ESSAY REVIEW

Using local–regional richness plots to test for species saturation: pitfalls and potentials

DIANE S. SRIVASTAVA*

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

Summary

1. Plots of local versus regional species richness are an exciting new tool for testing for species saturation in ecological communities. In this method, the local richness of a community is plotted as a function of its regional richness for different biogeographical regions. A proportional relationship between local and regional richness is interpreted as evidence for an unsaturated community, that is, a community with strong evolutionary limits to local richness. There will be no correlation between local and regional richness in a saturated community, that is, a community whose local species richness is limited largely by ecological processes.

2. Although at least 36 data sets have now been analysed using local–regional richness plots, there has not been much critical evaluation of the method. This paper provides such a critique, focusing on the selection of communities for comparison, the prevalence of pseudoreplication and multiple null models, and the effects of differing region size.

3. Local–regional richness plots are best suited for comparing similar habitats between different regions, not different habitats in a single region. In the latter, taxa effects and species pool effects are confounded.

4. Four very different types of local–regional richness plots have been published. Each type of plot has important underlying assumptions which are often not addressed by ecologists.

5. Of the 36 data sets reviewed in this paper, 13 were spatially pseudoreplicated, and 2 were temporally pseudoreplicated. Furthermore, ecologists differ in their choice of null model, with the result that the same local–regional plot could be interpreted as evidence for saturation by one ecologist, and for lack of saturation by another.

6. Differences in region size can result in pseudosaturation, the appearance of saturation by an unsaturated community. A simple model demonstrates this phenomenon. Other sources of error in estimating the regional species pool are also of concern.

7. In conclusion, local–regional richness plots are a potentially useful tool for distinguishing saturated from unsaturated communities, but should be used cautiously, and in conjunction with other supporting evidence (such as the presence or absence of competitive exclusion, resource limitation, density compensation and the effects of species invasions).

Key-words: pseudoreplication, region size, species pool, species richness, unsaturated communities.

Journal of Animal Eccology (1999) **68**, 1–16

Introduction

Broadly speaking, species richness at a local scale could be limited by either ecological or evolutionary causes. Ecological limitations result in species being actively excluded from communities by local effects such as upper limits to niche packing and minimum viable population sizes. Increases in the regional species pool have little impact on this ecological ceiling to local richness, and the community is said to be saturated with species. Evolutionary limitations involve

*Present address: Centre for Biodiversity Research, Faculty of Science, University of British Columbia, 6270 University Blvd., Vancouver BC, Canada V6T 1Z4.

strong regional controls on local community richness, so key parameters are rates of speciation and the fixation rate of mutations which allow new habitats (or hosts) to be invaded. Unsaturated communities are those whose local richness tracks that of the regional species pool. Evolutionary limitations were favoured as explanations of local diversity in the first part of this century, whereas ecological limitations have been emphasized in the last few decades (Kingsland 1985; Ricklefs 1987; Schluter & Ricklefs 1993), but until recently there have been few attempts to formally compare the two perspectives.

Accompanying this recent synthesis of perspectives (e.g. Ricklefs 1987; Cornell & Lawton 1992; Ricklefs & Schluter 1993) is the development of a new technique to separate saturated and unsaturated communities. In local–regional richness plots, the local and regional richness of a habitat is compared between different biogeographical regions. Local richness is the dependent variable, and regional richness the explanatory variable in this analysis. In a saturated community, local richness is anticipated to already be at the maximal level permitted by ecological constraints, like limits to niche packing, so will be largely independent of regional richness. In an unsaturated community, local richness is anticipated to be directly limited by regional richness, and therefore proportional to regional richness. This approach was first explicitly outlined by Terborgh & Faaborg (1980), although Pearson (1977) cited as evidence of the importance of regional history, the similarity in rank order of the local and regional richnesses of tropical forest bird communities. Further refinements to the analysis of local–regional plots came with studies on Californian gall wasps by Cornell (1985a,b), and then by a number of studies in the current decade which have used the technique.

Perhaps because it is still such a young technique, researchers have used a bewildering array of different methods for constructing and analysing local–regional plots. This paper reviews and assesses these different approaches. Although there have been useful reviews of the history (Ricklefs 1987; Schluter & Ricklefs 1993) and theoretical underpinnings of local–regional plots (Ricklefs 1987; Cornell & Lawton 1992), there has not yet been a comprehensive review of the methodology of local–regional plots. Only two other articles focus on methodology (Cresswell, Vidal-Martinez & Crichton 1995 discuss some statistical issues; Caley & Schluter 1997 examine the effect of locality size) and both complement rather than overlap with this review. Many previous studies do, however, include useful observations on methodology, and this paper both synthesizes and builds on the insights of these authors. It begins with an overview of the theory behind local–regional richness plots and discusses which questions such plots are most suited to answering, before examining in detail four very different methods of constructing local–

regional plots. Special attention is given to spatial pseudoreplication and variation in region area, both of which may cause artefactual results. This review concludes with a summary of avoidable and unavoidable pitfalls, and suggestions about appropriate supportive evidence.

Definitions of local and regional richness

Local and regional richness are differentiated by spatial scale. Local richness is measured on a spatial scale small enough that all the species could encounter each other within ecological time, and so possibly interact. This is the scale which community ecologists usually consider. Examples of local richness are the number of fish species in a lake, the number of grass species in a meadow and the number of gut parasites in a host. Regional richness, or the richness of the species pool, is measured on a larger spatial scale. The regional species pool contains all the species which could eventually colonise a location if competitive exclusion was unimportant. Dispersal of species within a region may be slow in ecological time, but is substantially greater than rates of speciation, host-shifting or long-distance dispersal which affect differences in regional richness between regions. Examples of regional richness are the number of fish species in Britain, the grassland flora of the Serengeti, and the total number of parasite species recorded from the guts of black bears. Neither of these definitions is particularly concrete, and the implications of the ambiguity for local–regional plots is discussed later.

The rationale of local–regional plots

If the local and regional richness of a standard habitat is compared across several regions, the possible outcomes are bounded by two possibilities (Fig. 1a). Either there is a linear relationship between local and regional richness, or local richness is generally constant over a wide range of regional richnesses (except as very low levels of regional richness). In the first scenario, local richness appears to be determined primarily by regional richness, and the community is unsaturated. In the second scenario, local richness is limited by other (i.e. ecological) factors, and the community is saturated. Somewhere between these two boundaries is a set of curvilinear functions, which reflect neither absolutely saturated nor unsaturated communities, but rather communities approaching saturation but still influenced by historical effects (Cornell & Lawton 1992). Many authors interpret curvilinear functions as evidence for saturation since local and regional richness will be uncorrelated at least in some regions. Other authors interpret any correlation, both linear and curvilinear, between local and regional richness as evidence for unsaturated communities since local richness is dependent on regional richness at least in some regions. This is lar-
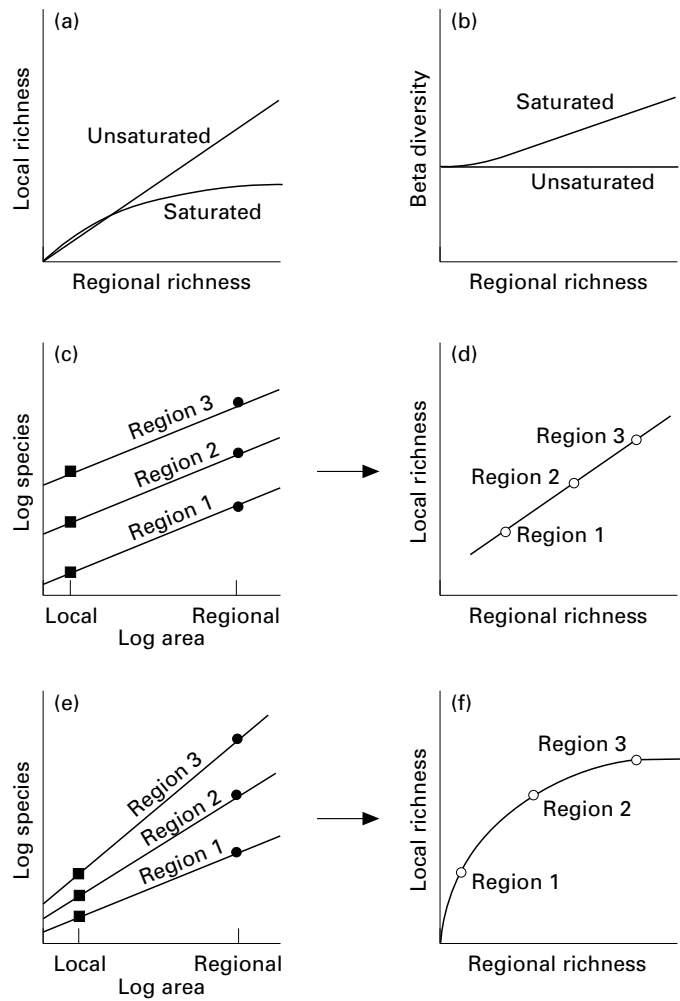
Fig. 1. In unsaturated communities, local richness is predicted to be a fixed proportion of regional richness. In saturated communities, local richness may increase with regional richness at very low levels of regional richness, but will quickly approach an upper limit at higher regional richnesses (a). Consequently, beta diversity is similar between regions in an unsaturated community, but will increase with regional richness in a saturated community (except at very low levels of regional richness) (b). In an unsaturated community, parallel lines will connect the local richness (solid squares) and regional richness (solid circles) of each region on a log species-log area plot (c). When species–area curves are parallel in log–log space, the ratio of local to regional richness will be similar between regions, yielding an unsaturated local–regional plot (d). In a saturated community, as regional richness increases, the slope of the log species–log area lines will also become steeper (e). As this log–log slope increases, the ratio between local and regional richness becomes smaller, yielding the curvilinear local–regional plot typical of saturated communities (f).

gely a question of semantics for communities summarized by curvilinear local–regional functions grade from unsaturated to saturated: the differences between the two views do, however, have implications for the statistical analysis of local–regional richness plots which I describe later. In this review, I take the majority view that the difference between linear and curvilinear functions is of ecological interest and that the latter is evidence for saturation.

