distributions. The success of these three model here spotlights this general approach as one of great
735 promise for future SAR research.

736
737 Considering the diverse classes of models tested here (Figure 2), shows a high level of performance for
738 those based on species occupancy (Hui ORC, Šizling) and related (Ulrich & Ollik, Hui HDC)
739 approaches. Conversely, methods based around extrapolating specific curves (Power-law, Logarithmic,
740 Lomolino, and even MaxEnt) were far less successful. There was mixed success in approaches based
741 on subsampling and spatial species turnover, and there remains significant potential for further
742 developing such approaches.

743 744 **Combining models**

745 As noted above, consensus models combining more than one of the more promising approaches often
746 outperformed any single "best" model for predicting the total species richness or SAR shape. This
747 generally occurred because different methods showed contrasting errors. Such combinations come at a
748 cost (Levins, 1966); there is often a trade-off in modelling between precision (which requires
749 complexity) and insight (which requires simplicity). Developing hybrids of multiple incommensurate
750 approaches runs the risk of producing a method that works well, but which has no compelling logic.
751 Such approaches may prove useful, but they are intellectually ugly. We can only hope that they will be
752 supplanted in time by models that are both accurate and meaningful.

753
754 There are additional unexplored opportunities for methodological hybrids amongst the methods
755 presented here, given the wide differences in approach set out above. Note for instance that the Šizling
756 model requires the user to have a prior estimate of S_0 , the total species richness in the focal region (as
757 does the original Harte et al. (2008) MaxEnt approach), while the Shen and He (2008) model estimates
758 that quantity but cannot estimate diversity at finer scales with any accuracy. Feeding the Shen and He
759 (2008) TSR estimate into the new Šizling or Harte et al. (2008) model would then provide credible

760 estimates of both. Thus for example, if we incorporate the Shen & He estimate of S_0 into the Šizling
761 approach and then downscale, the resulting SAR has a mean relative error score substantially better than
762 any of the individual models tested (Figure 5).

763

764 **Reducing survey effort**

765 Our focal dataset may represent a tiny fraction of the whole British land surface (roughly one part in
766 500,000), but it nonetheless requires an impressive investment in time and money to survey. It would
767 obviously be advantageous to have methods that could be nearly as effective with much lower survey
768 effort. We explored this issue at three spatial scales: (a) reducing the total number of 1 km cells
769 surveyed (represented by the “Narrow-deep” subsamples), (b) reducing the number of quadrats sampled
770 in each focal 1 km cell (represented by the “Wide-shallow” subsamples), and in one case (c) surveying a
771 smaller total area for each quadrat (Shen and He’s 4 m² analysis compared to the 200 m² analyses of the
772 same model). Our results clearly suggest that reducing local sampling intensity is far less serious than
773 reducing the number of sites examined. Wide-shallow sub-samples showed much less variation in
774 estimates and (in many cases) notably less bias (relative to the full dataset) than did the equally large
775 (but coarse-scale) Narrow-deep samples (Figure 6). Reducing sample size at still finer scales (by
776 changing the size of the local sample plot) may have even less impact: for the one model that was tried
777 at multiple scales (Shen and He 2008), the predictive accuracy of the model was virtually identical when
778 fit using 4 m² scale occupancy data than when fit using 200 m² data, despite the 50-fold smaller area
779 surveyed.

780

781 One issue with reduced sampling intensity in many models was the introduction of a bias: many of the
782 methods made systematically lower species richness predictions when fit to random subsamples of the
783 dataset than when fit to the set as a whole, despite the fact that each combined set of 5 subsamples
784 comprised the full dataset. This behaviour was displayed by most methods considered, with the
785 exception of the Power-law and Logarithmic extrapolations and the Hui ODC model (where subsample
786 estimates and full set estimates were virtually identical), and the Smith and Hui Zeta models (which
787 behaved inconsistently in this regard). Two possible explanations for the general trend suggest
788 themselves: one statistical, the other biological. On one hand, the smaller datasets may be noisier
789 (relative to their information content), and this will tend to flatten the regression relationships for small
790 samples [a possible solution would be to use Model II regression or equivalent techniques]. A more
791 biologically meaningful explanation is that one needs relatively large samples to encounter rare species,
792 and it is the rarer species that cause the SAR to rise, especially at the coarser scales (see e.g. Tjorve et al.
793 2009).

794

795 **Ideal and empirical models**

This article is protected by copyright. All rights reserved

796 Looking back over the full set of methods explored here, one useful albeit *post hoc* distinction is
797 between “ideal” and “empirical” SAR models. Ideal models are based on theoretical attempts to
798 understand the appropriate shape that the SAR should be expected to take in natural communities. As
799 such, they have the potential to provide mechanistic insight into potential processes underlying SAR
800 shape, but they tend to be most appropriately applied to natural diversity patterns (rather than
801 anthropogenic ones) where such mechanisms may be thought to determine diversity patterns. Ideal
802 SAR model predictions tend to be relatively inflexible in shape, and as a consequence, they require
803 relatively little data to parameterise; examples range from the canonical power-law SAR (Arrhenius
804 1921, Preston 1962) to the recent development of Maximum Entropy models (Harte et al. 2008, 2009).
805 The inflexibility of such models makes them intrinsically ill-suited to monitoring, e.g. changes of
806 biodiversity in response to management or other human interventions, since they are insensitive (by
807 design) to precisely the sorts of shifts in SAR shape that we would wish to detect. At the other extreme
808 are models designed to assess the empirical SAR whatever its shape happens to be. Such approaches
809 pay for their flexibility by requiring substantially more information. Nonetheless, this flexibility is
810 needed for some applications; for example, if upscaling methods are to be used for multi-scale
811 biodiversity monitoring (see introduction) they will need to be flexible enough to allow anthropogenic
812 shifts in biodiversity scaling to be reflected in their results.

