ECOGRAPHY 27: 532–544, 2004

Rarefaction method for assessing plant species diversity on a regional scale

Thomas Koellner, Anna M. Hersperger and Thomas Wohlgemuth

Koellner, T., Hersperger, A. M. and Wohlgemuth, T. 2004. Rarefaction method for assessing plant species diversity on a regional scale. – Ecography 27: 532–544.

In national conservation plans, it is necessary to comparatively assess species pools of different regions and monitor their changes over time. Two specific problems arise: i) species diversity must be standardized per area, because regions differ in size, and ii) the diversity measure should take into account how common or rare the species are on the regional scale. We used the rarefaction method combined with a fitting procedure to calculate the expected number of species E(S). The method takes into account the nonlinearity of species and area, as well as how common or rare each species is and allows analysis of species groups' contribution to total species diversity. The slope parameter of the fitted power function is used as an indicator of species turnover, and thus, of β-diversity. For the analysis, Switzerland was divided into seven biogeographic regions (256–10 642 km^2). The diversity of the total species pool and of six ecological species groups was investigated for each region. In every biogeographic region, we find the lowest species turnover in the fertilized meadow group, and the highest species turnover in the pioneer/weedy species and the mountain species groups. The results show that among Swiss regions, differences in E(S) are mainly due to the presence or absence of mountain species. Other species groups show a rather constant contribution to the regional species pools. We found the rarefaction method to be a very useful tool for assessing Swiss plant species diversity on a regional scale.

T. Koellner (thomas.koellner@env.ethz.ch), Dept of Environmental Sciences, HES Inst. for Human-Environmental Systems, ETH-Zentrum HAD, CH-8092 Zurich, Switzerland. – A. M. Hersperger and T. Wohlgemuth, WSL Swizz Federal Inst. for Forest, Snow and Landscape Research, Zürucherstrasse 111, CH-8903 Birmendorf, Switzerland.

Biodiversity policies and conservation efforts are increasingly focusing on the landscape level rather than the dominantly local focus of the past (Heywood 1995). This is partly due to article seven of the Convention on Biological Diversity (Anon. 1992), which requires participating countries to identify and monitor biodiversity nationwide. In this context, there needs to be a way to monitor the development of the regional species diversity over time and to compare species diversity among regions or countries (Anon. 2001, Annex 1). Such comparison will facilitate the assessment of anthropogenic impacts on biodiversity as well as the control of strategies and action plans.

To investigate the diversity of a national species pool (in sensu Pärtel et al. 1996, Zobel 1997), it is necessary to divide the area of a country into subregions. The species diversity in these subregions can be referred to as α-diversity and the diversity of the entire country as γ-diversity (Whittaker 1972, Balvanera et al. 2002). In this paper, we use the term β-diversity for the species turnover between subregions (sensu MacArthur 1965, Whittaker 1972, Whittaker et al. 2001). The β-diversity of a national species pool generally increases as the diversity of habitats and, hence, the environmental heterogeneity increase (Whittaker 1972, Alard and Podevigne 2000, Balvanera et al. 2002). Scale effects

Accepted 20 February 2004

ISSN 0906-7590

(Arita and Rodríguez 2002), isolation, and distance effects (Nekola and White 1999, Balvanera et al. 2002) affect β-diversity as well. If national grids are used to determine how common or rare species are, then the influence of grid characteristics (e.g. grain, extent and number of samples) on diversity measures and species area curves have to be taken into account (Palmer and White 1994, Witte and Torfs 2003).

Quantifying species diversity on a regional scale is quite challenging because of difficulties in measuring species abundance and distribution. Species richness (i.e. the number of species per sample plot) was used as a proxy for diversity (Magurran 1996) in experimental settings (e.g. Hector et al. 1999) and landscape ecology (e.g. Wohlgemuth 1998). Several authors (Hurlbert 1971, Simberloff 1978, Palmer 1990) proposed and discussed probabilistic methods for estimating species diversity based on the presence or absence of data on species in sample plots. They calculated the expected number of species in a sub-set of samples using a rarefaction function. Although data requirements for the rarefaction function are less demanding than for indices like the Shannon-Wiener Index (Shannon 1948) and the Simpson Index (Simpson 1949) (both of which require data on abundance), the rarefaction function integrates data on the each species' commonness or rarity in a given geographic region. Based on these features of the rarefaction function, Ricotta et al. (2002) proposed using the slope of the discontinuous function as an indicator for the β-diversity of plant communities. So far, the rarefaction method has been mainly used for investigations of animal species groups (e.g. Abele and Walters 1979, Achtziger et al. 1992, Douglas and Lake 1994, Boucher and Lambshead 1995, Caley and Schluter 1997, Gjerde and Sætersdal 1997).

The main goal of this paper is to present a method for quantifying plant species diversity on a regional level with a simple but adequate index. This method is based on the rarefaction function and allows: 1) comparison of the seven biogeographic regions of Switzerland in terms of species diversity and 2) comparison of six ecological plant species groups in terms of their contribution to regional species diversity.

Data and methods

Regions, data sets and species groups

Switzerland is located in the middle of Europe and has a total size of 41 244 km^2. The country includes a wide range of different climates, topography, geology, and types of land use types. Five large biogeographic regions can be distinguished (Gutersohn 1973). These regions are the Jura, the Swiss Plateau, and the Northern, Central and Southern Alps (Fig. 1). Areas above timberline in the Alps and the Jura were analyzed separately, because considerable parts of this zone are covered with vegetation consisting of highly adapted mountain species.

The regions' boundaries are drawn according to Wohlgemuth's analysis of the distribution of vascular plants (1996). A more recent biogeographical breakdown (not used in this research) takes political borders and the distribution of molluscs and insects into account as well

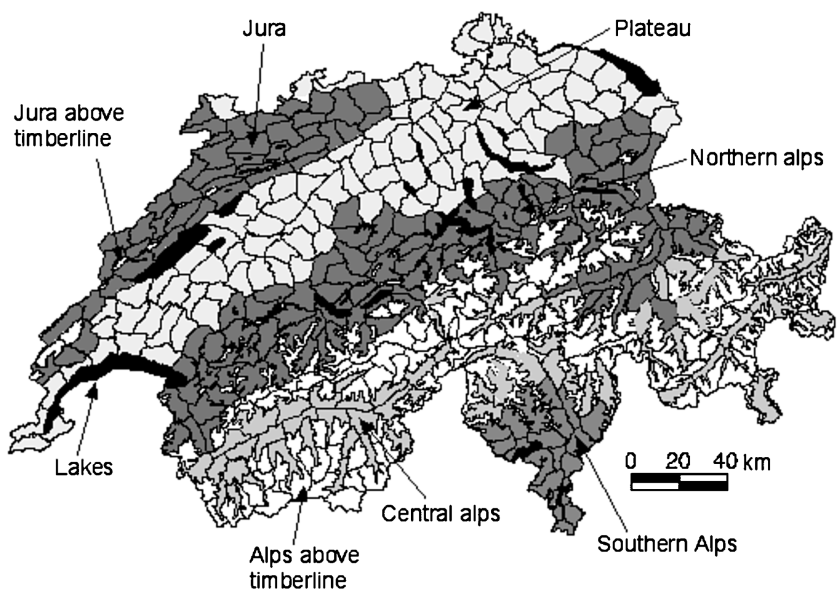

Fig. 1. Switzerland and its biogeographic regions: Jura, Plateau, Northern Alps, Central Alps, and Southern Alps (adapted from Wohlgemuth 1998).