The slope of these local–regional plots should be identical to:

$$1/(1 + \text{beta diversity}) \tag{1}$$

using Whittaker's definition (Whittaker 1960) of beta diversity as:

$$(\text{regional richness}/\text{local richness}) - 1 \tag{2}$$

Beta diversity is therefore not expected to vary amongst regions in an unsaturated community, as regions have the same ratio of local to regional richness. In saturated communities, increases in regional richness will be accompanied by increases in beta diversity (Fig. 1b; Cornell & Lawton 1992).

Local and regional scales are distinguished on the basis of area. As area increases from the local to the regional scale in any one region, the species richness of that area is expected to increase according to the standard species–area curve, a power function. In a given region i, therefore, local species richness (S_L) will be related to local area (A_L) by the equation:

$$\log S_{L,i} = z_i(\log A_{L,i}) + c_i \tag{3}$$

and regional richness (S_R) will be related to regional area (A_R) by the equation:

$$\log S_{R,i} = z_i(\log A_{R,i}) + c_i \qquad (4)$$

where c_i and z_i are constants. Although the species–area power function may break down over very large scales, this typically only occurs at extremely small local areas or areas larger than regions (Rosenzweig 1995). For this region i, the ratio of local to regional richness ($S_{L,i}/S_{R,i}$) is therefore related to the exponent, z_i, of the species–area curve as follows:

$$\log (S_{L,i}/S_{R,i}) = z_i(\log A_{L,i} - \log A_{R,i}) \qquad (5)$$

If we assume that $A_{L,i}$ and $A_{R,i}$ are the same for each region, then the ratio of local to regional richness will be identical between regions when z_i is identical between regions (the situation for different $A_{R,i}$ is examined later). In other words, unsaturated communities should have parallel species–area curves in log–log space, but these lines can differ in the intercept, c_i (Fig. 1c,d). Saturated communities will differ in z_i between regions (z_i will increase with regional richness) but have minimal differences in c_i (Fig. 1e,f). This representation of the local–regional richness plot has been used previously by Westoby (1993), and is the basis of some of the assessment of methodology in this review.

Which question? Within-habitat or between-habitat comparisons?

There are two profoundly different types of local–regional richness plots in the literature. The first type compares the same habitat in different regions (within-habitat comparisons), and is the main focus of this paper. The second type compares different habitats in the same region (between-habitat comparisons), and examples are just starting to be published (Pärtel *et al.* 1996). One would imagine that two such different approaches would have very different aims, but both attempt to disentangle ecological and evolutionary limits to local diversity. In the between-habitat comparison, regional control of the local richness of different habitats has been dubbed the 'species pool' hypothesis by Taylor, Aarssen & Loehle (1990), and the use of local–regional correlations to test it has been advocated by Eriksson (1993). Although the species pool hypothesis may be a useful null model for community ecologists, the simple local–regional plot is inappropriate and insufficient to test this hypothesis, as illustrated by the following example.

Pärtel *et al.* (1996) describe the local and regional richness of 14 different plant communities in Estonia (8 forest types, 4 grassland types, 1 bog, and 1 heath). The ratio of local to regional richness is roughly the same for each community type, which the authors take as evidence that local richness is 'largely determined by the regional species pool'. I suggest that this conclusion is more appropriate for a single habitat study, than for one which incorporates 14 different communities in a single region. Correlation never

equals causation, but correlation suggests causation when every other possible cause has been ruled out. In the case of multiple habitats, every other cause has not been ruled out. Forests are different from grasslands in many ways, and some of these differences may hold across different spatial scales. For example, whatever the spatial scale, the density of plants is usually lower in forests than grasslands, simply because trees are bigger than grass tillers. If species richness is related to the number of individuals, it is not particularly surprising that grasslands have more species than forests on both local and regional scales (especially as local richness was determined for a larger area for grasslands than forest communities in Pärtel *et al.*'s study). When there are such credible alternative hypotheses on offer, there needs to be additional evidence of evolutionary and historical limits to local diversity before any conclusion can be reached. Such evidence might include linear within-habitat local–regional plots (comparing species richness of a habitat across regions) indicating that each community type is unsaturated. Obviously, if most of the communities are saturated, the 'species pool' hypothesis cannot hold, and any between-habitat correlations between local and regional richness must be artefactual (Fig. 2; see also Method 2 section below).

Explanations aside, it is clear that Estonian forests and grasslands have roughly the same ratio of local to regional richness. According to the mathematical reasoning presented in the above section, this should mean that the exponent of their species–area curves are similar, which is indeed the case, as can be seen by replotting the data (Fig. 3). Perhaps the more interesting question, then, is why species tend to accumulate across areas in similar ways for very different communities.

In summary, between-habitat comparisons of local and regional richness are inappropriate tests for saturation. The remainder of this paper concentrates on within-habitat comparisons.

Four general methods of testing for saturation

Assuming that one wishes to test for saturation within a single type of community, the procedure is as follows. Several regions are identified whose species pools are as independent as possible from each other. The local richness of the community is determined in each region, as is the regional species pool. Local and regional estimates for each region are plotted against each other, and the relationship between the two is statistically analysed. In general terms, the procedure is straightforward. In practical terms, there are many ways to define regions and their species pools, as described below, and many ways to analyse the results as described in the next section. There are four general ways to define regions and their species pools:
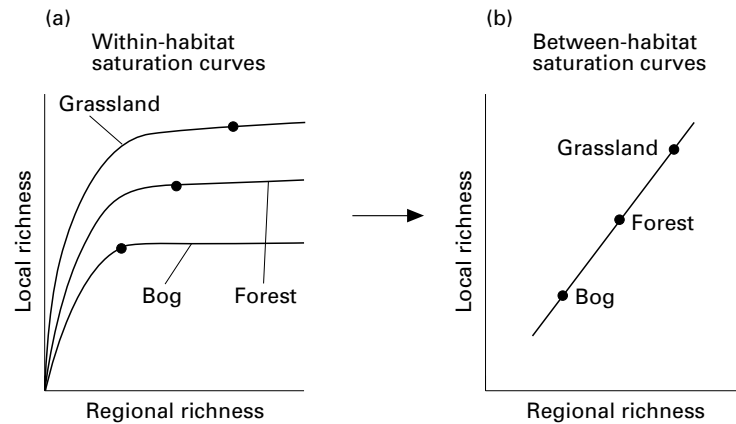
Fig. 2. The forests, grasslands, and bog examined by Pärtel *et al.* (1996) in Estonia could each be saturated communities, as shown by within-habitat plots of local vs. regional richness (a), but still show a linear relationship between local and regional richness when plotted together in a between-habitat analysis (b).
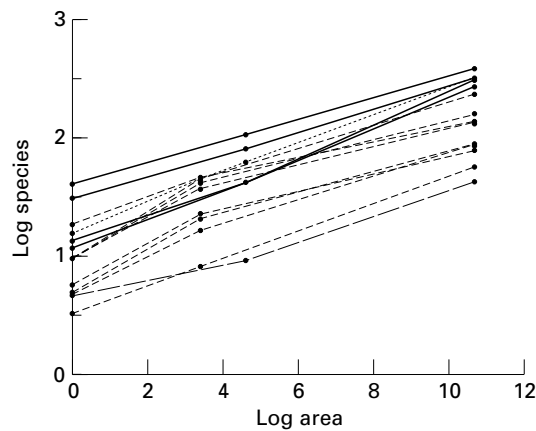
Fig. 3. Species–area relationships, plotted in log–log space, for 14 Estonian plant communities (medium dashed line = forests, solid line = grasslands, very long dashed line = bog, dotted line = heath). The largest areas are the region areas (approximated by the area of Estonia), the smallest areas are 1 m² plots, and the two intermediate areas are local areas. Data are from Pärtel *et al.* (1996).