813
814 It is not surprising, given the highly anthropogenic nature of the British landscape, that the best
815 performing models in this analysis (Shen and He 2008, Hui’s HDC and ORC models, Ulrich and Ollik
816 2005) were all empirical approaches. It would be interesting to see how the relative performance of the
817 various approaches explored here would shift were they to be tested on data from more natural
818 landscapes. Several of the methods that performed relatively poorly here have already been shown to
819 behave quite well in such applications (e.g. Ugland et al 2003, Krishnamani et al. 2004, Jobe 2008).
820 Indeed, the contrasts between ideal and empirical models may be instructive if well tested methods for
821 each can be employed. In well studied areas with good historical species richness records, a reasonable
822 estimate of the natural SAR might be computed using an ideal model (such as that of Harte et al. 2008).
823 This may then be compared to a current SAR computed using one of the empirical models based on
824 current monitoring data. The difference between the two could be interpreted as the “footprint” of
825 anthropogenic activities on biodiversity across spatial scales.

826 827 **Conclusions**

828 The topic of biodiversity upscaling has been largely of theoretical interest to date, but it is an area that
829 has tremendous potential practical value. Robust and tested upscaling methods would allow the
830 assessment of species richness in poorly studied regions and taxa; they would also make it possible to
831 monitor multi-scale biodiversity change over time, and might allow the coarse-scale implications of

832 environmental or management changes to be inferred from (necessarily fine-scale) experimental results
833 if replicated across multiple sites. To do so we need methods that can be fit using sets of point survey
834 data, and that will be responsive to any anthropogenic changes in local richness and spatial turnover,
835 giving robust and accurate predictions. To test these methods, we need excellent ground-truthed
836 biodiversity survey data from diverse natural and anthropogenic communities across the globe. We
837 have brought together most existing methods for biodiversity upscaling, and have set them an ambitious
838 target: to estimate the total species richness and Species-area relationship of a sizeable land mass, using
839 scattered point biodiversity samples from only a tiny fraction of the total area. While methods differed
840 dramatically in their performance, the best of them did reasonably well. Despite a ca. 500,000-fold
841 increase in scale from the total area surveyed to the area to be assessed, the best of the approaches
842 reliably predicted total species richness within about 10%, and estimated the full species-area
843 relationship within about 18% of the true values. Combining contrasting methods allowed even better
844 accuracy, allowing the SAR to be estimated within 16%. While there is still substantial room for
845 improvement (in particular, in estimating SAR slope) and additional tests on other datasets (ideally
846 involving contrasting regions and taxa) would be welcome, our results suggest that biodiversity
847 upscaling has begun to come of age. It is notable that of the three best methods for SAR estimation, 2 ½
848 (Hui's ORC and HDC and methods, and Šizling's downscaling) are novel methods published here for
849 the first time, suggesting that the field is progressing rapidly. Additional tools are still in development,
850 but our results suggest that existing methods can begin being applied with some confidence.

851

852 **Acknowledgements**

853

854 This work was originally conceived and shaped by WEK, with all authors contributing to the analyses
855 and manuscript preparation and hence listed alphabetically. This work was supported by a UKPopNet
856 grant (Dispersed platforms for biodiversity research: developing methods and networks for multi-scale
857 research) to WEK, with further support from the EU FP7 SCALES (Securing the Conservation of
858 biodiversity across Administrative Levels and spatial, temporal and Ecological Scales ENV-2008-
859 226852) and EU BON (Building the European Biodiversity Observation Network; ENV-2012-308454)
860 projects, and by a fellowship at the Stellenbosch Institute for Advanced Studies. Atlas data were
861 provided by Chris Preston of CEH's Biological Records Centre, Mark Hill of the Bryophyte Recording
862 Scheme and Janet Simkin of the Lichen Recording Scheme. CH was supported by the National
863 Research Foundation of South Africa (nos. 81825 and 76912) and the Australian Research Council
864 (DP150103017). CP's research was supported by the Marie Curie BIOCONS (European Centre for
865 Biodiversity and Conservation Research) EST programme (MEST-CT-2004-514350). Work by AS and
866 DS was further supported by grant from the Czech Science Foundation no. 14-36098G. WU was
867 supported by grants from the Polish Science Committee (KBN 3 P04F 03422 and KBN 2 P04F 039 29).
This article is protected by copyright. All rights reserved

868

869

870 **References cited**

871

872 Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems,
873 prevalence, and an alternative. *Journal of Wildlife Management* **64**: 912-923.

874 Arfken, G. 1985. *Mathematical Methods for Physics*, 3rd edition. Academic Press, Orlando FL.

875 Arrhenius, O. 1921. Species and area. *Journal of Ecology* **9**: 95-99.

876 Azaele, S., A. Maritan, S. J. Cornell, S. Suweis, J. R. Banavar, D. Gabriel, and W. E. Kunin. 2015.
877 Towards a unified descriptive theory for spatial ecology: predicting biodiversity patterns across
878 spatial scales. *Methods in Ecology and Evolution* **6**: 324-332.

879 Chao, A. 2005. Species richness estimation. Pages 7907-7916 in N. Balakrishnan, C. B. Read and B.
880 Vidakovic, editors. *Encyclopedia of Statistical Sciences* Second edition vol. 12, Wiley, New York.

881 Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. *The*
882 *American Naturalist* **113**: 791-833

883 Dewdney, A. K. 1998. A general theory of the sampling process with application to the “veil line”.
884 *Theoretical Population Biology* **54**: 294-302.

885 Drakare, S., J. J. Lennon, and H. Hillebrand. 2006. The imprint of the geographical, evolutionary and
886 ecological context on species-area relationships. *Ecology Letters* **9**:215-227.

887 Erwin, T. L. 1982. Tropical forests: their richness in coleopteran and other arthropod species.
888 *Coleopterists Bulletin* **36**: 74-75.