(Gonseth et al. 2001). The intensity of land use varies among the biographical regions according to their characteristic climate and topography. The Swiss Plateau is lowland in character and used intensively for agriculture, forestry, and settlements. The Alps and the Jura Mountains, in contrast, are used less intensively and have more non-productive areas.

The information on vascular plant species richness was taken from the distribution atlas Pteridophytes and Phanerogams of Switzerland (Welten and Sutter 1982, Wohlgemuth 1993). The atlas includes data from an extensive and systematic ground survey of Switzerland. Between 1967 and 1979 about 170 botanists inventoried the vascular plant species in 593 polygons of, on average, 70 km^2 (Table 1). The shape of the polygons is consistent with topographical characteristics: therefore, the polygons varied in size. The mean polygon size, standard deviation, and minimum and maximum polygon size for each region are shown in Table 1. For each polygon, a list of vascular plant species was compiled. The species pool of Switzerland, in terms of total number of residential plant species, has 2586 species. This number refers to the nationwide survey and excludes species known only from literature or herbaria. In this list, many sub-species have been merged into species aggregates. The new Red List of Switzerland (not used in this research) lists 3144 species (Moser et al. 2002).

For further analysis, we used Landolt's species classification (1991). He classified vascular plant species into eight ecological species groups based on expert opinion. Because of the coarse similarity of their ecological requirements, we joined pioneer and weed species, and marsh and water species, respectively. We distinguished the following six species groups in our analysis (percentage of vascular plants in Switzerland according to Landolt 1991 is given in brackets): 1) forest species (17.1%): all vascular plant species associated with forests, forest edges, and shrubs; 2) fertilized meadow species (2.7%): species found in regularly mowed and fertilized meadows; 3) unfertilized meadow species (13.1%): species found in unfertilized and dry meadows; 4) mountain species (23.2%): species with their main occurrence above timberline and in sub-alpine forests (above 1200–1500 m); 5) marsh species (12.3%)/water species (4.4%): species found in inland wetlands (marshes, peat bogs) and bodies of water (brooks, rivers, ponds, lakes) 6) pioneer species (5.7%)/weed species (20.8%): species found in industrial fallows; gravel pits; deposits of sand, gravel and rubble; stone walls; floodplains; arable land; agricultural fallow; embankments.

Total and partial rarefaction functions

To compare the species pools of the Swiss regions, total and partial rarefaction functions were computed. Frequently, species area relationships are calculated for plots (e.g. true islands or habitat islands), that vary in size (Connor and McCoy 1979, Wisheu and Keddy 1996). Species richness S of n plots can then be determined and species area relationships can be directly fitted because of the variation in plot sizes A (cf. eq. 2, where species number S is a function of the plot size A). The direct fit procedure, however, is not possible if there is no variation in plot sizes, because species numbers are measured for a standardized plot size. In such a case, the rarefaction method (Hurlbert 1971, Heck et al. 1975) can be used instead. This method explicitly recognizes the non-linearity of the relationship between area and species number.

Rarefaction is based on a statistical procedure and is used to calculate the number of species expected in a sub-sample of individuals selected at random from a larger sample. In this paper, however, we estimate the number of species expected in a sub-sample of the total number of sample plots. Rarefaction technique can be used to standardize samples, that differ in terms of individual size or plot size. The method is derived from hypergeometric distribution and results in a hyperbolic curve showing the expected number of species for a given sample size.

The expected numbers of species $E(S_n)$ is calculated by randomly choosing a sub-sample n from all N plots in the sample:

Table 1. Area A of biogeographic region j of Switzerland and its polygons i (see Fig. 1) in km^2 (N: number of plots, Std Dev: standard deviation, min: minimum, max: maximum).

Region	A_j^{total}	A_i^{mean}	$A_i^{Std\ Dev}$	A_i^{min}	A_i^{max}	$N_{j,i}$
Jura	3757	85	27	36	172	44
Jura above timberline	256	11	16	0	62	23
Plateau	10642	102	23	30	161	104
Northern Alps	9073	82	25	23	150	110
Central Alps	5222	74	26	28	152	71
Southern Alps	1449	69	37	11	153	21
Alps above timberline	9597	50	31	0	134	192
Lakes	1248	45	48	5	229	28
Total Switzerland	41244	70	37	0	229	593

$$E(S_n) = S - \frac{\sum_{i=1}^{S} \binom{N - N_i}{n}}{\binom{N}{n}} \quad (1)$$

where N is the total number of plots in the sample; N_i, the number of plots where species i is found; n, the number of randomly chosen plots; and S, the total number of species on all the plots. Given a fixed number of species S and a fixed number of plots n, the expected number of species is dependent on the species' abundance (the more abundant a species i is, the more plots N_i are inhabited by it). If, for example, rare species become more abundant (i.e. N_i gets larger) while the total number of species S remains constant, the expected number of species $E(S_n)$ for n plots will increase. This is because it is more likely that a common species is found than a rare species. Thus, this method explicitly embraces the commonness or the rarity of species in the investigated area. The calculated $E(S_n)$ for randomly choosing one plot ($n = 1$) is equal to the mean species number and can indicate α-diversity in sensu MacArthur or within habitat diversity in sensu Whittaker (Ricotta et al. 2002). γ-diversity refers to the diversity of a total region or country, which is indicated by the number of species expected when all plots are sampled.

The rarefaction method was applied to the sum of all of the species in the pool (total rarefaction function) as well as partially to the six ecological species groups described above (partial rarefaction function). The partial rarefaction functions allow assessment of the contribution of different plant species groups to the total species pool. The calculation of partial rarefaction curves is possible, because the total expected number of species is additively composed of the single expected values for each species (cf. Appendix and Ricotta et al. (in press).

In general, rarefaction works with any kind of data set, but the most accurate results are obtained for data without spatial autocorrelation. Because rarefaction implicitly takes spatial autocorrelation into account, we also computed the spatial autocorrelation; i.e. any nonrandom spatial pattern of the individuals of one or more species. This allows the dependence on distance of the polygons' floristic composition in the seven regions to be checked. The basic assumption underlying rarefaction is that individuals of one species are randomly dispersed, while individuals of different species are distributed independently. In general, rarefaction works with any kind of dataset, but the most accurate results are obtained for data without spatial autocorrelation. In field data, however, this requirement can hardly be fulfilled.

Spatial autocorrelation was computed, as in Sokal (1986) and Legendre and Fortin (1989). This means species lists of the polygons were compared to one another using Euclidian distances. The resulting nXn-resemblance matrix S was then compared to a nXn-distance matrix D, involving the coordinates x and y at the centres of the polygons. We performed directed autocorrelations using MULVA-5 (Wildi and Orlóci (1996) in the coarse direction of the topographic/geological texture from the northwest to the southeast. Floristic changes along these directions are the smallest.