Method 1. In each of several islands, or other geographically distinct areas, obtain local species richness in several different habitats: there are multiple dependent variables as the data for each habitat is analysed separately. Use the total species richness for each island, undifferentiated into habitats, as the independent variable (e.g. Terborgh & Faaborg 1980; Ricklefs 1987).

Method 2. In one geographic area, examine richness on analogous but not identical habitats, such as insects on taxonomically related host-plant species or parasites on congeneric fish. Local and regional richness are analysed for these different host species, which act as geographically overlapping regions (e.g. Cornell, 1985a,b; Aho & Bush 1993).

Method 3. In each of several geographically distinct regions, obtain local and regional richness in a single, carefully defined habitat (e.g. Lawton 1990; Pearson & Juliano 1993).

Method 4. In a given geographic region, the local and regional richness of a particular community is compared through evolutionary time (van Valkenburgh & Janis 1993).

Each of these methods is assessed below, using examples summarized in Table 1.

METHOD 1: REGIONAL POOL SUMS OVER HABITATS

In the first method, the local richness of each habitat is compared to the number of species which occur in a region *irrespective of habitat*. In other words, the 'regional pool' of a habitat can contain species which could never occur in that particular habitat. This rather counterintuitive definition of the species pool is found only in three studies, all analyses of the bird assemblages of Caribbean islands (Terborgh & Faaborg 1980; Ricklefs 1987; Wiens 1989, reanalysing Terborgh & Faaborg's data). Most other studies, by contrast, restrict the species pool to those species which could potentially occur in a given habitat (see definitions of Cornell & Lawton 1992; Eriksson 1993; Zobel 1997, etc.). According to this widely accepted definition of the species pool, the Caribbean island studies overestimate the true species pool for each habitat. It could be argued that many bird species are capable of existing in a variety of habitats; indeed these same Caribbean islands have provided some of the best evidence for ecological release (Terborgh & Faaborg 1973; Cox & Ricklefs 1977; Terborgh, Faaborg & Brockmann 1978; Terborgh & Faaborg 1980). It is unlikely, though, that all species on an island are capable of existing in every habitat on that island. Some species that are restricted to clearings

Table 1. A summary of published local–regional richness plots, indicating the method used (see text), number of regions, and type of statistical analysis (pseudoreplication indicated in all cases where it appears). Curvilinear regressions include polynomial and log–log regressions. Studies which test for saturation but do not explicitly use the local–regional plot method are not included, with the exception of Pearson (1977) and Bohnsack & Talbot (1980); these two early studies capture the spirit of the method closely enough to warrant inclusion. Communities are classified as either saturated and unsaturated, as reported by the authors of each study: note that authors differ in definitions of these terms (see text)

Community	Study	Method	Regions	Statistical analysis
(a) Evidence for unsaturated communities				
Birds of tropical dry forests	Pearson 1977, analysed in this review*	3	6	Originally none, now linear vs. curvilinear regressions
Cynipine gall wasps on Californian oak	Cornell 1985a,b	2	8	Linear vs. curvilinear regressions, pseudoreplicated
Songbirds on Caribbean islands	Ricklefs 1987	1	5	None given
Insects on Asteraceae	Zwölfer 1987, interpreted by Cornell & Lawton 1992, as unsaturated	2	14	Correlation, no comparison with curvilinear regressions
Insects on Asteraceae	Lewinsohn 1991	2	44	Linear regression (no curvilinear regressions)
Wood-boring scolytid beetles in eastern USA	Cornell & Lawton 1991, analysed data from Stevens 1986	2	40	Linear vs. curvilinear regressions; pseudoreplicated
Fig wasps in South Africa	Hawkins & Compton 1992	2	15	Linear vs. curvilinear regressions
Parasitoids of fig wasps in South Africa	Hawkins & Compton 1992	2	15	Linear vs. curvilinear regressions
Bracken herbivores, world-wide	Lawton *et al.* 1993	3	4	None possible (range of values)
Pimplines (Hymenoptera), world-wide	Gaston & Gauld 1993	3	5	None possible (range of values)
Coral reef fish of the Great Barrier Reef and the Caribbean	Westoby 1993	3	2	None possible (2 data points)
Arid-zone lizards of Australia and North America	Westoby 1993	3	2	None possible (2 data points)
Deep-sea gastropods	Stuart & Rex 1994	3	5 & 6	Linear regression (no curvilinear regressions), pseudoreplicated
Helminth parasites of introduced British fish	Kennedy & Guégan 1994	2	7	Linear vs. curvilinear regressions
Parasitoids of British chalcid wasps	Dawah *et al.* 1995	2	>15	Linear regression, pseudoreplicated
Fish in riverine pools, Côte d'Ivoire	Hugueny & Paugy 1995	3	10	Linear vs. curvilinear regressions, pseudoreplicated
Hermatypic coral, world wide	Cornell & Karlson 1996	3	<40	Linear vs. curvilinear regressions, pseudoreplicated
Mussels of North American rivers	Vaughn 1997	3	14	Curvilinear regression only
Birds, world-wide	Caley & Schluter 1997	3	5	Linear regression only on this dataset, linear vs. curvilinear regressions on a larger multitaxa dataset
Mammals, world-wide	Caley & Schluter 1997	3	3	As above
Reptiles, world-wide	Caley & Schluter 1997	3	4	As above
Freshwater fish, world-wide	Caley & Schluter 1997	3	3	As above
Fish in rivers of North-Western France	(a) Oberdorff *et al.* 1998 (b) Belkessam *et al.* 1997	(a) 3 (b) 3	(a) 9 (b) 6	(a) & (b) Linear vs. curvilinear regressions, pseudoreplicated
(b) Evidence for saturated communities				
Reef fish, Australia and Florida	Bohnsack & Talbot 1980	3	2	None, or Mann–Whitney *U*-test (depending on which richness measure used), latter pseudoreplicated
Small-lake fish, Wisconsin and Finland	Tonn *et al.* 1990	3	2	Kruskal–Wallis *H* statistic, pseudoreplicated
Helminth parasites of amphibians, separately analysed for 7 genera	Aho 1990	2	3 to 5	ANOVA with linear regression, pseudoreplicated
Aquatic macrophytes in Norway	Rørslett 1991	3	11 or 18†	Linear regression, and ANOVA, pseudoreplicated
Parasites of North American fish, separately for two genera	Aho & Bush 1993	2	3 & 9	ANOVA & ANOVA with linear regression, pseudoreplicated
Sclerophyll woodland plants, Australia and South Africa	Westoby 1993	3	2	None possible (two data points)
Mammals in North America over last 44 million years	Van Valkenburgh & Janis 1993	4	25	Linear vs. curvilinear regressions (however, former has higher r^2), pseudoreplicated
Helminth parasites of endemic British fish	Kennedy & Guégan 1994	2	32	Linear vs. curvilinear regressions

Table 1. (*Continued*)

Community	Study	Method	Regions	Statistical analysis
Australian *Banksia*	Richardson *et al*. 1995	3	40	Curvilinear regression
North American lacustrine fish	Griffiths 1997	3	15	Linear vs. curvilinear regressions
(c) Evidence for both saturated and unsaturated communities				
Birds of Caribbean islands	(a) Terborgh & Faaborgh 1980 (b) reanalysed by Wiens 1989 (c) reanalysed by Griffiths 1997	1	12	(a) Saturated, no statistical analysis (b) Unsaturated, linear vs. curvilinear regressions (c) Saturated, linear vs. curvilinear regressions, constrained to pass through origin
Tiger beetles of India and North America	Pearson & Juliano 1993	3	2	Depends on habitat whether saturated or not; *t*-tests, pseudoreplicated

* See Appendix 2.

† 18 regions cited in text, but 11 regions according to F statistic.

and fields in species-rich islands still do not expand into the coastal scrub or rainforests of species-poor islands (Terborgh *et al.* 1978). Terborgh & Faaborgh (1980) suggest that a 'substantial number' of species may be committed habitat specialists.

Given that the species pool for each habitat is overestimated by this approach, does this affect the results? If the species pool of each region was overestimated by a constant or proportional amount, the difference between saturated and unsaturated local–regional plots would be maintained (see later section for a more detailed analysis of regional overestimation). Unfortunately, this assumption cannot be made. The total number of species per island could increase independently of a habitat's true species pool by either increasing habitat number (island richness and habitat number are strongly correlated: Fig. 2 in Terborgh & Faarborgh 1980; see also Caley & Schluter 1997), or by increasing the available species pool in another habitat.