889 Firbank, L. G., C. J. Barr, R. G. H. Bunce, M. T. Furse, R. Haires-Young, M. Hornung, D. C. Howard,
890 J. Sheail, A. Sier, and S. M. Smart. 2003. Assessing stock and change in land cover and
891 biodiversity in GB: an introduction to Countryside Survey 2000. *Journal of Environmental*
892 *Management* **67**: 207-218

893 Geijzendorffer, I. R., E. C. Regan, H. M. Pereira, L. Brotons, N. Brummitt, Y. Gavish, P. Haase, C. S.
894 Martins, J. B. Mihoub, C. Secades, D. S. Schmeller, S. Stoll, F. T. Wetzler, and M. Walters. 2016.
895 Bridging the gap between biodiversity policy data and policy reporting needs: an Essential
896 Biodiversity Variables approach. *Journal of Applied Ecology* **53**: 1341-1350.

- 897 Gleason, H. A. 1922. On the relation between species and area. *Ecology* **3**: 158-162.
- 898 Green, J. L., and J. B. Plotkin. 2007. A statistical theory for sampling species abundances. *Ecology*
899 *Letters* **10**: 1037-1045.
- 900 Gritti, E. S., A. Deputie, F. Massol, and I. Chuine. 2013. Estimating consensus and associated
901 uncertainty between inherently different species distribution models. *Methods in Ecology and*
902 *Evolution* **4**: 442-452.
- 903 Harte, J. 2011. *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*.
904 Oxford University Press, Oxford UK.
- 905 Harte, J., and A. P. Kinzig. 1997. On the implications of species-area relationships for endemism, spatial
906 turnover, and food web patterns. *Oikos* **80**: 417-427.
- 907 Harte, J., S. McCarthy, A. Taylor, A. Kinzig, and M. L. Fischer. 1999. Estimating species-area
908 relationships from plot to landscape scale using spatial-turnover data. *Oikos* **86**: 45-54.
- 909 Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. Maximum entropy and the state-variable
910 approach to macroecology. *Ecology* **89**: 2700-2711.
- 911 Harte, J., A. B. Smith, and D. Storch. 2009. Biodiversity scales from plots to biomes with a universal
912 species-area curve. *Ecology Letters* **12**: 789-797.
- 913 Harte, J., and E. Newman. 2014. Maximum entropy as a framework for ecological theory. *Trends in*
914 *Ecology and Evolution* **29**: 384-389.
- 915 Harte, J., and J. Kitzes. 2015. Inferring regional-scale species diversity from small-plot censuses.
916 *PLOS ONE* | DOI:10.1371/journal.pone.0117527.
- 917 He, F. L., and S. P. Hubbell. 2011. Species-area relationships always overestimate extinction rates from
918 habitat loss. *Nature* **473**: 368-371.
- 919 Hortal, J., P. A. V. Borges, and C. Gaspar. 2006. Evaluating the performance of species richness
920 estimators: sensitivity to sample grain size. *Journal of Animal Ecology* **75**: 274-287.
- 921 Hui, C. 2012. Scale effect and bimodality in the frequency distribution of species occupancy.
922 *Community Ecology* **13**: 30-35.
- 923 Hui, C., and M. A. McGeoch. 2007. Modelling species distributions by breaking the assumption of self-
924 similarity. *Oikos* **116**: 2097-2107.

- 925 Hui, C., and M. A. McGeoch. 2014. Zeta diversity as a concept and metric that unifies incidence-based
926 biodiversity patterns. *The American Naturalist* **184**: 684-694.
- 927 Jaynes, E. T. 1982. On the rationale of maximum-entropy methods. *Proceedings of the IEEE* **70**: 939-
928 952.
- 929 Jobe, R. T. 2008. Estimating landscape-scale species richness: reconciling frequency- and turnover-
930 based approaches. *Ecology* **89**:174-182.
- 931 Keil, P., J. C. Biesmeijer, A. Barendregt, M. Reemer, and W. E. Kunin. 2011. Biodiversity change is
932 scale-dependent: An example from Dutch and UK hoverflies (Diptera, Syrphidae). *Ecography* **34**:
933 392-401.
- 934 Keith, S. A., A. C. Newton, M. D. Morecroft, C. E. Bealey, and J. M. Bullock. 2009. Taxonomic
935 homogenization of woodland plant communities over 70 years. *Proceedings of the Royal Society*
936 *B – Biological Sciences* **276**: 3539-3544.
- 937 Kleijn, D., F. Berendse, R. Smit, and N. Gilissen. 2001. Agri-environment schemes do not effectively
938 protect biodiversity in Dutch agricultural landscapes. *Nature* **413**: 723-725.
- 939 Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-absence data.
940 *Journal of Animal Ecology* **72**: 367-382.
- 941 Krishnamani, R., A. Kumar, and J. Harte. 2004. Estimating species richness at large spatial scales using
942 data from small discrete plots. *Ecography* **27**: 637-642.
- 943 Latombe, G., C. Hui, and M. A. McGeoch. 2017. Multi-site generalised dissimilarity modelling: Using
944 zeta diversity to differentiate drivers of turnover in rare and widespread species. *Methods in*
945 *Ecology and Evolution* **8**: 431-442.
- 946 Lomolino, M. V. 2001. The species-area relationship: New challenges for an old pattern. *Progress in*
947 *Physical Geography* **25**:1-21.
- 948 Lennon, J. J., P. Koleff, J. J. D. Greenwood, and K. J. Gaston. 2001. The geographical structure of
949 British bird distributions: diversity, spatial turnover and scale. *Journal of Animal Ecology* **70**: 966-
950 979.
- 951 Levins, R. 1966. The strategy of model building in population biology. *American Scientist* **54**: 421-431.