Fitting discontinuous rarefaction functions

Continuous species area relationships were fitted to the discontinuous rarefaction functions and the resulting parameters were applied for comparative assessment of regions and species groups. Models used for fitting species area samples depict a monotonically increasing curve, which is steep at the beginning and gradually becomes flat (He and Legendre 1996). That is, the first deviation of the functions is decreasing. Three models are commonly used for such fittings: the power model (Arrhenius 1921), the exponential model (Gleason 1922, 1925), and the logistic model (Archibald 1949). The species rarefaction function (eq. 1) is a monotonically increasing function, and it is reasonable to expect that such a function is rather straight in a semilog space or a log-log space with high R squared values.

We used the power (log-log) model

$$\ln S = \ln c + z \ln A \quad (2)$$

(where S: species number, A: area of the plot, c: measure for species richness, and z: measure for species accumulation rate) and the exponential model to fit area versus the expected number of species:

$$S = c + z \ln A \quad (3)$$

This allows species diversity for an area of a specific size to be calculated. An advantage of the power function is that its parameters c and z can be interpreted in terms of species diversity. The parameter c (y-intercept) of the power function indicates the expected number of species within a sample standardized for $A = 1$. The parameter z (slope) denotes the species accumulation rate, which indicates β-diversity (Ricotta et al. 2002).

Results

Total rarefaction functions: comparison of regions

The total rarefaction functions for the seven biogeographic regions (Fig. 2) resemble a species area curve, where the addition of one unit area can add (could add, might add) many new species if the samples are small (i.e. at the beginning of the curve). For samples of many plots n, the addition of one unit area only adds a few new species. The endpoint of the curve is reached when

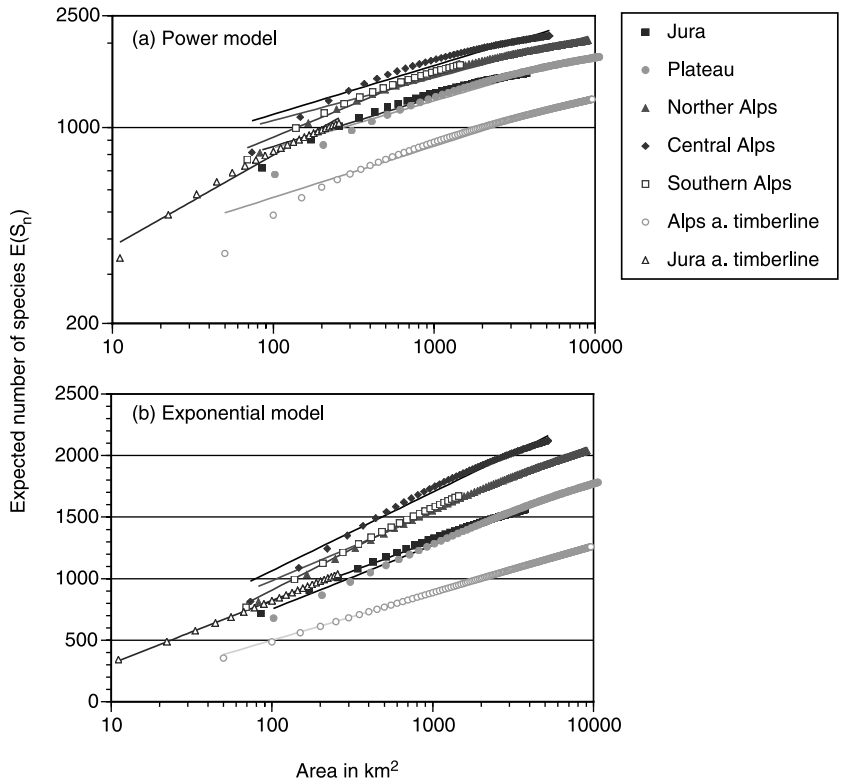

Fig. 2. Rarefaction functions for the biogeographic regions in Switzerland showing the expected number of species $E(S_n)$ as a function of area. The discontinuous rarefaction functions are fitted with a) the power model and b) the exponential model.

all of the plots are sampled. At this point the expected number of species is equal to the total number of species in the region, according to Table 2. Comparison of the regions shows that the species number per region is highest in the Central and Northern Alps, with 2120 species and 2041 species respectively, followed by the Plateau, the Southern Alps, and the Jura with 1561–1782 species (Table 2). The two regions above timberline show 1037 and 1259 species.

In terms of the number of species projected per 100 km^2, the species pool is highest in the Central Alps; slightly lower, but approximately equal, in the Northern and Southern Alps; and even lower and approximately equal in the Jura and the Plateau. For example, in a 1000 km^2 area we expect to find 1703 species in the Central Alps, 1579 species in the Southern Alps, 1537 species in the Northern Alps, 1306 species in the Jura, 1272 species in the Plateau, and 884 species in the Alps above timberline (Fig. 2). The accumulation rates are rather similar among the biogeographic regions (indicated by the nearly parallel right part of the curves). While the shape of the total rarefaction curves generally resemble one another, the curves for the Jura and Plateau intersect at approximately 2700 km^2. In general, the rank of the biogeographic regions remains rather stable over the modeled range in square kilometers (100–10 000 km^2).

The correlograms (Fig. 3) reveal the distance dependence of the polygons' floristic compositions in the seven groups. In the Jura, the Plateau, and the Northern Alps, polygons are autocorrelated to a distance of 80 km. Polygons of the central Alps have low correlation values, relative to the Southern Alps and the Jura above timberland. Polygons of the Alps above the timberline are autocorrelated to a distance of 190 km.

Table 2. Number of vascular plant species S in Swiss regions j.

Region	S_j^{total}	S_i^{mean}	$S_i^{Std\ Dev}$	S_i^{min}	S_i^{max}	$N_{j,i}$
Jura	1561	718	114	490	950	44
Jura above timberline	1037	342	104	159	519	23
Plateau	1782	679	106	478	1078	104
Northern Alps	2041	813	142	542	1411	110
Central Alps	2120	814	127	517	1116	71
Southern Alps	1670	768	66	641	907	21
Alps above timberline	1259	356	73	134	551	192