METHOD 2: ANALOGOUS HABITATS IN THE SAME GEOGRAPHIC AREA

This is a popular method, first used by Cornell (1985a,b) for cynipine gall wasp communities on different oak species in California, and since used for various insect herbivore and animal parasite communities (Table 1). In this method, taxonomically related hosts serve as analogous, but obviously not completely identical, habitats for the herbivore/ parasite community. As the hosts exist in the same geographic area, the regions in the analysis are not geographically separated, unlike methods 1 and 3. The key assumption in this analysis is that communities on different hosts saturate in a similar way. Essentially, several saturation curves are being compared (one for each host species), and the less similar

these curves are, the more likely that differences in local–regional ratios between hosts just reflect differences in their saturation curves, not differences in the actual degree of saturation (Fig. 4). In some ways, this method represents one endpoint of the between-habitat comparison discussed earlier, and its validity is contingent on the habitats being so similar that objections of confounding factors disappear. This caveat has been recognized previously by Cornell (1993), who writes : 'A strong local–regional correlation may result simply from larger niche space on some host populations than on others, and all assemblages may be saturated within their individual niche species, some large, some small'. Godfray (pers. comm. in Hawkins & Compton 1992) also suggests that differences in host size may result in different saturation levels, confounding local–regional plots. Differences in host size have been explicitly factored into certain studies (Zwölfer 1987; Lewinsohn 1991), presumably for this reason. The majority of studies, however, do not mention the assumption of identical saturation curves underlying this method.

METHOD 3: IDENTICAL HABITATS IN GEOGRAPHICALLY DIFFERENT REGIONS

This method compares the local and regional richness of a particular habitat between several geographically distinct regions. It was first used by Lawton (1990), to compare bracken herbivores between several countries, and by Tonn *et al.* (1990), to compare fish communities between Wisconsin and Finland, and has since been used in numerous other studies. Analyses of plant communities are obviously restricted to this method (few plants are parasites). The same habitat is examined in different regions, rather than the several analogous habitats of method 2. It is assumed, however, that the habitat remains identical between

Within-host
saturation curves

Between-host
saturation curves
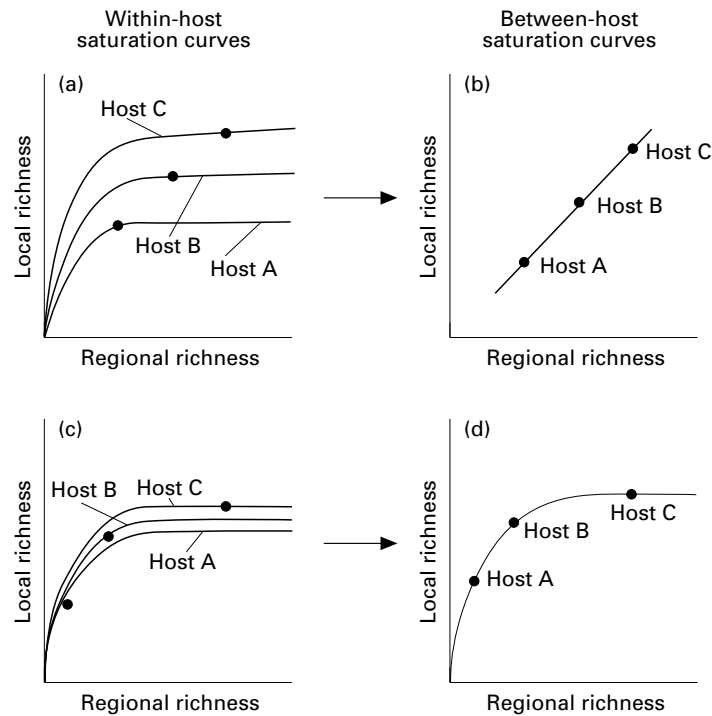
Fig. 4. Each host may have a very different saturation curve for parasite or herbivore species, which will affect the between-host saturation curve. As a result, even if all hosts are saturated with species (a), an unsaturated between-host saturation curve may occur (b). The converse is also possible. Only if the saturation curves for each host are nearly identical (c) will the between-host saturation curve reflect the true saturation status of the hosts (d).

regions (for example, that there is no effect of latitude on the number of available niches). Some authors have presented evidence supporting this assumption. Lawton (1982), for example, shows that bracken growth and patch area are similar between two regions, and argues that 'legitimate worries about climate, enemies or plant chemistry are arguments about "second order" effects...they may modify, but are not primarily responsible for, major differences in faunal richness between geographical regions' (see also Lawton, Lewinsohn & Compton 1993). Other studies have recognized that their habitats differ between regions, but have attempted to control for this by including an explanatory environmental variable in the analysis, for example, depth in deep sea prosobranch snail communities (Stuart & Rex 1994), depth and habitat type in coral reef communities (Cornell & Karlson 1996), river characteristics – as summarized by PCA scores–for French riverine fish (Oberdorff *et al.* 1998), habitat type for tiger beetles (Pearson & Juliano 1993) and season and downstream position in African fish communities (Hugueny & Paugy 1995).

METHOD 4: COMPARISONS IN A GIVEN REGION OF A PARTICULAR COMMUNITY THROUGH EVOLUTIONARY TIME

If a community is saturated with species, one would expect further additions to the regional species pool

to occur only with increases in beta diversity. That is, through evolutionary time, increases in regional richness will not be mirrored at the local scale. Obviously, this assumes that the community basically encounters the same sort of habitat through evolutionary time, and that the non-independence of data points is of minor importance (see comments on temporal pseudoreplication below). Despite such shortcomings, temporal slices of local–regional richness (such as van Valkenburgh & Janis 1993) can usefully complement the geographical slices of the above three methods.

Statistical analysis and pseudoreplication

Pseudoreplication is probably the most common statistical error in the analysis of local–regional plots. Pseudoreplication occurs when the replicates used in the analysis are not truly independent from each other, either because they are spatially correlated (e.g. within the same block) or temporally correlated (e.g. when repeated measures are used as replicates). I argue that when local–regional plots have been statistically analysed, the data are frequently pseudoreplicated. Of the 30 datasets in Table 1 which have been statistically analysed (no analysis possible or given for 6 additional datasets), 13 appear to be spatially pseudoreplicated, and 2 appear to be temporally pseudoreplicated. Analyses of pseudoreplicated data confuse differences between regions with differences between regional

species pools, as explained in detail below. This is a subtle distinction, which may explain why pseudo-replication is so widespread.

Spatial pseudoreplication occurs when the local richness of a region is not plotted as a single mean value, but rather as a cluster of values obtained from several localities in the same region (method 3), or several host populations of the same host species (method 2). A simple thought experiment will illustrate why this constitutes pseudoreplication (this is based on method 3 analysis, but applies as easily to method 2). Consider the case of 15 separate regions, each with an associated mean local richness. There should be a total of 14 degrees of freedom, some of which will subsequently be used by the regression analysis. Now consider the case of three separate regions, with five measurements of local richness for each region. There clearly cannot be a total of 14 degrees of freedom for the regression analysis, for this case is quite different from the above one. There are, in fact, only 2 degrees of freedom, for there are only 3 regions. The key assumption of regression analysis is that all the data points are independent, and the 3 regions with 5 locations scenario contravenes this assumption. This is, of course, equally true for any form of regression analysis, including ANOVA and its analogues. In fact, as ANOVA – by definition – requires multiple measurements per treatment level, any local–regional plots analysed using ANOVA-type tests *must* be pseudoreplicated. Measuring local richness repeatedly in the same region confounds site (region) and regional richness effects as surely as having only one block per treatment level (see also Westoby 1993). Essentially one is testing for an effect of region, not of regional richness, two subtly different questions.

Spatial pseudoreplication can be avoided using a single local richness value (such as the mean) and a single regional richness value for each region, which will preclude using ANOVA-type analyses. The most appropriate regression analyses compare the fit of linear and curvilinear models and assess the significance of either model, and the statistical technicalities of the procedures have been discussed earlier by Cresswell *et al.* (1995). It is difficult to generalize about the effects of correcting for pseudoreplication, as this depends on the exact distribution of the data. As a specific example, correcting for pseudoreplication in data on the parasites of lepomid sunfish (data from Aho & Bush 1993) does not change the conclusion of saturation but does decrease the significance of both linear and curvilinear regressions (Table 2; details of calculations in Appendix 1).

The above comments on spatial pseudoreplication apply to analyses where the regions are geographically distinct and separate (presently the most common situation). The waters become muddied when regional richness is estimated separately for each local site in what might otherwise be called a biogeographic region (for example by seed trays at each locality, or by using range maps to separately calculate the species pool within a fixed radius of each site). These may be sensible methods for estimating the species pool: as will be argued later, regional boundaries are often far from clear-cut, and species pool for different localities may grade into each other. The validity of such analyses will depend on minimising spatial autocorrelation between sites.

Temporal pseudoreplication is much rarer than spatial pseudoreplication (method 4 has rarely been used). Van Valkenburgh & Janis (1993) compared local and regional North American mammal diversity at various times over the last 44 million years but they explicitly recognized the non-independence of their data points are urged caution in the interpretation of the statistical results.