- 952 Magnussen, S., R. Pelissier, FL. He, and B. R. Ramesh. 2006. An assessment of sample-based
953 estimators of tree species richness in two wet tropical forest compartments in Panama and India.
954 *International Forestry Review* **8**: 417-431.
- 955 May, R. M. 1990. How many species? *Philosophical Transactions of the Royal Society of London B*
956 **330**: 292-304.
- 957 Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomasd, R. Valencia, and J. L. Green.
958 2008. A general framework for the distance-decay of similarity in ecological communities.
959 *Ecology Letters* **11**: 904-917.
- 960 Ostling, A., J. Harte, J. L. Green, and A. P. Kinzig. 2004. Self-similarity, the power law form of the
961 species-area relationship and a probability rule: a reply to Maddux. *The American Naturalist* **163**:
962 627-633.
- 963 Palmer, M. W. 1990. The estimation of species richness by extrapolation. *Ecology* **71**: 1195-1198.
- 964 Pereira, H. M., S. Ferrier, M. Walters, G. N. Geller, R. H. G. Jongman, R. J. Scholes, M. W. Bruford, N.
965 Brummitt, S. H. M. Butchart, A. C. Cardoso, N. C. Coops, E. Dulloo, D. P. Faith, J. Freyhof, R.
966 D. Gregory, C. Heip, R. Holt, G. Hurtt, W. Jetz, D. S. Karp, M. A. McGeoch, D. Obura, Y.
967 Onoda, N. Pettorelli, B. Reyers, R. Sayre, J. P. W. Scharlemann, S. N. Stuart, E. Turak, M.
968 Walpole, and M. Wegmann. 2013. Essential Biodiversity Variables. *Science* **339**: 277-278.
- 969 Perring, F. H., and S. M. Walters. 1962. Atlas of the British flora. EP publishing Ltd, Wakefield.
- 970 Plotkin, J. B., M. D. Potts, D. W. Yu, S. Bunyavejchewin, R. Condit, R. Foster, S. Hubbell, J.
971 LaFrankie, N. Manokaran, L. H. Seng, R. Sukumar, M. A. Nowak, and P. S. Ashton. 2000.
972 Predicting species diversity in tropical forests. *Proceedings of the National Academy of Sciences*
973 USA **97**: 10850-10854.
- 974 Plotkin, J. B., and H. C. Muller-Landau. 2002. Sampling the species composition of a landscape.
975 *Ecology* **83**: 3344-3356.
- 976 Polce, C. 2009. Dynamics of native and alien plant assemblages: the role of scale. PhD dissertation,
977 University of Leeds
- 978 Polce C and W.E.Kunin. 2017. SAR dataset for British plants. University of Leeds.
979 <https://doi.org/10.5518/264>.

- 980 Powell, K. I., J. M. Chase, and T. M. Knight. 2013. Invasive plants have scale-dependent effects on
981 diversity by altering Species-Area Relationships. *Science* **339**: 316-318.
- 982 Preston, F. W. 1960. Time and space and the variation of species. *Ecology* **41**:612-627.
- 983 Preston, F. W. 1962. The canonical distribution of commonness and rarity. *Ecology* **43**: 185-215.
- 984 Preston, C. D., D. A. Pearman, and T. D. Dines. 2002. *New Atlas of the British and Irish Flora*. Oxford
985 University Press, Oxford.
- 986 Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge.
- 987 Rosenzweig, M. L. 2001. The four questions: what does the introduction of exotic species do to
988 diversity? *Evolutionary Ecology Research* **3**: 361-367.
- 989 Scheiner, S. M. 2003. Six types of species-area curves. *Global Ecology and Biogeography* **12**: 441-447.
- 990 Scheiner, S. M., A. Chiarucci, G. A. Fox, M. R. Helmus, D. J. McGlinn, and M. R. Willig. 2011. The
991 underpinnings of the relationship of species richness with space and time. *Ecological Monographs*
992 **81**: 195-213.
- 993 Shen, T. J., and FL. He. 2008. An incidence-based richness estimator for quadrats sampled without
994 replacement. *Ecology* **89**: 2052-2060.
- 995 Shmida, A., and M. V. Wilson. 1985. Biological determinants of species-diversity. *Journal of*
996 *Biogeography* **12**: 1-20.
- 997 Šizling, A. L., D. Storch, J. Reif, and K. J. Gaston. 2009. Invariance in species-abundance distributions.
998 *Theoretical Ecology* **2**: 89-103.
- 999 Šizling, A. L., and D. Storch. 2004. Power-law species-area relationships and self-similar species
1000 distributions within finite areas. *Ecology Letters* **7**:60–68.
- 1001 Šizling, A. L., and D. Storch. 2007. Geometry of species distributions: random clustering and scale
1002 invariance. Pages 77–100 *in* D. Storch, P. A. Marquet, and J. H. Brown, editors. *Scaling*
1003 *Biodiversity*. Cambridge University Press, Cambridge.
- 1004 Smart, S. M., K. Thomspson, R. H. Marrs, M. G. Le Duc, Maskell, L. C. and L. G. Firbank. 2006a.
1005 Biotic homogenization and changes in species diversity across human-modified ecosystems.
1006 *Proceedings of the Royal Society. London. B.* **263**: 2659-2665.

- 1007 Smart, S. M., R. H. Marrs, M. G. Le Duc, K. Thompson, R. G. H. Bunce, L. G. Firbank and M. J. Rossall. 2006b.
1008 Spatial relationships between intensive land cover and residual plant species diversity in temperate, farmed
1009 landscapes. *Journal of Applied Ecology* **43**: 1128-1137.
- 1010 Smith, K. T. 2008. On the measurement of beta diversity: an analog of the species-area relationship for
1011 point sources. *Evolutionary Ecology Research* **10**: 987-1006.
- 1012 Socolar, J. B., J. J. Gilroy, W. E. Kunin, and D. P. Edwards. 2016. How should beta-diversity inform
1013 biodiversity conservation? *Trends in Ecology and Evolution* **31**: 67-80.
- 1014 Stohlgren, T. J., C. Jarnevich, G. W. Chong, and P. H. Evangelista. 2006. Scale and plant invasions: a
1015 theory of biotic acceptance. *Preslia* **78**: 405-426.
- 1016 Storch, D. 2016. The theory of the nested species-area relationship: geometric foundations of
1017 biodiversity scaling. *Journal of Vegetation Science* **27**: 880-891.
- 1018 Tjørve, E. 2003. Shapes and functions of species-area curves: A review of possible models. *Journal of*
1019 *Biogeography* **30**: 827-835.
- 1020 Tjørve, E. 2009. Shapes and functions of species-area curves (II): a review of new models and
1021 parameterizations. *Journal of Biogeography* **36**: 1435-1445.
- 1022 Tjørve, E., W. E. Kunin, C. Polce, and K. M. C. Tjørve. 2008. The species-area relationship: separating
1023 the effects of species-abundance and spatial distribution. *Journal of Ecology* **96**: 1141-1151.
- 1024 Tjørve, E., and K. M. C. Tjørve. 2008. The species-area relationship, self-similarity, and the true
1025 meaning of the z-value. *Ecology* **89**: 3528-3533.
- 1026 Tjørve, E. and W. R. Turner. 2009. The importance of samples and isolates for species-area
1027 relationships. *Ecography* **32**: 391-400.
- 1028 Ugland, K. I., J. Gray, and K. E. Ellingsen. 2003. The species-accumulation curve and
1029 estimation of species richness. *Journal of Animal Ecology* **72**: 888 – 897.
- 1030 Ulrich, W., and M. Ollik. 2005. Limits to the estimation of species richness: The use of relative
1031 abundance distributions. *Diversity and Distributions* **11**: 265-273.
- 1032 Ulrich, W., M. Ollik, and K. I. Ugland. 2010. A meta-analysis of species - abundance distributions.
1033 *Oikos* **119**: 1149-1155.