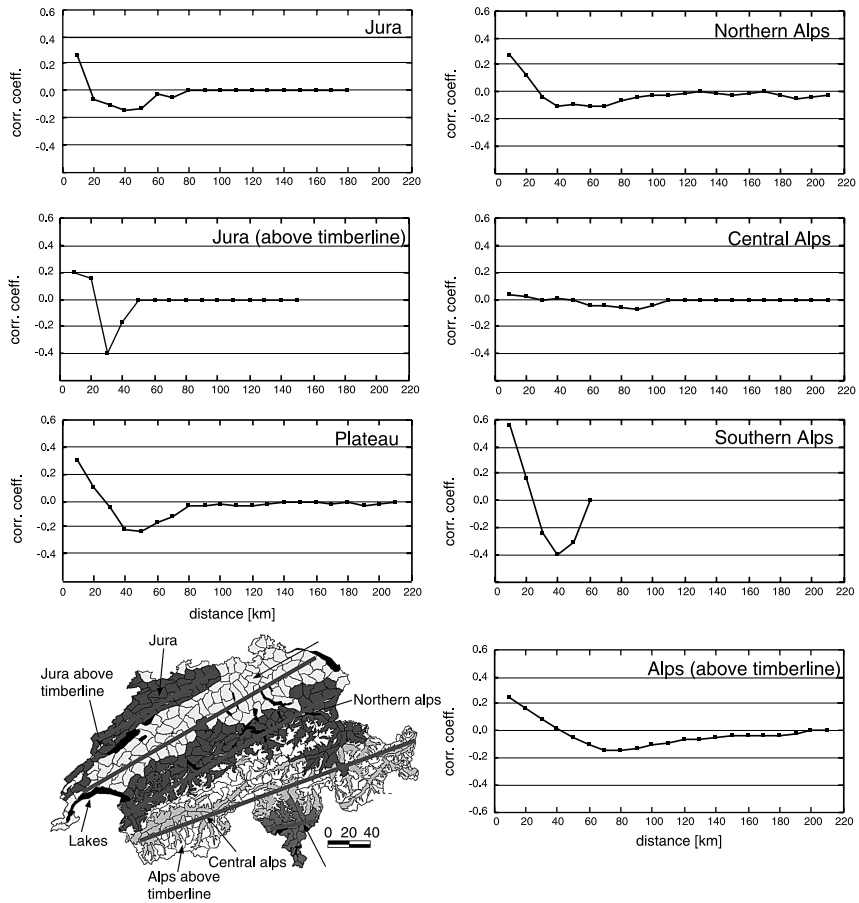

Fig. 3. Autocorrelation for Swiss regions.

Partial rarefaction functions: comparison of ecological groups

Partial rarefaction functions analyze the species pool with respect to ecological species groups (Fig. 4). In all of the biogeographic regions and for areas of 500 km^2 or less, we find the fewest species in the fertilized meadow group. We find the most species accumulated in either the pioneer/weedy species group or the mountain species group. The forest species group is ranked either second or third; the marsh/water species group is ranked either third, fourth, or fifth; and the unfertilized meadow group is ranked fourth or fifth.

The partial rarefaction curve for the fertilized meadow species is almost a straight horizontal line in all biogeographic regions. There are about 70 of these fertilized meadow species and even any small sample area of about 80 km^2 tends to contain all of them. In addition, the curves for forest species saturate rather quickly relative to other species groups. The curves for the pioneer/weedy species, in contrast, are less steep and their slope for the largest range continuously increases in all of the biogeographic regions. Therefore, in a small sample region, one can only expect to find a rather small number of the potential pioneer/weedy species. The area above timberline clearly differs from other regions. In it, we find mountain species and more than 40 fertilized meadow species are widespread, but there are few pioneer/weedy species.

Fitting species area relationships to rarefaction functions

The fitted parameters c and z in equations 2 and 3 differ from region to region and across species groups (Tables 3 and 4). The R squared values for the power model are generally between 0.80 and 0.98, since the power function is not fitted to a scatter of data points, but rather to a discontinuous function (Fig. 2a). For the central Alps and the Alps above timberland, the power fitted line is biased in comparison to a rarefaction function, especially for small areas. The R squared values for the exponential fit (0.98–1.00) are even higher, demonstrated by the straight line of points in the semilog

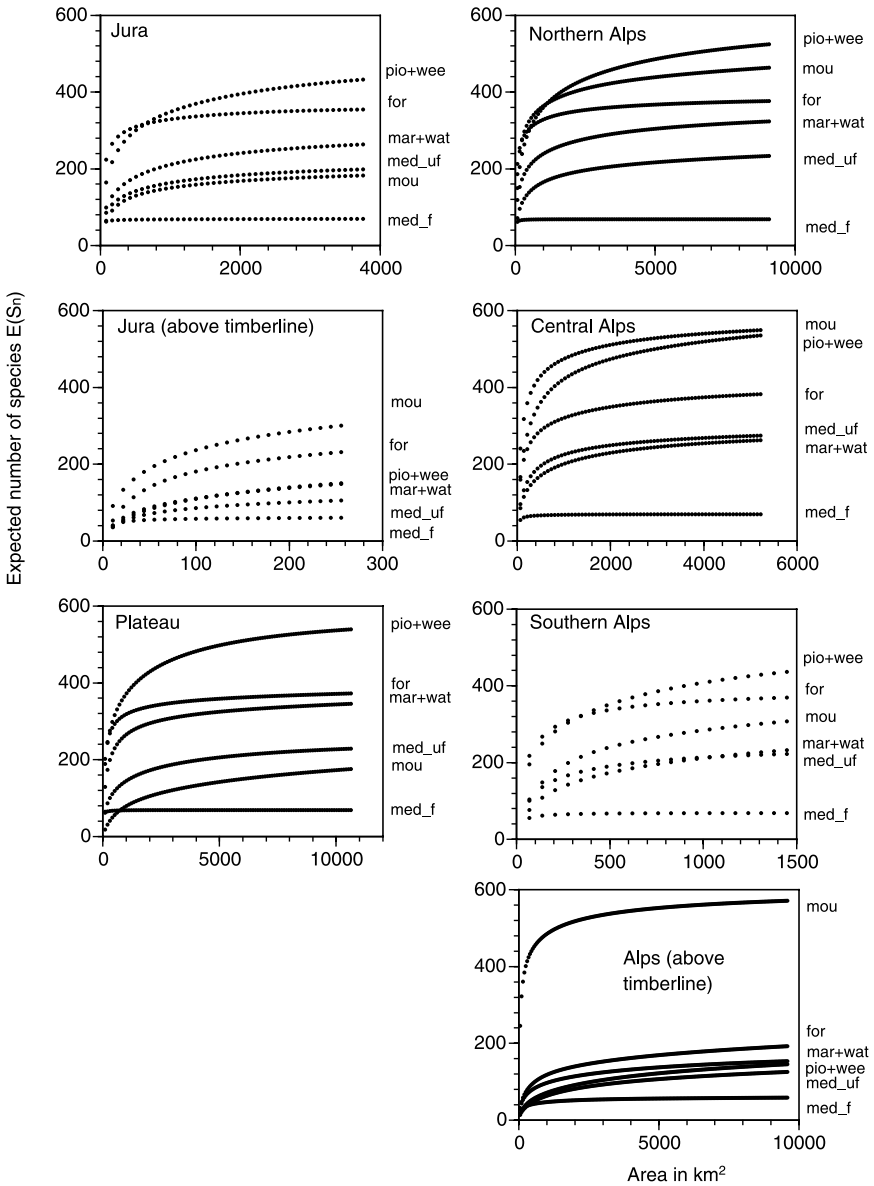

Fig. 4. Expected number of species $E(S_n)$ based on partial rarefaction for different ecological plant species groups (for: forest species, med_f: fertilised meadow species, med_uf: unfertilised meadow species, mou: mountain species, mar+wat: marsh/water species, pio+wee: pioneer/weed species) and regions.

plot (Fig. 2b). The exponential function results in a much better fit, especially for small regions. Based on the fitted exponential function, the total expected number of species as well as the expected number of species associated with the various ecological groups was calculated for a standardized area of 100 km^2 (Fig. 5).