Bohnsack & Talbot (1980) introduced eight identical artificial reefs off the coast of both Australia and Florida, and compared fish species richness of the artificial reefs to the two regional pools. Their 'species/reef/month' variable seems to have elements of temporal pseudoreplication (multiple values, apparently one per month). The correct number of values for comparison is two, one for each regional richness, but of course two values cannot be compared statistically. Nor could the problem be solved by assuming that the regression line passes through the origin and so increasing the number of data points to three. Several authors have argued that, as regions with a species pool of zero must also have a local richness of zero, the regression line should be constrained to pass through the origin (Hugueny & Paugy 1995; Caley & Schluter 1997). Constraining the intercept is, however, always of dubious statistical merit, for it inflates the degrees of freedom and extrapolates beyond the range of the data (Wiens 1989; Hawkins & Compton 1992; M. Crawley, personal communication). If local richness in extremely depauperate regions is of interest, it seems it would be far better to actually measure it than imagine it.

In general, when the number of datapoints is restricted to avoid pseudoreplication, it may well emerge that there are insufficient degrees of freedom to decide if a community is saturated or unsaturated. This underlines the difficulty of obtaining local and regional richness estimates for many different regions, a major practical restriction to the general applicability of the method.

Not only is it important to ensure independence between local richness estimates but, ideally, also the estimates of local and regional richness measures for each region. Most commonly, local and regional estimates for a region are not independent: they are estimated from the same data set of site-specific species lists (regional richness is estimated as the cumulative number of species, local richness as the mean number of species). This could conceivably lead to spurious correlations (for example, if a species was missed from all the sites where it occurred, both local and regional

richness would be depressed; Stuart & Rex 1994; Cresswell *et al.* 1995). This problem could be resolved by calculating regional richness from distribution maps or regional floras and faunas, or by partitioning the data set into a subset for local richness estimates and a subset used for regional richness measures (e.g. Stuart & Rex, 1994; Cresswell *et al.* 1995).

There is a final subtle statistical question to consider, with regard to local–regional plots: the choice of null model. Ecologists who expect all communities to be saturated will be interested in determining if communities in the most depauperate regions are unsaturated. Ecologists who expect no communities to be saturated will be interested in determining if communities in the richest regions begin to approach saturation. In the first case, any correlation (linear or curvilinear) between local and regional richness is interpreted as evidence for lack of saturation, while in the second case the community is unsaturated only if a linear function fits the data better than a curvilinear function. This dichotomy in definitions is evident in the statistical tests summarized for each study in Table 1. It is possible that the same local–regional plot could be interpreted as evidence for saturation by one ecologist, and for lack of saturation by another. Obviously, the terms 'saturated' and 'unsaturated' need to be qualified, and the null hypothesis explicitly described, when reporting or citing analyses of local–regional plots.

Uncertainty in local richness

Local richness values will depend on sample sizes. Recently, Caley & Schluter (1997) investigated the effect of differing local sample size on local–regional plots, both theoretically and empirically. Small sample sizes tended to underestimate local richness proportionately more in the richest localities (given a log-normal distribution of species abundances, the rarest species is least abundant, and so most easily overlooked, when species richness is greatest). Griffiths (1997) suggests that a similar effect may occur when richer regions have a greater range of body sizes, as the smallest species are also most easily overlooked. This systematic bias could result in misleading curvilinear local–regional plots for unsaturated communities. Possible solutions to this problem are described in detail by Caley & Schulter (1997).

Uncertainty in regional richness

THE EFFECTS OF DIFFERENCES IN REGION SIZE

The analysis of local–regional plots requires data from several regions which differ substantially in regional richness. Ideally, these regions would be identical in every other way (same geographical area, disturbance frequency, etc.). Realistically, regions will differ in

these other variables, which may confound some of the results.

Consider the case of regions which differ in geographical area as well as regional richness. In an unsaturated community the ability of each species in the region to invade a given local community will be a decreasing function of the distance it must disperse. Species occurring on the edge of a large region are therefore much less likely to be able to disperse to a central locality than those on the edge of a small region. Such distant species in the large region are barely contributing to the effective species pool of that locality. However, all species are treated equally in local–regional plots, regardless of these distance effects, so the effective species pool of large regions may be overestimated. If such large regions are also the richer regions, this may result in an erroneous asymptotic curve in an unsaturated community (Fig. 5; see also Caley & Schluter 1997).

It is quite likely that larger regions are also richer. In the case of method 3, island biogeography theory predicts that larger regions will contain more species than smaller regions. In the case of method 2, host range has frequently been found to be a good predictor of its herbivore/parasite species pool (e.g. Strong, Lawton & Southwood 1984; Gregory 1990). Of the studies that use local–regional plots, summarized in Table 1, those looking at oak galls, wood-boring beetles, insects on Asteraceae, bracken herbivores, freshwater fish, and fish parasites all show an increase in regional diversity with host range or region size (Cornell 1985a; Stevens 1986; Lewinsohn 1991; Lawton *et al.*, 1993; Aho & Bush 1993; Kennedy & Guégan 1994; Hugueny & Paugy 1995; Belkessam, Oberdoff & Hugueny 1997). The effects of correcting for host range or region size in three of these studies are described later in this paper.

I have just summarized verbally how variation in region size could lead to an erroneous asymptotic curve for an unsaturated community. I will now present a more specific model of this phenomenon. This model will show that the local richness of unsaturated communities appears to be a power function of regional richness when regions differ in geographic size. The main assumption of the model is that the effective species pool of each region i ($S_{E,i}$) is best estimated by the number of species in the same standard area (A_E) in each region (for example, an area equivalent to the size of the smallest region). Any species occurring outside this area are assumed to be too far away to be potential contributors to local richness (this assumption is discussed further in the next section). This assumption underlies the local–regional plots of Westoby (1993) and Caley & Schluter (1997). The only other assumptions of the model are that standard (power function) species–area curves apply within regions, and a separate (and also a power function) species–area curve applies between regions. These seem to be reasonable assumptions, as power
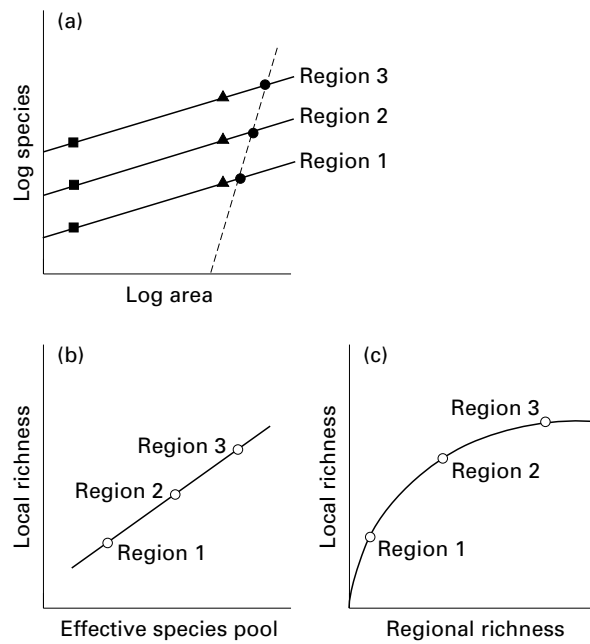
Fig. 5. (a) If three regions differ considerably in regional area (solid circles), the effective species pool (triangles) may be best represented by the number of species found in a constant area in each region. An unsaturated community is illustrated, with both within-region and between-region species–area curves plotted in log–log space. A plot of local richness (squares in (a)) against the effective species pool for the three regions is expected to be linear (b), but a misleading curvilinear relationship (c) between local and regional richness may occur for this unsaturated community; for details see text.

function species–area curves are well established, both within and between regions (for example, Rosenzweig 1995). A similar starting point was used by Holt (1993), albeit for a somewhat different question (see below).

The between-region relationship between the uncorrected regional richness (S_R) and the full region area (A_R) can be written as:

$$\log S_R = k(\log A_R) + b \qquad (6)$$

where both k and b are constants.

The within-region relationship between species (S) and area (A) for a given region i can be denoted by:

$$\log S = z_i(\log A) + c_i \qquad (7)$$

where area can vary from local area ($A_{L,i}$) to the full area of the region ($A_{R,i}$), corresponding to a range in species richness from the local ($S_{L,i}$) to regional ($S_{R,i}$). The effective species pool ($S_{E,i}$) and its area ($A_{E,i}$) will occur within this range. I now digress slightly to show that the slope, z_i, will be the same for all regions if the community is unsaturated. Recall that, if the community is unsaturated, the ratio of local richness ($S_{L,i}$) to the effective species pool ($S_{E,i}$) should, by definition, be the same for each region. Using the above within-region equation (eqn 7), the log of this ratio is equivalent to:

$$\log S_{L,i} - \log S_{E,i} = z_i(\log A_{L,i} - \log A_{E,i}) \qquad (8)$$

As local area ($A_{L,i}$) and effective region area ($A_{E,i}$) are also the same for each region, z_i must also be the same for each region, and can be simply denoted z (Fig. 5).