- 1034 Watson, H. C. 1835. Remarks on the geographical distribution of British plants. Longman, Rees, Orme,
1035 Brown, Green and Longman, London.
- 1036 Xu, H., S. Liu, Y. Li, R. Zang, and FL. He. 2012. Assessing non-parametric and area-based methods for
1037 estimating regional species richness. *Journal of Vegetation Science* **23**: 1006-1012.
- 1038 Zillio, T., I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan. 2005. Spatial scaling in model plant
1039 communities. *Physics Review Letters* **95**: 098101.

1040

1041

1042

1043

1044 **Data Availability**

1045 Data associated with this study are available from the Research Data Leeds Repository:

1046 <https://doi.org/10.5518/264>

Author Manuscript

1047 **Figure legends:**

1048

1049 **Figure 1.** The location of CS survey sites and Atlas cells, and of the regional subsets used in
1050 the analyses. The number of samples in each region are indicated in the legend. A
1051 hypothetical 1×1 km focal landscape is shown at higher magnification on the right,
1052 containing X-plots and Linear samples (not to scale), and the nature of (multi-scaled) “X-
1053 only” and (composite) “X+Linear” samples is displayed.

1054

1055 **Figure 2.** Conceptual groupings of the methods employed. See text for further explanation.

1056

1057 **Figure 3.** Model predictions for the full UK dataset, based on (a, b) X-only samples, (c)
1058 X+Linear samples, and (d) randomised subsets. For clarity, a subset of the best-fitting
1059 models are plotted in (b), with an expanded Y-axis. Note that several models (MaxEnt, and
1060 fitted versions of Power and Logarithmic models) could not be estimated on X+Linear
1061 samples (see text and Figures 4 and 5). Plots in (d) represent means of X-only and X+Linear
1062 data from both WS and ND samples. Error distributions around each curve (with matching
1063 line color) represent trimmed ranges: the central 18 of the 20 datapoints (roughly
1064 corresponding to 90% confidence intervals). The true SAR is indicated by bold lines in each
1065 panel, for clarity.

1066

1067 **Figure 4.** Compilation of Total Species Richness fits of the various upscaling models
1068 tested. Figures represent proportional absolute errors $[|S_{\text{predicted}} - S_{\text{true}}| / S_{\text{true}}]$, with
1069 underscored numbers indicating the best and second-best fitting model for a particular
1070 dataset. Combined models are underscored relative to the set of individual models.
1071 Shading represents fit, with cutoff values 0.05, 0.1, 0.25, 0.5, and 1. Rank correlation
1072 coefficients (Spearman’s ρ) for the relationship between true and estimated richness are
1073 listed in the final row.

1074

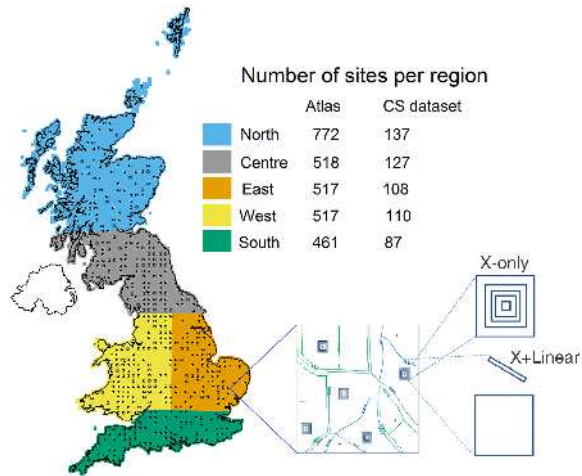
1075 **Figure 5.** Quality of SAR fit, as indicated by mean relative absolute error. Underscores
1076 indicate the best and second best models for each dataset, as in Figure 4. Shading is as in

1077 Figure 4, to aid comparison. The final row lists Spearman's rank correlation coefficients
1078 between true and estimated SAR slopes across the different datasets tested.