Table 3a. Fitted curve parameters for power model $E(S) = cA^z$ with E(S) for 10, 100, 1000 km^2.

Region	c	z	R^2	$E(S) = c10^z$	$E(S) = c100^z$	$E(S) = c1000^z$
Jura	380	0.176	0.96	570	855	1282
Jura above timberline	178	0.325	0.96	377	797	1685
Plateau	392	0.167	0.96	576	846	1243
Northern Alps	523	0.153	0.95	744	1058	1504
Central Alps	498	0.174	0.93	743	1109	1656
Southern Alps	312	0.235	0.97	537	923	1587
Alps above timberline	243	0.183	0.98	370	563	857

Table 3b. Fitted parameters for exponential model E(S) = c+zln(A) and E(S) for 100, 1000 km^2.

Region	c	z	R^2	E(S) = c+zln(10)	E(S) = c+zln(100)	E(S) = c+zln(1000)
Jura	−156	212	0.99	331	819	1306
Jura above timberline	−205	223	1.00	309	824	1338
Plateau	−285	225	0.99	234	753	1272
Northern Alps	−124	240	0.99	430	983	1537
Central Alps	−211	277	0.98	427	1065	1703
Southern Alps	−437	292	1.00	235	907	1579
Alps above timberline	−268	167	1.00	116	384	884

Discussion

Comparison of regional species pools

The differences among biogeographic regions in number of species (Fig. 2) come from differences in species richness per habitat or differences in habitat diversity. We expect there to be a combination of the two effects. For example, the reason why the Central Alps, where we find the highest species richness, has more habitat types than the other regions might be because it stretches over a wider topographic and climatic range (Wohlgemuth 1998). Similarly, the fact that the range of soils, topography, and climate in the Northern and Southern Alps is wider than in the Jura and Plateau contributes to the species pool being larger (Wohlgemuth 2002).

Species area relationships are generally described for true islands, for habitat islands such as forest fragments within a landscape, and for heterogeneous regions consisting of habitat mosaics. It is interesting to compare our results for the slope parameter z of the power function with literature data. Parameter z typically ranges from 0.20 to 0.40 for habitat islands and heterogeneous regions (Connor and McCoy 1979, Connor et al. 1983). For true islands with a high degree of isolation, the slope is expected to be lower and vary from 0.12 to 0.19 (Preston 1960, Connor and McCoy 1979). Differences are explained by the habitat-diversity hypothesis and the area-per-se hypothesis (Preston 1962, Connor and McCoy 1979, Kohn and Walsh 1994, Tjørve 2002a, Triantis et al. 2003). In our investigation, the slope parameter varies from 0.15 to 0.33 for total rarefaction functions (Table 3), confirming that the landscapes studied do not consist of true islands. One has to be cautious, however, when interpreting these results, because the slope parameter can be scale dependent (He and Legendre 1996).

The importance of a regional species pool for local species diversity has been investigated and discussed extensively (Cornell and Lawton 1992, Pärtel et al. 1996, Zobel 1997, 2001, Caley and Schluter 1997, Zobel and Liira 1997). In most cases, the relation between local and regional pools was investigated using local-regional richness plots for two or three scales. To overcome the problem of estimating the overall relationship based on two or three measurements, Srivastava (1999) uses species area curves (log-log power model) and suggests using their parameters as measures for local-regional relationships. When curves in log-log plots are parallel with a constant slope parameter z, the ratio of local to regional richness remains constant. The constant slope parameter is interpreted as an indicator of unsaturated local communities. If the communities are saturated, the ratio between local and regional richness varies, and thus, slope parameter z is altered. Our results suggest local species pools to be unsaturated, since the slope parameters z for all species show low variability (Table 3). In contrast, ecological species groups show a large degree of variability in their slope parameters, which would indicate saturation of local communities. One concern with this interpretation is that the meaning of parameters of species area curves is still the subject of lively debate (e.g. Bartha and Ittzés 2001). Differences in slope parameters can be dependent on factors other than un/-saturation of local communities. In addition, the

Table 4. Fitted curve parameters for E(S) = c+zln(A) for each ecological species group separately (for: forest species, med_f: fertilised meadow species, med_uf: unfertilised meadow species, mou: mountain species, mar+wat: marsh/water species, pio+wee: pioneer/weed species).

Region	Group for		med_f		med_uf		mou		mar+wat		pio+wee	
	c	z	c	z	c	z	c	z	c	z	c	z
Jura	123	29	59	1.4	−43	30	−55	29	−94	44	−135	70
Jura above timberline	−86	58	33	5.3	−16	22	−76	68	−49	35	−63	38
Plateau	117	28	63	0.6	−105	36	−182	38	−13	39	−143	75
Northern Alps	132	28	64	0.6	−65	33	6	51	−49	41	−185	79
Central Alps	−3	46	53	2.1	−43	38	42	61	−85	41	−147	81
Southern Alps	43	46	45	3.4	−49	38	−182	68	−148	52	−143	80
Alps above timberline	−103	32	3	6.3	−108	25	152	47	−72	25	−128	29

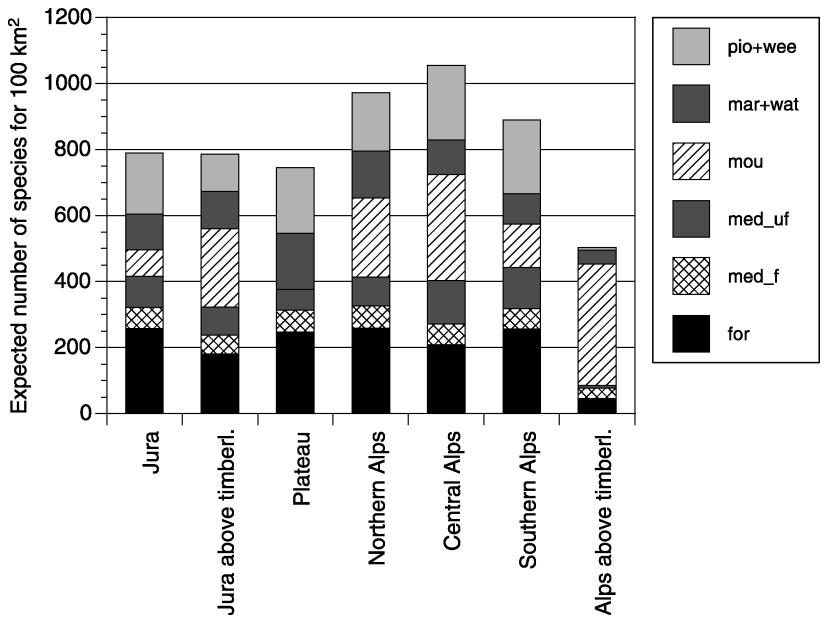

Fig. 5. Expected number of species calculated for 100 km^2. Contribution of different plant species groups in the investigated regions (see Fig. 4 for abbreviations).

definition of local scale is ambiguous. Srivastava (1999) defined it as being small enough that all of the species contained within could potentially interact over ecological time scales. Typically, this would correspond to the number of fish species in a lake or the number of plant species in a forest. In our investigation, however, the local scale is determined by the size of the polygons, which have an average area of 70 km^2 and are typically heterogeneous in terms of habitat.