Returning to the within-region equation (eqn 7), the ratio of local richness ($S_{L,i}$) to the original uncorrected regional richness ($S_{R,i}$) will be related to the ratio of local area ($A_{L,i}$) to the full regional area ($A_{R,i}$) as follows:

$$\log S_{L,i} - \log S_{R,i} = z(\log A_{L,i} - \log A_{R,i}) \qquad (9)$$

or substituting the between-region equation (eqn 6):

$$\log S_{L,i} - \log S_{R,i} = zk^{-1}(k(\log A_{L,i}) - \log S_{R,i} + b) \qquad (10)$$

Since A_L is presumably the same for all regions, and k, z and b are constants as well, rearrangement gives:

$$\log S_{L,i} = (1 - zk^{-1})\log S_{R,i} + d \qquad (11)$$

where the constant $d = z(\log A_L) + bzk^{-1}$. Local species richness will appear to saturate ($1 > 1 - zk^{-1} > 0$) with regional richness when $z < k$, that is, when the within-region exponent is less than the between-region exponent. Species–area curves are normally steeper between regions than within regions (Rosenzweig 1995), so such 'pseudosaturation' would appear to be the most common situation.

There are a few other interesting scenarios. When regions are of the same area but differ in richness, k approaches infinity, and the familiar unsaturated linear relationship occurs ($1 - zk^{-1} = 1$). When within-region and between-region species–area curves are identical ($z = k$), local richness is the same between regions ($1 - zk^{-1} = 0$): this is the passive sampling effect described previously by Holt (1993).

In summary, the simple model described above indicates that unsaturated communities could appear to approach saturation when region size and richness are positively correlated. When the effective species pool is best estimated by a standard area in each region, a power function is anticipated between local and regional richness. Although I do not explicitly model the case for saturated communities, it is easy to see that, by similar logic, increases in regional richness with regional area will make curvilinear local regional plots look even less linear.

HOW SHOULD DIFFERENCES IN REGION SIZE BE DEALT WITH?

Given that correlations between region size and regional richness could lead to artefactual curvilinear local–regional relationships, it seems necessary to account for differences in region size. One possible approach is to restrict regional richness to that found in a standard area in each region (e.g. the area of the smallest region). This assumes that species which occur outside this standard area do not contribute at all to the effective species pool. Not correcting for regional area at all assumes that all species contribute equally to the effective species pool. Perhaps the best approach is to separately plot regional richness both corrected and uncorrected for area, with the rationale that the effective region size will be bracketed by these extremes, and conclude that the pattern is robust only when both analyses yield the same type of local–regional plot. This approach has been used by Kennedy & Guégan (1994) for method 2, and Gaston & Gauld (1993) for method 3. In both cases, correcting for host range and region area did not affect the basic shape of the local–regional plots. A practical example of correcting for host range, using data on the parasites of lepomid sunfish from Aho & Bush (1993) is given in Appendix 1, and summarized in Table 2. The parasite species pool of lepomid sunfish species is positively correlated with the size of the host range.

I use the expected power function between species richness and area to estimate the species pool if all hosts had the same range size. In this dataset, the best-fit regression between local and regional richness is a power function, using either full or equal-area host ranges suggesting that the conclusion of saturation is robust. Note that the difference between linear and curvilinear regressions decreases when regional richness values for equal areas are used (Table 2) as predicted in the previous section.

The above discussion suggests that the effective species pool should depend not just on geographical/host range limits, but also on dispersal distances. This conclusion is illustrated by a study of prosobranch snails (Stuart & Rex 1994), in which local richness is linearly related to regional richness. The residuals of local richness from this linear relationship correlate positively with the dispersal ability (proportion of plankotrophic species) of different species pools. One interpretation of this result is that the greater dispersal ability of these regions increases the effective species pool beyond that predicted by the standardized (equal sample size) species pool.

EFFECTS OF OVERESTIMATING AND UNDERESTIMATING REGIONAL RICHNESS

Determining the regional species pool is obviously not always a simple matter. Cornell (1985b) describes it as 'a vexing problem with no simple solution beyond experimental defaunation and introduction'. To take a concrete example, consider the bracken herbivores of South Africa. If South Africa as a whole is defined as a single region, the regional species pool would be 13 species (Compton, Lawton & Rashbrook 1989). Some bracken herbivores, however, are restricted to either northern or southern latitudes within South Africa, with the result that only 7 of these 13 span the whole of South Africa (Lawton et al. 1993).

Regional 'boundaries' could conceivably be identified by abrupt changes in beta diversity (for example,

Table 2. Two corrections advocated in the text are illustrated with the sunfish parasite data of Aho & Bush (1993). Local richness (L) data is corrected for pseudoreplication (averaged over sites) unless indicated ($n = 64$ before correction, $n = 9$ after). Regional richness (R) is based on range sizes of different areas except for the equal-area analysis (in this analysis the richness of a standard area is used, equal to the size of the smallest region). Note that neither correction results in the linear model (linear L–linear R) fitting the data better than curvilinear models (all others)

Regression		Pseudoreplicated L	Different-area R	Equal-area R
Linear L–linear R	r^2	0·077	0·344	0·429
	$F(P)$	5·13 ($P < 0.05$)	3·67 (NS)*	5·25 (NS)*
Linear L–log R	r^2	0·105	0·455	0·516
	$F(P)$	7·30 ($P < 0.01$)	5·85 ($P < 0.05$)	7·46 ($P < 0.05$)
Log L–linear R	r^2	0·092	0·425	0·508
	$F(P)$	6·31 ($P < 0.05$)	5·17 (NS)*	7·23 ($P < 0.05$)
Log L–log R	r^2	0·137	0·572	0·632
	$F(P)$	9·86 ($P < 0.01$)	9·37 ($P < 0.05$)	11·99 ($P < 0.05$)

*NS = not significant at the 5% level.

in plots of beta diversity against increasing size of the putative region). In the case of South African bracken herbivores, however, species accumulate evenly as latitudinal range is increased (either north to south or south to north; Lawton *et al*. 1993). It is not clear how dissimilar the species composition of the northern and the southern areas of South Africa would have to be before they could be considered separate regions. Nor is it clear how similar more northerly faunas would have to be in order to be considered part of the South African region. These are complex questions with no easy answers, and perhaps best left to those with expertise in each particular system. Rather than outlining 'rules' for demarcating regions, therefore, I have chosen instead to examine the consequences of overestimating and underestimating regional richness.

Consider first overestimation. Only a subset of the perceived species pool may actually be able to colonize a host in the short term, a phenomena called 'pool exhaustion' by Lawton & Strong (1981). Cornell (1985b) first suggested that pool exhaustion could lead to an unsaturated community appearing saturated, an explanation later used by Aho (1990) to explain an asymptotic relationship in amphibian parasites. Method 1 similarly overestimates regional richness by including species in different habitats other than the one of interest. Whatever the cause, though, overestimation can only lead to an asymptotic relationship in an unsaturated community when the species pool of truly richer regions are overestimated proportionately more than impoverished regions (see also Cornell & Lawton 1992; Griffiths 1997; Caley & Schluter 1997). This is exemplified by the case of a power function between regional size and regional richness discussed above. If regional richness is overestimated by a constant or proportional amount, an unsaturated community will still have a linear local–regional plot (the intercept will be lower for a constant amount, the slope will be lower for a proportional amount).

The effects of overestimation of regional richness on a saturated community do not appear to have been discussed before. Overestimation of regional richness by a greater than proportional amount will increase the curvature of the saturation curve, and so increase the gap between the fit of linear and curvilinear (polynomial or power function) regression lines (and so reduce the probability of mistaking a saturated community for an unsaturated one). Overestimation by a proportional or constant amount will have little effect on the results, unless a power function is used to assess a curve that has been horizontally translated to the right of the origin (power functions are constrained to pass through the origin, so the fit will be poor).

A likely scenario for the underestimation of regional richness might be the use of too small an area for regions. Regardless of whether communities are saturated or unsaturated, as long as their species–area curves conform to Fig. 1c and Fig. 1e respectively, reduction in region area by a constant amount will

have the same proportional effect on local to regional ratios. That is, a linear relationship between local and regional richness will remain linear (just steeper), and a curvilinear relationship will remain curvilinear (just horizontally compressed).

Synthesis

It is evidently easy to reach the wrong conclusion about species saturation by analysing local–regional richness plots. The data could be at the between-habitat scale, not the within-habitat scale suitable for this analysis. The host species used could have very different saturation curves, or local and regional richness could covary with some unknown environmental variable. The local richness values could be pseudoreplicated. The regional pool could be overestimated by summing across habitats, or need to be corrected for area and dispersal effects. There may not be enough data to really decide between models. Some of these pitfalls can be avoided by careful construction of the data set and tight statistics, but other potential problems are much more difficult to deal with.