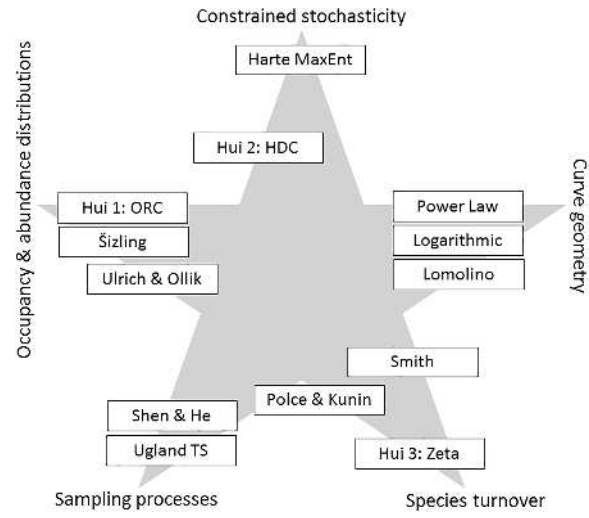
1079

1080 **Figure 6.** Variation in statistical subsample runs. For each model, the Coefficient of
1081 Variation (Standard error/mean) is given for both Wide-shallow and Narrow-deep subsample
1082 sets. Shading reflects CV values, with cutoff values of 0.01, 0.03, 0.1 and 0.3. "Ratio
1083 WS:ND" indicates the CV of Narrow-deep divided by that of Wide-shallow samples. The
1084 mean value of subsample projections relative to those of the full sample set are indicated as
1085 "relative

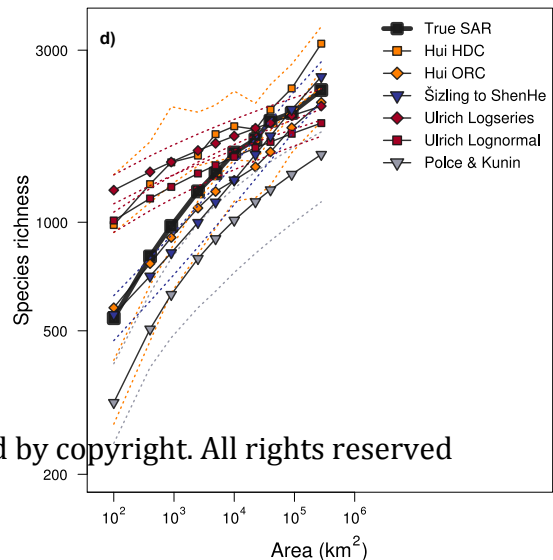
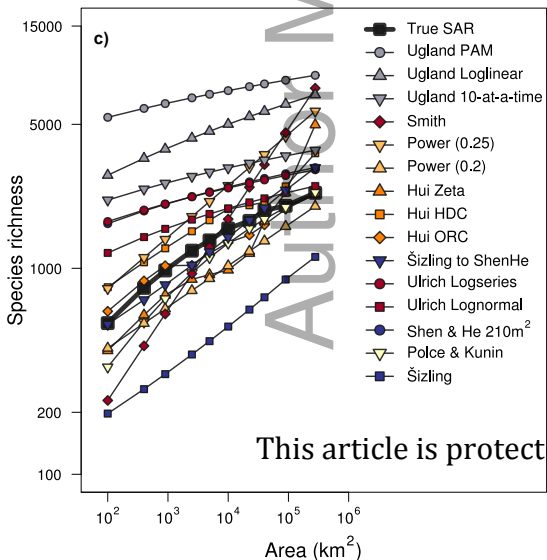
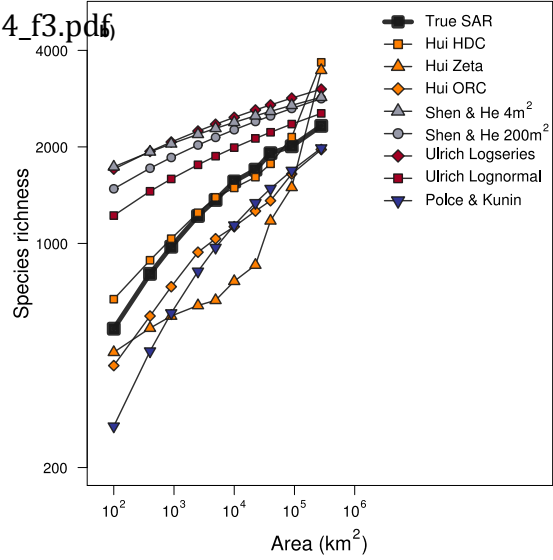
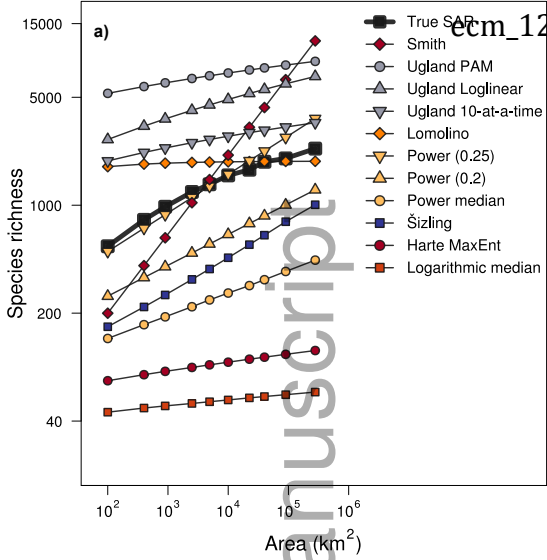
Author Manuscript