Contribution of plant species groups to regional species pools

The partial rarefaction functions make assessing the contribution of different plant species groups to the total regional species pool possible. In multi-habitat landscapes, the total species pool is composed of different ecological species groups bound to different habitat types. The theoretical work of Tjørve (2002a) suggested that it is possible to combine the individual species area curves of the habitat types in a landscape. Because of the potential for a species overlap between habitats, we suggest calculating partial species area curves for ecological species groups rather than for habitat types.

Differences in the total number of species expected between regions (Fig. 5) mainly have to do with the absence or occurrence of the mountain species. This number is far higher in the Northern Alps and Central Alps than in the Swiss Plateau. The alpine region above timberline shows the number of species expected to be low, which can be explained by altitudinal richness patterns (e.g. Grytnes 2003). In contrast, the numbers of marsh/water marsh/water species are largest in the plateau and lowest in the alpine regions. The other species groups show the expected number of species among regions to be similar. For example, in each region (the area above timberline excluded) about 65 fertilized meadow species can be found in the species pool. The partial rarefaction functions for fertilized meadow species show a horizontal asymptote (Fig. 4) and a low value for slope parameter z (Table 4) of the fitted exponential function. This indicates that species of this group are ubiquitary, occupying a rather small, but frequented habitat segment defined by regular cutting and fertilizing. They contribute little to the species pools' diversity; in other words, the β-diversity – the between habitat diversity – of this species group is low. Having high ubiquitarity in tandem with low β-diversity enhances the likelihood of occurrence of species at any given site of the corresponding habitat type, as the random model for avian diversity of Blackburn and Gaston (2001) suggests. Many of the plant species in this group probably only appeared in the last few thousand years through ecological differentiation (Landolt 1991, p. 34). Interestingly, when the scale is expanded to 500 km^2, the number of mountain species exceeds those of fertilized meadow species, even on the Plateau or in the Jura. Many mountain species in these regions are found in relictic places where land use is absent, or on oligotrophic grasslands. The curves for forest species also saturate rather quickly. This might be due to the

relatively narrow ecological amplitude of forest ecosystems, shaped by shadow and moderate temperature changes. In contrast, the species accumulation rate z is rather high for the marsh/water species groups and for the mountain species groups. The shape of species area curves can vary considerably from species group to species group and might have a horizontal asymptote as well, as in the case of fertilized-meadow species (Williamson et al. 2001).

Rarefaction function with fitting procedure for constructing species area relationships

The nonlinear shape of species area relationships has been widely discussed (Arrhenius 1921, Connor and McCoy 1979, Connor et al. 1983, Lomolino 2000, Williamson et al. 2001, Tjørve 2002b). In general, curves with a convex shape, like linear, log-log power functions (Arrhenius 1921, Preston 1962) or exponential functions are used for regression of the field data. According to Tjørve (2002b), recent research has suggested that species area curves are sigmoid (e.g. logistic function, Hill function, or Lomolino function) with an upper asymptote (He and Legendre 1996, Lomolino 2000). Until further research has confirmed the advantages of these functions, however, we consider application of the log-log power function and exponential functions to be appropriate.

Diversity measures can be scale dependent, a fact not addressed in this research. Indexes based on multifractal geometry can be appropriate tools for gaining a deeper understanding of the spatial implications and scale dependency of diversity measures (Ricotta 2000, Yue et al. 2001).

Using rarefaction together with curve fitting methods for investigating regional species pools bears some specific problems. The parameters of the species area function are based on smoothed rarefaction curves, not on scattered species area relationships. Unlike conventional measures for β-diversity, based on presence/absence data (for a review, see Koleff et al. 2003), our measure is calculated indirectly, which can lead to the following errors.

The smoothed rarefaction curves approximate the statistically expected species richness while subsequently sampling more plots (Gotelli and Colwell 2001). Slope parameter z can be biased, because the rarefaction method generates an artificial mixing of plots (i.e. geographically separate plots are sampled and pooled together). If one were to investigate joined plots in the landscape ranging from small to large for constructing species area relationships, one would find less heterogeneity of habitats within a single plot than in pooled plots of the same size (for further discussion of this issue, see Palmer 1990, 1991). According to Scheiner (2003), we are dealing with a type II data set, which is an array of equal sub-areas in a continuous grid. Information about spatial autocorrelation among polygons (i.e. closer polygons are probably more similar in terms of species richness than distant ones) is eliminated through applying the rarefaction technique to this kind of data (Wagner 2003). In this sense, the rarefaction technique can be considered a neutral model for diversity estimation. Calculated with nested subregions ranging from small to large (Type I), the accumulation rate z will fully integrate the spatial variability.

Strong autocorrelation in the polygons of the Alps above timberline is plausible. This is, because in these extreme habitats with frost harshness, the regional plant composition is fairly constant, except for the few species disjunctions that resulted from isolation processes due to glaciation. The number of plant species that are adapted to these conditions is smaller when compared to lists of polygons below timberline, and many of these adapted plant species are distributed along the whole Alpine arc. Most polygons are distinctly autocorrelated to a distance of about 40 km in the direction of the weakest floristic gradient.

A specific problem, which results from the data set used in this paper, is the differences in polygon size (Table 1). Frequently, the rarefaction method (eq. 1) is used to compare samples that contain different numbers of organisms N (Hurlbert 1971). The method makes it possible to calculate the expected number of species for standardized sample sizes. In this situation, the advantage of the rarefaction method is that differences in species number no longer depend on differences in sample size; they depend on the variety of the community structure. In our application, however, rather than assess the number of organisms of a species i in samples, we assessed the number of plots N_i where species i occurs. The probability of occurrence, and thus N_i, is dependent on the size of the sample plot. Particularly when assuming a heterogeneous distribution of species in space, this means that differences in sample sizes result in under or overestimation of species. Species distributed in areas with many small sample plots are overestimated, while those in areas with few large plots are underestimated. The size of sample plots in km^2 should, therefore, be similar for gaining best estimates with the rarefaction method.

In our data set, the plots were chosen according to topographical features (Fig. 1) and vary in size considerably. For example, small mountaintops can be found in the front ranges of the Alps, and larger mountain areas, in the more elevated central ranges. Many mountain species in the front ranges of the Alps are located merely on the top of isolated mountains (e.g. Brienzer Rothorn, Mythen, Pilatus). The definition of sample plots followed the principle of best distribution display and aimed at measuring high resolutions of plant

species distribution rather than species richness. As a consequence, small sample plots were defined for summit areas in the front ranges of the Alps and large plots for the mountain areas of the Grisons and the Valais (Wohlgemuth 2002). There is generally an overbalance of small sample plots with respect to mountaintops in the front ranges of the Alps. Therefore, rarefaction overestimates species occurring in these areas. This effect is expected to increase with the standard deviation in plot sizes. That is, the effect is expected to be large for the Southern Alps and the areas above timberline, but more moderate for the other regions (Table 1).