On the other hand, local–regional richness plots are one of the most direct ways of separating the processes that influence local diversity into ecological and evolutionary components. This technique has a valuable role to play alongside the more established methods in this field such as analyses of species introductions or biotic interchange, and experimental analyses of competitive processes. It is also a very versatile technique, applicable to a wide variety of communities as the literature review in Table 1 shows. Furthermore, the relationship between beta diversity, species–area curves and local–regional plots outlined in this review suggest new ways of extending the method: for example, by comparing the spatial turnover of species between regions.

A pluralistic approach is always better than relying on a single technique, and the potential problems outlined in this review emphasize the importance of substantiating the results of local–regional plots with other evidence. I conclude with a summary of useful supportive evidence for both unsaturated and saturated local–regional plots, giving examples from the literature.

Unsaturated local–regional plots will be supported by evidence of the non-interactive nature of the community. According to Cornell & Lawton (1992), all non-interactive community models yield unsaturated communities; but the reverse is not true because apparently unsaturated communities (linear local–regional plots) can also be predicted by interactive models (see also Caswell & Cohen 1993). Non-interaction is therefore an asymmetric test for lack of saturation. Hawkins & Compton (1992) support their unsaturated local–regional plot of fig wasps by citing evidence of non-interaction, including lack of resource limitation (abundant uncolonized figs) and stochastic

colonization. Similarly, Lawton *et al.* (1993) support their unsaturated local–regional plot of bracken herbivores with evidence of weak interspecific interactions and obvious vacant niches. Huegeny & Paugy (1995) cite, as an explanation for an unsaturated local–regional plot of African fish, the frequency of local disturbance and the importance of colonization from the surrounding region (a river). Lastly, Oberdorff *et al.* (1998) support an unsaturated local–regional plot for riverine fish not only with evidence of negligible density compensation (i.e. population size is independent of local richness), but also show that all calculable competition coefficients are near zero.

Saturated local–regional plots are not predicted by any of the non-interactive community models examined by Cornell & Lawton (1992) or the non-interactive model of Caswell & Cohen (1993). Strong interspecific interactions are therefore necessary but not sufficient supporting evidence for saturation. Kennedy & Guégan (1994) cite evidence of competitive exclusion to support an explanation based on community saturation for their curvilinear local–regional plot of fish helminth communities. Further evidence comes from their analyses of subsets of the dataset; only the subset containing recently introduced fish species has a linear local–regional plot, as would be expected if such new fish species are in the process of acquiring parasites and so competitive exclusion is still rare. Other evidence that could be cited in support of a saturated community includes resistance to invasion, resource-limitation, niche shifting, density compensation and convergence in guild structure.

It is important for ecologists to understand the evolutionary processes underlying community structure and diversity. If communities are saturated with species in some regions, but not others, it may not be possible to extrapolate the results of ecological experiments across regions. Ecological theory has largely assumed that communities are both strongly interactive and saturated with species. Such theory (e.g. many explanations for maximum diversity at intermediate levels of productivity and disturbance) may need to be reinterpreted for non-interactive and unsaturated communities. It is crucial, therefore, for ecologists to have robust tests for species saturation. Only continued discussion of the methodology and theory of local–regional plots will ensure that this technique achieves its potential to be one of these robust tests.

Acknowledgements

References

Aho, J.M. (1990) Helmith communities of amphibians and reptiles: comparative approaches to understanding patterns and processes. *Parasite Communities: Patterns and Processes* (eds G.W. Esch, A.O. Bush & J.M. Aho), pp. 157–195. Chapman and Hall, London.

Aho, J.M. & Bush, A.O. (1993) Community richness in parasites of some freshwater fishes from North America. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 185–193. University of Chigaco Press, Chicago.

Belkessam, D., Oberdorff, T. & Hugueny, B. (1997) Unsaturated fish assemblages in rivers of North-Western France: potential consequences for species introductions. *Bulletin Français de Pêche et Pisciculture*, **344/345**, 193–204.

Bohnsack, J.A. & Talbot, F.H. (1980) Species-packing by reef fishes on Australian and Caribbean reefs: an experimental approach. *Bulletin of Marine Science*, **30**, 710–7213.

Caley, M.J. & Schluter, D. (1997) The relationship between local and regional diversity. *Ecology*, **78**, 70–80.

Caswell, H. & Cohen, J.E. (1993) Local and regional regulation of species–area relations: a patch occupancy model. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 99–107. University of Chicago Press, Chicago.

Compton, S.G., Lawton, J.H. & Rashbrook, V.K. (1989) Regional diversity, local community structure, and vacant niches: the herbivorous insects of bracken in South Africa. *Ecological Entomology*, **14**, 365–373.

Cornell, H.V. (1985a) Local and regional richness of cynipine gall wasps on Californian oaks. *Ecology*, **66**, 1247–1260.

Cornell, H.V. (1985b) Species assemblages of cynipid gall wasps are not saturated. *American Naturalist*, **126**, 565–569.

Cornell, H.V. (1993) Unsaturated patterns in species assemblages: the role of regional processes in setting local species richness. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 243–252. University of Chicago Press, Chicago.

Cornell, H.V. & Karlson, R.H. (1996) Species richness of reef-building corals determined by local and regional processes. *Journal of Animal Ecology*, **65**, 233–241.

Cornell, H.V. & Lawton, J.H. (1992) Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. *Journal of Animal Ecology*, **61**, 1–12.

Cox, G.W. & Ricklefs, R.E. (1977) Species diversity and ecological release in Caribbean land bird faunas. *Oikos*, **28**, 113–122.

Cresswell, J.E., Vidal-Martinez, V.M. & Crichton, N.J. (1995) The investigation of saturation in the species richness of communities: some comments on methodology. *Oikos*, **72**, 301–304.

Dawah, H.A., Hawkins, B.A. & Claridge, M.F. (1995) Structure of the parasitoid communities of grass-feeding chalcid wasps. *Journal of Animal Ecology*, **64**, 708–720.

Eriksson, O. (1993) The species-pool hypothesis and plant community diversity. *Oikos*, **68**, 371–374.

Gaston, K.J. & Gauld, I.D. (1993) How many species of pimplines (Hymenoptera: Ichneumonidae) are there in Costa Rica? *Journal of Tropical Ecology*, **9**, 491–499.

Gregory, R.D. (1990) Parasites and host geographic range as illustrated by wildfowl. *Functional Ecology*, **4**, 645–654.

Griffiths, D. (1997) Local and regional species richness in North American lacustrine fish. *Journal of Animal Ecology*, **66**, 49–56.

Hawkins, B.A. & Compton, S.G. (1992) African fig wasp communities: undersaturation and latitudinal gradients in species richness. *Journal of Animal Ecology*, **61**, 361–372.

Holt, R.D. (1993) Ecology at the mesoscale: the influence of regional processes on local communities. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 77–88. University of Chicago Press, Chicago.

Hugueny, B. & Paugy, D. (1995) Unsaturated fish communities in African rivers. *American Naturalist*, **146**, 162–169.

Kennedy, C.R. & Guégan, J.-F. (1994) Regional versus local helminth parasite richness in British freshwater fish: saturated or unsaturated parasite communities? *Parasitology*, **109**, 175–185.

Kingsland, S.E. (1985) *Modelling Nature*: *Episodes in the History of Population Ecology*. University of Chicago Press, Chicago.

Lawton, J.H. (1982) Vacant niches and unsaturated communities: a comparison of bracken herbivores at sites on two continents. *Journal of Animal Ecology*, **51**, 573–595.

Lawton, J.H. (1990) Local and regional species–richness of bracken–feeding insects. *Bracken Biology and Management* (eds J.A. Thomson & R.T. Smith). Australian Institute of Agricultural Science, Sydney.

Lawton, J.H., Lewinsohn, T.M. & Compton, S.G. (1993) Patterns of diversity for the insect herbivores on bracken. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.F. Ricklefs & D. Schluter), pp. 178–184. University of Chicago Press, Chicago.

Lawton, J.H. & Strong, D.R. (1981) Community patterns and competition in folivorous insects. *American Naturalist*, **118**, 317–338.

Lewinsohn, T.M. (1991) Insects in flower heads of Asteraceae in southeast Brazil: a case study in tropical species richness. *Plant–Animal Interactions*: *Evolutionary Ecology in Tropical and Temperate Regions* (eds P.W. Price, T.M. Lewinsohn, G.W. Fernandes & W.W. Benson), pp. 525–559. John Wiley & Sons, Inc., Chichester.

Oberdorff, T., Hugueny, B., Compin, A. & Belkessam, D. (1998) Non-interactive fish communities in the coastal streams of North-Western France. *Journal of Animal Ecology*, **67**, 472–484.

Pärtel, M., Zobel, M., Zobel, K. & van der Maarel, E. (1996) The species pool and its relation to species richness: evidence from Estonian plant communities. *Oikos*, **75**, 111–117.