ecm_1284_f1.tif



ecm_1284_f2.tif



Model:		Harte MaxEnt	Hui HDC	Hui ORC	Hui Zeta	Logarithmic Median	Lomolino	Polce & Kunin	Power 0.2	Power 0.25	Power median	Shen He 4 m ²	Shen He 200/210 m ²	Šizling	Smith	Ugland: Loglinear	Ugland: 10-at-a-time	Ugland: PAM	Ulrich Ollik lognormal	Ulrich Ollik logseries	SH+UO1 Mix	SH+UO2 Mix	OU1+OU2	
X-only	Full British	0.951	0.578	0.156	0.490	0.974	0.172	0.148	0.459	0.566	0.810	0.228	0.216	0.567	3.984	1.934	0.471	2.673	0.094	0.301	0.155	0.258	0.198	
	Subset	Wide-shallow	0.952	0.282	0.145	0.402	0.973	0.400	0.335	0.454	0.566	0.811	0.049	0.069	0.714	13.533	1.987	0.195	1.629	0.185	0.090	0.058	0.010	0.138
		Narrow-deep	0.951	0.242	0.143	1.144	0.973	0.508	0.425	0.453	0.566	0.808	0.023	0.003	0.735	3.813	1.578	0.111	1.424	0.179	0.084	0.091	0.044	0.132
	Regional	South	0.955	0.282	0.219	0.035	0.976	0.558	0.580	0.617	0.003	0.862	0.051	0.051	0.738	0.848	1.365	0.067	1.391	0.258	0.159	0.155	0.105	0.209
		East	*	0.014	0.135	0.357	0.980	0.571	0.493	0.758	0.363	0.856	0.245	0.148	0.772	0.191	1.330	0.055	1.309	0.327	0.215	0.238	0.182	0.271
		West	0.951	0.033	0.215	0.488	0.974	0.519	0.478	0.560	0.158	0.857	0.166	0.132	0.740	0.673	1.161	0.004	1.282	0.302	0.228	0.217	0.180	0.265
		Centre	0.946	0.026	0.322	0.527	0.971	0.282	0.307	0.511	0.289	0.830	0.091	0.029	0.699	0.313	1.614	0.130	1.634	0.189	0.091	0.109	0.060	0.140
	North	0.916	0.214	0.239	0.389	0.951	0.511	0.551	0.200	1.149	0.704	0.026	0.061	0.686	2.080	0.947	0.109	1.595	0.147	0.052	0.104	0.057	0.099	
Full British		0.556	0.015	1.138		0.060	0.004	0.140	1.490			0.325	0.512	2.219	1.992	0.608	2.717	0.079	0.294	0.202	0.309	0.186		
X + Linear	Subset	Wide-shallow	0.436	0.006	0.832		0.317	0.240	0.139	1.490			0.178	0.666	13.767	2.154	0.368	1.823	0.182	0.088	0.002	0.045	0.135	
		Narrow-deep	0.417	0.009	0.761		0.437	0.350	0.139	1.490			0.111	0.682	3.673	1.704	0.276	1.732	0.214	0.127	0.051	0.008	0.171	
	Regional	South		0.381	0.161	0.424		0.532	0.558	0.369	0.668			0.018	0.688	0.698	1.243	0.182	1.776	0.255	0.157	0.118	0.070	0.206
		East		0.166	0.252	0.500		0.452	0.340	0.540	0.222			0.051	0.720	0.168	1.628	0.127	0.781	0.333	0.219	0.192	0.135	0.276
		West		0.144	0.107	0.386		0.439	0.436	0.331	0.778			0.037	0.692	0.862	1.271	0.128	0.651	0.308	0.231	0.173	0.134	0.269
		Centre		0.287	0.194	0.046		0.194	0.274	0.235	1.035			0.091	0.633	0.540	1.806	0.305	0.900	0.204	0.098	0.056	0.003	0.151
North		0.300	0.174	0.134		0.439	0.485	0.192	2.234			0.041	0.637	3.574	0.967	0.225	1.898	0.125	0.042	0.042	0.0004	0.084		
Overall: Mean (SD)		0.972 (0.145)	0.272 (0.173)	0.156 (0.089)	0.503 (0.328)	0.972 (0.009)	0.400 (0.152)	0.375 (0.158)	0.381 (0.192)	0.817 (0.615)	0.817 (0.050)	0.110 (0.091)	0.097 (0.085)	0.680 (0.067)	3.183 (4.315)	1.543 (0.375)	0.210 (0.162)	1.576 (0.572)	0.211 (0.080)	0.155 (0.083)	0.122 (0.070)	0.100 (0.093)	0.183 (0.063)	
Rank correl.		0.074	0.800	0.697	0.752	0.146	0.576	0.728	0.121	0.261	0.122	0.805	0.825	0.600	0.661	0.764	0.788	0.679	0.655	0.655	0.782	0.764	0.655	