There are clear advantages to the use of a combination of the rarefaction method with fitting procedure. In many countries, distribution maps for plant species are available showing the occurrence for equal subregions in a continuous grid (type II data in Scheiner's systematic). The rarefaction method uses these data sets and takes fully into account how common or rare a plant species is. The use of a fit procedure allows continuous species-area relationships to be constructed out of this type of data. Standardization of the expected number of species for a specific size area is then possible.

Conclusion

Since the rarefaction method takes the species area relationship explicitly into account, it is more suitable for assessing species richness than simple richness measures. The method proved helpful for the comparative assessment of regional species pools and for the assessment of the contribution of species groups to overall diversity. Existing data on species distribution can be analyzed in a nested manner; that is, the species area curves for different species groups or habitats can be combined (Tjørve 2002a). Knowledge about species richness within a defined area on a regional level, as provided by the rarefaction curve, is useful for land use management and decisions on a national or international level. For example, one can compare the number of species in a national park, a preserve, or a nation using the rarefaction curve to assess whether the area of interest contains more species or less species than an ecologically similar area of the same size. Forecasting species richness of unknown areas, however, might be problematic due to regional differences in environmental conditions that influence species diversity (Ulrich and Buszko 2003).

Species turnover between subregions and species distribution within a region can best be quantified by the measurement of β-diversity. Species diversity depends on the heterogeneity of habitats. The decrease in habitat heterogeneity caused by human impact (especially silvocultural and agricultural land use activities) has been identified as an important driver for the decline of species distribution (Korneck and Sukopp 1988). Specifically in national or international monitoring schemes, β-diversity is a suitable indicator for quantifying habitat mosaic changes, because it reflects efforts to mitigate the effects of reductions in habitat heterogeneity caused by humans. Policy decisions and financial transfers should be based on comparative assessments of species diversity like this (Köllner et al. 2002).

Acknowledgements – We wish to thank Carlo Ricotta, Dept of Plant Biology, Univ. of Rome, for his valuable input and discussion on an earlier version of the manuscript.

References

Abele, L. G. and Walters, K. 1979. Marine benthic diversity: a critique and alternative explanations. – J. Biogeogr. 6: 115–126.

Achtziger, R., Nigmann, U. and Zwölfer, H. 1992. Rarefaction-Methoden und ihre Einsatzmöglichkeiten bei der zooökologischen Zustandsanalyse und Bewertung von Biotopen. – Z. Ökol. Natursch. 1: 89–105.

Alard, D. and Podevigne, I. 2000. Diversity patterns in grassland along a lanscape gradient in nortwestern France. – J. Veg. Sci. 11: 287–294.

Anon. 1992. Convention on biological diversity. – UN Environment Programme (UNEP) Nairobi, Kenya.

Anon. 2001. Global biodiversity outlook. – UN Environment Programme (UNEP), Montreal, Quebec.

Archibald, E. E. A. 1949. The species character of plant communities. II. A quantitative approach. – J. Ecol. 37: 260–274.

Arita, H. T. and Rodríguez, P. 2002. Geographic range, turnover rate and the scaling of species diversity. – Ecography 25: 541–550.

Arrhenius, O. 1921. Species and area. – J. Ecol. 9: 95–99.

Balvanera, P. et al. 2002. Patterns of bidiversity in a Mexican tropical dry forest. – J. Veg. Sci. 13: 145–158.

Bartha, S. and Ittzés, P. 2001. Local richness-species-pool ration: a consequence of the species-area relationship. – Folia Geobot. 36: 9–23.

Blackburn, T. M. and Gaston, K. J. 2001. Local avian assemblages as random draws from regional pools. – Ecography 24: 50–58.

Boucher, G. and Lambshead, P. J. D. 1995. Ecological biodiversity of marine nematodes in samples from temperate tropical and deep-sea regions. – Conserv. Biol. 9: 1594–1604.

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

Connor, E. F. and McCoy, E. D. 1979. The statistics and biology of the species-area relationship. – Am. Nat. 113: 791–833.

Connor, E. F., McCoy, E. D. and Cosby, B. J. 1983. Model discrimination and expected slope values in species-area studies. – Am. Nat. 122: 789–798.

Cornell, H. and Lawton, J. H. 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. – J. Anim. Ecol. 61: 1–12.

Douglas, M. and Lake, P. S. 1994. Species richness of stream stones: an investigation of the mechanisms generating the species-area relationship. – Oikos 69: 387–396.

Gjerde, I. and Sætersdal, M. 1997. Effects on avian diversity of introducing spruce *Picea* spp. plantations in the native pine *Pinus sylvestris* forests of western Norway. – Biol. Conserv. 79: 241–250.

Gleason, H. A. 1922. On the relation between species and area. – Ecology 3: 158–162.

Gleason, H. A. 1925. Species and area. – Ecology 6: 66–74.