Pearson, D.L. (1977) A pantropical comparison of bird community structure on six lowland forest sites. *Condor*, **79**, 232–244.

Pearson, D.L. & Juliano, S.A. (1993) Evidence for the influence of historical processes in co-occurrence and diversity of tiger beetle species. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 194–202. University of Chicago Press, Chicago.

Richardson, D.M., Cowling, R.M., Lamont, B.B. & van Hensbergen, H.J. (1995) Coexistence of *Banksia* species in southwestern Australia: the role of regional and local processes. *Journal of Vegetation Science*, **6**, 329–342.

Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Ricklefs, R.E. & Schluter, D. (1993) *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives*. University of Chicago Press, Chicago.

Rørslett, B. (1991) Principal determinants of aquatic microphyte richness in northern European lakes. *Aquatic Botany*, **39**, 173–193.

Rosenzweig, M.L. (1995) *Species Diversity in Time and Space*. Cambridge University Press, Cambridge.

Schluter, D. & Ricklefs, R.E. (1993) Species diversity: an introduction to the problem. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 1–10. University of Chicago Press, Chicago.

Stevens, G.C. (1986) Dissection of the species–area relationship among wood-boring insects and their host plants. *American Naturalist*, **128**, 35–46.

Strong, D.R., Lawton, J.H. & Southwood, T.R.E. (1984) *Insects on Plants*: *Community Patterns and Mechanisms*. Blackwell Science, Oxford.

Stuart, C.T. & Rex, M.A. (1994) The relationship between development pattern and species diversity in deep-sea prosobranch snails. *Reproduction, Larval Biology and Recruitment in the Deep-Sea Benthos* (eds C.M. Young & K.J. Eckelbarger), pp. 118–136. Columbia University Press, New York.

Taylor, D.R., Aarssen, L.W. & Loehle, C. (1990) On the relationship between r/K selection and environmental carrying capacity: a new habitat templet for plant life history strategies. *Oikos*, **58**, 239–250.

Terborgh, J., Faaborg, J. & Brockmann, H.J. (1978) Island colonization by lesser antillean birds. *Auk*, **95**, 59–72.

Terborgh, J. & Faaborg, J. (1973) Turnover and ecological release in the avifauna of Mona Island, Puerto Rico. *Auk*, **90**, 759–779.

Terborgh, J.W. & Faaborg, J. (1980) Saturation of bird communities in the West Indies. *American Naturalist*, **116**, 178–195.

Tonn, W.M., Magnuson, J.J., Rask, M. & Toivonen, J. (1990) Intercontinental comparison of small-lake fish assemblages: the balance between local and regional processes. *American Naturalist*, **136**, 345–375.

Van Valkenburgh, B. & Janis, C.M. (1993) Historical diversity patterns in North American large herbivores and carnivores. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 330–340. University of Chicago Press, Chicago.

Vaughn, C.C. (1997) Regional patterns of mussel species distributions in North American rivers. *Ecography*, **20**, 107–115.

Westoby, M. (1993) Biodiversity in Australia compared with other continents. *Species Diversity in Ecological Communities*: *Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 170–177. University of Chicago Press, Chicago.

Whittaker, R.H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, **30**, 279–338.

Wiens, J.A. (1989) *The Ecology of Bird Communities*, 1. Cambridge University Press, Cambridge.

Zobel, M. (1997) The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? *Trends in Ecology and Evolution*, **12**, 266–269.

Zwölfer, H. (1987) Species richness, species packing, and evolution in insect-plant systems. *Ecological Studies*, **61**, 301–319.

Appendix 1

REANALYSIS OF SUNFISH LOCAL–REGIONAL RICHNESS PLOTS

Aho & Bush (1993) present data on the parasites of lepomid sunfish. Several of the corrections advocated

in the text will be illustrated with this data set. Any number of studies could have been reanalysed in this way, and this data set was simply chosen because it was convenient for the analysis.

The parasite richness of nine sunfish species are presented by Aho & Bush (1993; data read from Fig. 17.6 for total parasites). Data are given for numerous (3 to 22) localities for each species, for a total of 64 data points in the original analysis. I argue in the text that this constitutes spatial pseudoreplication. The local–regional plot is analysed using mean values of local richness (Table 2). Although curvilinear regressions fit the data better than a linear regression both before and after correction for pseudo-replication, the significance of all regressions is reduced when mean values are used.

The lepomid sunfish have different range sizes. The range of each species was estimated using regional richness–host range relationships given by Aho & Bush (Table 17.3, 1993). Each local richness value is from a 'discrete body of water (or from a specific collection site of a large body of water)', which is assumed to be approximately 5 km². Species richness can be expressed as a function of area for each region given that (a) the relationship is expected to be linear in log–log space, and (b) species richness and area is known for two points on this line, namely at the local and regional scale. Using these equations, the parasite richness of each species was estimated for an area of 58,500 km², the smallest range size (the results do not depend on exactly which standard area is chosen). This procedure is illustrated in Fig. 6. When this equal-area richness is used in place of the original regional richness values, the fit of the linear regression increases more than that of the power function regression (although the latter remains the best fit; Table 2).
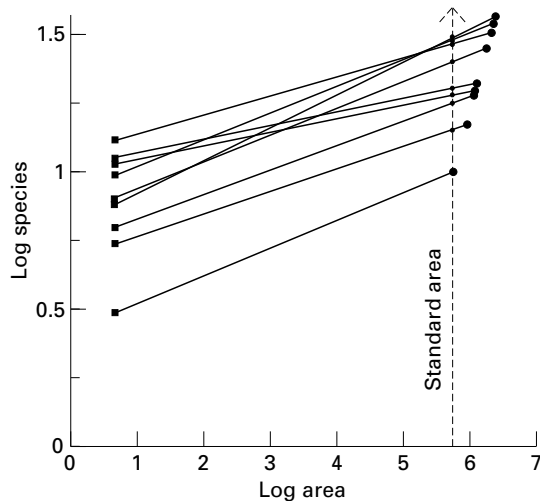
Fig. 6. Calculation of species richness for an area of equal size (dashed line) in several regions, using data on sunfish parasites (Aho & Bush 1993). The relationship between species richness and area (solid line) is assumed to be linear in log–log space from the local (squares) to regional (large circles) scale. The richness of a standard area (small circles) can then be estimated from these linear equations.

Appendix 2

STATISTICAL ANALYSIS OF THE FOREST BIRD DATA OF PEARSON (1977)

Pearson (1977) estimated local and regional richness of birds in the tropical forests of six regions. He suggested that the (eyeballed) similarity of local to regional ratios between the regions made 'historical factors a vital part of understanding these communities'. A formal statistical analysis of his data is presented here, using the methodology of local–regional plots which has been developed since his study. This analysis also used local richness values determined for a constant number of hours of observation, unlike Pearson's original comparison (variable sampling effort is problematic when some local richness estimates appear to still be increasing with sampling effort at the final sample). Local richness was estimated for 280 hours of observation (Fig. 1 in Pearson 1977). Regressions involving all combinations of both linear and logarithmic forms of local and regional richness were calculated (Table 3). Although some of the curvilinear regressions technically have a better fit than the linear regression, the difference is small and the curvature slight over the range of the data (Fig. 7). It is concluded, therefore, that the relationship is essentially unsaturated.

Table 3. Linear and curvilinear regressions describe equally well the relationship between the local richness (L) and regional richness (R), in tropical forest birds of six regions. Data are from Pearson (1977), with corrections to local richness as described in text

Regression	r^2	Slope	Intercept
Linear L–linear R	0·935	0·448	36·3
Linear L–log R	0·902	75·3	−271
Log L–linear R	0·958	$3·84 \times 10^{-3}$	57·7
Log L–log R	0·944	0·652	3·98
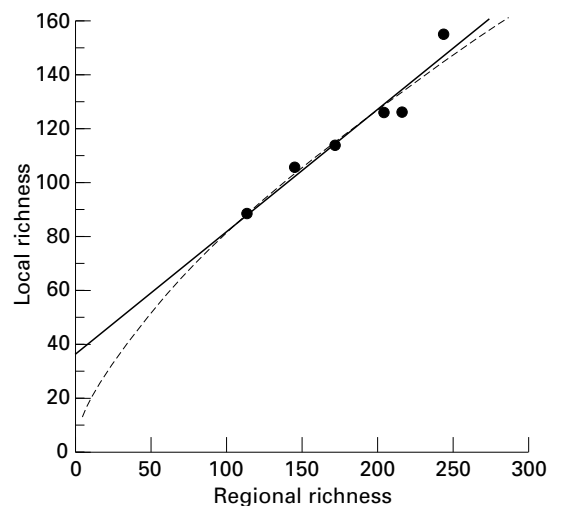
Fig. 7. Local–regional richness plots for tropical forest birds, using data from Pearson (1977). Both linear (solid line) and exponential (dashed line) regression lines are shown (equations in Table 3).