Model:		Harte MaxEnt	Hui HDC	Hui ORC	Hui Zeta	Logarithmic Median	Lomolino	Polce & Kunin	Power 0.2	Power 0.25	Power median	Shen He 4 m ²	Shen He 200/210 m ²	Sizling	Smith	Ugland: Loglinear	Ugland: 10-at-a-time	Ugland: PAM	Ulrich Ollik lognormal	Ulrich Ollik logseries	PK+UO1 Mix	U1+U2 mean	Sizling to SH	
X-only	Full British	1.296	<u>0.125</u>	0.238	0.400	0.956	0.601	0.294	0.553	<u>0.168</u>	0.809	0.808	0.669	0.683	1.042	2.446	1.094	4.470	0.446	0.848	<u>0.088</u>	0.647	0.168	
	Statistical	Wide-shallow	0.927	<u>0.100</u>	<u>0.137</u>	0.473	0.957	0.405	0.335	0.548	0.168	0.807	0.548	0.473	0.765	3.357	2.503	0.725	2.995	0.229	0.336	0.159	0.275	<u>0.137</u>
		Narrow-deep	0.926	0.335	<u>0.218</u>	0.580	0.957	0.381	0.466	0.548	<u>0.168</u>	0.808	0.449	0.377	0.779	0.998	2.063	0.608	2.686	0.232	0.341	0.198	0.281	<u>0.169</u>
	Regional	South	0.944	0.182	<u>0.144</u>	0.378	0.968	0.399	0.520	0.670	0.230	0.864	0.199	<u>0.148</u>	0.779	0.519	1.525	0.302	2.108	0.176	0.170	0.315	0.165	0.200
		East	*	0.249	<u>0.180</u>	0.380	0.972	0.400	0.555	0.786	0.500	0.861	0.216	<u>0.171</u>	0.810	0.580	1.550	0.237	2.072	0.231	0.200	0.367	0.205	0.289
		West	0.932	<u>0.094</u>	0.253	0.573	0.963	0.358	0.441	0.600	<u>0.137</u>	0.813	0.236	0.198	0.769	0.447	1.484	0.330	1.929	0.211	0.214	0.269	0.211	0.206
		Centre	0.922	<u>0.150</u>	0.284	0.642	0.954	0.429	0.244	0.529	<u>0.108</u>	0.837	0.351	0.337	0.721	0.379	2.166	0.600	2.784	0.237	0.328	0.161	0.276	<u>0.102</u>
	North	0.860	0.700	<u>0.252</u>	0.270	0.913	0.440	0.358	<u>0.193</u>	0.873	0.646	0.619	0.494	0.678	0.709	1.741	0.792	2.170	0.411	0.525	0.268	0.463	<u>0.098</u>	
Full British		0.293	<u>0.106</u>	0.345		0.770	<u>0.172</u>	0.289	0.726			0.837	0.635	0.646	2.581	1.304	4.556	0.416	0.834	0.123	0.625	<u>0.120</u>		
X + Linear	Statistical	Wide-shallow		0.476	<u>0.132</u>	0.355		0.463	<u>0.229</u>	0.289	0.726		0.638	0.716	3.603	2.751	0.982	3.301	0.231	0.338	0.138	0.278	<u>0.083</u>	
		Narrow-deep		0.447	<u>0.128</u>	0.393		0.392	0.383	0.289	0.726		0.547	0.726	0.990	2.267	0.854	3.160	<u>0.225</u>	0.308	0.183	0.257	<u>0.095</u>	
	Regional	South		0.209	<u>0.145</u>	0.546		0.377	0.487	0.457	0.285		0.208	0.730	0.451	1.446	0.448	1.384	0.189	<u>0.171</u>	0.303	0.172	<u>0.134</u>	
		East		<u>0.130</u>	0.219	0.406		0.337	0.400	0.594	<u>0.144</u>		0.206	0.760	0.526	1.897	0.443	1.396	0.236	0.200	0.294	0.207	0.189	
		West		<u>0.210</u>	<u>0.167</u>	0.422		0.338	0.374	0.392	0.432		0.254	0.721	0.446	1.638	0.498	2.176	0.214	0.212	0.249	0.211	<u>0.108</u>	
		Centre		0.450	<u>0.103</u>	<u>0.224</u>		0.502	0.175	0.263	0.728		0.506	0.654	0.352	2.434	0.854	3.266	0.230	0.320	0.164	0.267	<u>0.059</u>	
	North		0.721	<u>0.125</u>	<u>0.137</u>		0.478	0.289	0.207	1.819		0.634	0.615	1.347	1.829	0.988	4.608	0.440	0.565	0.281	0.497	<u>0.115</u>		
Overall: Mean (SD)	0.972 (0.145)	0.304 (0.202)	<u>0.177</u> (0.059)	0.408 (0.134)	0.955 (0.018)	0.442 (0.110)	0.358 (0.118)	0.451 (0.179)	0.496 (0.444)	0.807 (0.069)	0.428 (0.219)	0.418 (0.212)	0.721 (0.056)	1.024 (1.000)	2.020 (0.438)	0.691 (0.309)	2.816 (1.043)	<u>0.272</u> (0.094)	0.369 (0.217)	0.222 (0.081)	0.315 (0.155)	<u>0.156</u> (0.062)		
Slope correl.	-0.037	-0.576	-0.497	-0.164	0.756	0.261	-0.146	0	0	0.244	-0.195	-0.176	-0.115	-0.361	-0.036	-0.042	-0.194	-0.194	-0.006	-0.097	-0.152	-0.115		

Model:			Harte MaxEnt	Hui HDC	Hui ORC	Hui Zeta	Logarithmic Median	Lomolino	Polce & Kunin	Power 0.2	Power 0.25	Power median	Shen He 4 m ²	Shen He 200/210	Sizzling	Smith	Ugland: Loglinear	Ugland: 10-at-a-time	Ugland: PAM	Ulrich & Ollik lognormal	Ulrich & Ollik logseries
X-only	CV:	Wide-shallow	0.0133	0.1186	0.1716	0.4156	0.0661	0.0243	0.0522	0.0448	0.0448	0.1271	0.0134	0.0182	0.0477	0.1608	0.0221	0.0160	0.0611	0.0170	0.0073
		Narrow-deep	0.0266	0.0989	0.2063	0.3275	0.0419	0.1389	0.1260	0.0336	0.0336	0.1181	0.0774	0.0780	0.0655	0.1688	0.1081	0.0779	0.0730	0.0742	0.0767
	Ratio ND:WS		1.9938	0.8334	1.2024	0.7881	0.6345	5.7163	2.4123	0.7495	0.7495	0.9289	5.7947	4.2958	1.3751	1.0501	4.8800	4.8769	1.1941	4.3605	10.477
	Rel. to whole	Wide-shallow	0.9901	1.0000	1.1743	0.8662	1.0044	0.7346	0.9629	1.0098	1.0000	1.0073	0.8556	0.8821	0.7470	2.1438	1.0168	0.8216	0.7276	0.7753	0.7093
		Narrow-deep	1.0058	1.1977	1.0261	0.8918	1.0040	0.6041	0.7691	1.0099	1.0001	1.0000	0.7982	0.8237	0.7041	0.9692	0.8878	0.7655	0.6712	0.7817	0.7138
X+Linear	CV:	Wide-shallow		0.1350	0.0829	0.2541		0.0423	0.0443	0.0169	0.0169			0.0200	0.0150	0.4212	0.0199	0.0109	0.0412	0.0086	0.0037
		Narrow-deep		0.1352	0.1449	0.2969		0.0922	0.0973	0.0341	0.0341			0.0595	0.0558	0.1624	0.0783	0.0593	0.1185	0.1327	0.1652
	Ratio ND:WS			1.0021	1.7482	1.1686		2.1787	2.1972	2.0198	2.0198			2.9760	3.7108	0.3857	3.9284	5.4218	2.8795	15.467	44.585
	Rel. to whole	Wide-shallow		1.1440	1.1367	0.9677		0.7394	0.9500	1.0000	1.0000			0.8911	0.7842	2.9627	1.0483	0.8584	0.7714	0.7947	0.7162
		Narrow-deep		1.1190	0.9285	0.8895		0.6105	0.7576	1.0000	1.0000			0.8415	0.7589	1.2934	0.9115	0.8025	0.7461	0.7664	0.6863