Gonseth, Y. et al. 2001. Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard. – Swiss Agency for Environment, Forests and Landscape, Bern.
Gotelli, N. J. and Colwell, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. – Ecol. Lett. 4: 379–391.
Grytnes, J. A. 2003. Species-richness patterns of vascular plants along seven altitudinal transects in Norway. – Ecography 26: 291–300.
Gutersohn, H. 1973. Naturräumliche Gliederung. – In: Anon (ed.), Atlas der Schweiz. – Haeupler, Bern, folio 78.
He, F. and Legendre, P. 1996. On species-area relations. – Am. Nat. 148: 719–737.
Heck, K. L., van Belle, G. and Simberloff, D. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. – Ecology 56: 1459–1461.
Hector, A. et al. 1999. Plant diversity and productivity experiments in European grasslands. – Science 286: 1123–1127.
Heywood, V. H. (ed.) 1995. Global Biodiversity Assessment. – UN Environment Programme (UNEP), Cambridge Univ. Press.
Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. – Ecology 52: 577–586.
Kohn, D. D. and Walsh, D. M. 1994. Plant species richness – the effect of island size and habitat diversity. – J. Ecol. 82: 367–377.
Koleff, P., Gaston, K. J. and Lennon, J. J. 2003. Measuring beta diversity for presence-absence data. – J. Anim. Ecol. 72: 367–382.
Köllner, T., Schelske, O. and Seidl, I. 2002. Integrating biodiversity into intergovernmental fiscal transfers based on cantonal benchmarking: a Swiss case study. – Basic Appl. Ecol. 3: 381–391.
Korneck, D. and Sukopp, H. 1988. Rote Listen der in der Bundesrepublik Deutschland ausgestorbenen, verschollenen und gefährdeten Farn- und Blütenpflanzen und ihre Auswertung für den Arten- und Biotopschutz. – Schriftenreihe Vegetationsk. Vol. 19, Bonn-Bad Godesberg.
Landolt, E. 1991. Gefährdung der Farn- und Blütenpflanzen in der Schweiz mit gesamtschweizerischen und regionalen roten Listen. – Eidgenöss. Drucksachen- und Materialzentrale, Bern.
Legendre, P. and Fortin, M.-J. 1989. Spatial analysis and ecological modelling. – Vegetatio 80: 107–138.
Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species-area relationship. – J. Biogeogr. 27: 17–26.
MacArthur, R. H. 1965. Patterns of species diversity. – Biol. Rev. 40: 510–533.
Magurran, A. E. 1996. Ecological diversity and its measurement. – Chapman and Hall.
Moser, D. M. et al. 2002. Rote Liste der gefährdeten Farn- und Blütenpflanzen der Schweiz. – Bundesamt für Umwelt, Wald und Landschaft, Bern.
Nekola, J. C. and White, P. S. 1999. The distance decay of similarity in biogeography and ecology. – J. Biogeogr. 26: 867–878.
Palmer, M. W. 1990. The estimation of species richness by extrapolation. – Ecology 71: 1195–1198.
Palmer, M. W. 1991. Estimating species richness: the second-order jackknife reconsidered. – Ecology 72: 1512–1513.
Palmer, M. W. and White, P. S. 1994. Scale dependency and the species-area relationship. – Am. Nat. 144: 717–740.
Pärtel, M. et al. 1996. The species pool and its relation to species richness: evidence from Estonian plant communities. – Oikos 75: 111–117.
Preston, F. W. 1960. Time and space and the variation of species. – Ecology 41: 611–627.
Preston, F. W. 1962. The canonical distribution of commonness and rarity: part I and II. – Ecology 43: 185–215, 410–432.
Ricotta, C. 2000. From theoretical ecology to statistical physics and back: self-similar landscape metrics as a synthesis of ecological diversity and geometrical complexity. – Ecol. Modell. 125: 245–253.
Ricotta, C., Carranza, M. L. and Avena, G. 2002. Computing β-diversity from species area curves. – Basic Appl. Ecol. 3: 15–18.
Ricotta, C., Chiarucci, A. and Avena, G. 2004. Quantifying the effects of nutrient addition on community diversity of serpentine vegetation using parametric entropy of type α. – Acta Oecol. 25: 61–65.
Scheiner, S. M. 2003. Six types of species-area curves. – Global Ecol. Biogeogr. 12: 441–447.
Shannon, C. 1948. A mathematical theory of communication. – Bell System Tech. J. 27: 379–423.
Simberloff, D. S. 1978. Use of rarefaction and related methods in ecology. – In: Dickson, K. L., Cairns Jr, J. and Livingston, R. J. (eds), Biological data in water pollution assessment: quantitative and statistical analysis. Am. Soc. Test. Mat., Philadelphia, STP 652, pp. 150–165.
Simpson, E. H. 1949. Measurement of diversity. – Nature 163: 688.
Sokal R. R. 1986. Spatial data analysis and historical pricesses. – In: Diday, E. et al. (eds), Data analysis and informatics, IV. Proc. Fourth Int. Symp. Data Analysis Informatics. North Holand pp. 29–43.
Srivastava, D. S. 1999. Using local-regional richness plots to test for species saturation: pitfalls and potentials. – J. Anim. Ecol. 68: 1–16.
Tjørve, E. 2002a. Habitat size and number in multi-habitat landscapes: a model approach based on species-area curves. – Ecography 25: 17–24.
Tjørve, E. 2002b. Shapes and functions of species–area curves: a review of possible models. – J. Biogeogr. 30: 827–835.
Triantis, K. A. et al. 2003. A model for the species–area–habitat relationship. – J. Biogeogr. 30: 19–27.
Ulrich, W. and Buszko, J. 2003. Species–area relationships of butterflies in Europe and species richness forecasting. – Ecography 26: 365–373.
Wagner, H. H. 2003. Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. – Ecology 84: 1045–1057.
Welten, M. and Sutter, R. 1982. Verbreitungsatlas der Farn- und Blütenpflanzen der Schweiz. – Birkhäuser, Basel.
Whittaker, R. H. 1972. Evolution and measurement of species diversity. – Taxon 21: 213–251.
Whittaker, R. J., Willis, K. J. and Field, R. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. – J. Biogeogr. 28: 453–470.
Wildi, O. and Orlóci, L. 1996. Numerical exploration of community patterns, 2nd ed. – SPB Acad. Publ., The Hague.
Williamson, M., Gaston, K. J. and Lonsdale, W. M. 2001. The species-area relationship does not have an asymptote! – J. Veg. Sci. 28: 827–830.
Wisheu, I. C. and Keddy, P. 1996. Three competing models for predicting the size of species pools: a test using eastern North American wetlands. – Oikos 76: 253–258.
Witte, J. P. M. and Torfs, P. J. J. F. 2003. Scale dependency and fractal dimension of rarity. – Ecography 26: 60–68.
Wohlgemuth, T. 1993. Der Verbreitungsatlas der Farn- und Blütenpflanzen der Schweiz (Welten und Sutter 1982) auf EDV: Die Artenzahl und ihre Abhängigkeit von verschiedenen Faktoren. – Bot. Helv. 103: 55–71.
Wohlgemuth, T. 1998. Modelling floristic species richness on a regional scale: a case study in Switzerland. – Biodiv. Conserv. 7: 159–177.
Wohlgemuth, T. 2002. Environmental determinants of vascular plant species richness in the Swiss alpine zone. – In: Körner, C. and Spehn, E. (eds), Mountain biodiversity. A global assessment. Parthenon, pp. 103–116.

Yue, T. et al. 2001. Changes of Holdridge life zone diversity in all of China over half a century. – Ecol. Modell. 144: 153–162.
Zobel, K. 2001. On the species-pool hypothesis and on the quasi-neutral concept of plant community diversity. – Folia Geobot. 36: 3–8.
Zobel, K. and Liira, J. 1997. A scale-independent appraoch to the richness vs biomass relationship in ground-layer plant communities. – Oikos 80: 325–332.
Zobel, M. 1997. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? – Trends Ecol. Evol. 12: 266–269.

Appendix: deduction of the rarefaction function

The rarefaction function is derived from hypergeometric distribution. This distribution defines the probability of having exactly k successes in a sub-sample of size n, which is chosen without replacement out of a sample of the size N with N_i successes. It is defined as follows

$$P(k) = \frac{\binom{N - N_i}{n - k}\binom{n}{k}}{\binom{N}{n}}$$

The probability $P(k = 1,\ldots,N_i)$ of choosing $k = 1,\ldots,N_i$ individuals of species S_i is equal to the probability $1 - P(k = 0)$ of not choosing $k = 0$ individuals of that species. It is

$$P_i(k = 1, ..., N_i) = 1 - P(k = 0)$$

$$P_i(k = 1, ..., N_i) = 1 - \frac{\binom{N - N_i}{n - 0}\binom{n}{0}}{\binom{N}{n}}$$

$$= 1 - \frac{\binom{N - N_i}{n}\frac{n!}{0!(n - 0)!}}{\binom{N}{n}} = 1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}}$$

The expected value for the number of species, when choosing n plots out of a sample of N plots, is the sum of the probabilities P_i times the value v_i of species occurrence/presence:

$$E(S_n) = \sum_{i=1}^{S} v_i * P_i(k = 1, ..., N_i)$$

where

$$v_i = \begin{cases} 0 \text{ for species absence} \\ 1 \text{ for species presence} \end{cases}$$

The resulting rarefaction curve, according to Hurlbert and Simberloff, is

$$E(S_n) = \sum_{i=1}^{S} 1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} = S - \frac{\sum_{i=1}^{S}\binom{N - N_i}{n}}{\binom{N}{n}}$$