

POLISH JOURNAL OF ECOLOGY (Pol. J. Ecol.)	48	1	21–35	2000
--	----	---	-------	------

Werner ULRICH

Department of Animal Ecology, Nicholas Copernicus University in Toruń, Gagarina 9, 87-100 Toruń;  
Poland, e-mail: ulrichw@cc.uni.torun.pl

## ON SPECIES–AREA RELATIONSHIPS II: SLOPE AND FACTOR VALUES OF POWER FUNCTION AND EXPONENTIAL MODEL

**ABSTRACT:** Using model assemblages generated by a FORTRAN program the parameter values of the slope of the power function and the factor of the exponential model of species–area relationships have been studied. It appeared that the slope value is not a constant independent of area and sampling method but depends strongly on grain, sampling method, and model fit. The fraction of singletons in the sample proved to be of major importance. A plot of slope against assemblage structure (estimated by the standard deviation of  $\log_2$  (densities)) was bell shaped with the highest slope values at intermediate SD values. A comparison of this plot with SD values from theoretical relative abundance distributions showed that log-normal distributed assemblages should have slope values that are higher than previously reported in the literature. Although it was impossible to predict the slope from the relative abundance distribution, the opposite was possible. At any given slope value there are two linked relative abundance distributions. The factor of the exponential model was more independent of sampling methods but linearly connected with sampling efficacy. A high non-linear correlation between factor and Shannon diversity was detected and a general function of this relationship developed and tested. The factor of the exponential species–area relationship may serve as an estimate of regional diversity.

**KEY WORDS:** species–area relationship, model species assemblages, diversity, evenness, relative abundance distributions, random sampling

### 1. INTRODUCTION

The question how to interpret the parameters of species–area relationships have intrigued ecologists from the beginning (Arrhenius 1923). Especially the slope value of the power function model caught much interest. Preston (1962), MacArthur and Wilson (1967), May (1975), and Sugihara (1981) derived a slope value around 0.25 if the underlying relative abundance distribution followed a log-normal distribution. Because of the popularity of the log-normal model and the seemingly mathematical strength of the derivation the slope parameter was since then widely and often uncritically used in papers and ecological textbooks.

However, a wide range of slope values appeared (Connor and McCoy 1979, Martin 1981), with values even above 1 (Strong 1974, Rosenzweig and Sandlin 1997), values not covered by former (Schoener 1976) and recent theories (Harte *et al.* 1999). It appeared also that mainland populations have frequently smaller slope values than islands (Preston 1962, Begon *et al.* 1986, Rosenzweig 1995, Hanski and Gyllenberg 1997, Ney-Nifle and Mangel 1999, Ulrich

1999), that the slope values are scale dependent (Palmer and White 1994, He and Legendre 1996, Rosenzweig 1999, Ney-Nifle and Mangel 1999) and that they rely on species turnover rates, evolutionary time, latitude, remoteness, and even food web structures (MacArthur 1965, Wilson and Taylor 1967, Whitehead and Jones 1969, Begon *et al.* 1986, Rosenzweig 1995, Harte and Kinzig 1997). Recently, Leitner and Rosenzweig (1997) examined nested species–area curves derived from model log-normal relative abundance distributions and found slope values between 0.48 and 0.83, which led them state that for such nested sampling a different theory is necessary which is in accordance to the observed lower values.

The biological interpretation of the slope value remained unclear. Connor and McCoy (1979) thought that the slope value is nothing more than a statistical artefact, resulting from what they called ‘passive sampling’. The statistical derivations of power function SPARs of Wissel and Maier (1992) and Harte *et al.* (1999) point to the same direction. However, extensive studies on natural or model assemblages relating slope values and other ecological parameters of interest (diversity, community structure, habitat properties, scale, species turnover and others) are largely missing and urgently needed to get a better understanding of the properties of the slope parameter.

Interestingly, the factor value of the exponential function has gained much less interest. Connor and McCoy (1979) in their extensive review on the matter even do not mention any possible interpretation of the factor value and this has not changed up to now. The factor is merely seen as a fitted constant describing a special case without any further biological meaning.

The aim of this study is to reveal the relation of the slope and the factor values with other ecological or sampling parameters. It will be shown that especially the factor of the

exponential function deserves much more interest and can be interpreted as an estimate of regional diversity tightly connected with other measures of diversity but also with community structure.

## 2. MATERIAL AND METHODS

The assemblages used in this study are the same as in part I (Ulrich 2000a), the generation procedure, the properties of these assemblages, and the generating variables are listed in Table 1 of Part I (Assemblages 1 to 10). For each of these 728 assemblages species area curves of the exponential and the power function type were computed and the slope value  $z$  (power function) and the factor  $b$  (exponential function) recorded (Formulas 1 and 2 in Part I). For all assemblages the total number of species  $S_a$ , the species density  $S_{unit}$ , the relative species density  $rel. S_a$ , the number of singletons in the assemblage, the number of species found ( $S_s$ ), the standard deviation of  $\log_2$  (densities) (SD), Shannon diversity and evenness (in the latter three cases separately for the total assemblage and the sample) were computed. Again three different sampling regimes (sequential adding, nested, non-nested) will be studied.

## 3. RESULTS

### 3.1. INFLUENCE OF SAMPLING METHOD AND GOODNESS OF FIT ON THE PARAMETER VALUES

Figure 1 shows that fit of the power function model and slope value are not independent. Below values of variance explanation ( $R^2$ ) of 0.95 both variables are clearly correlated. In the model assemblages used in this study low slope values below 0.3 were often combined with a weak fit of the power function model. The exponential model, on the other hand, did also show such a correlation, but to a much lesser degree.

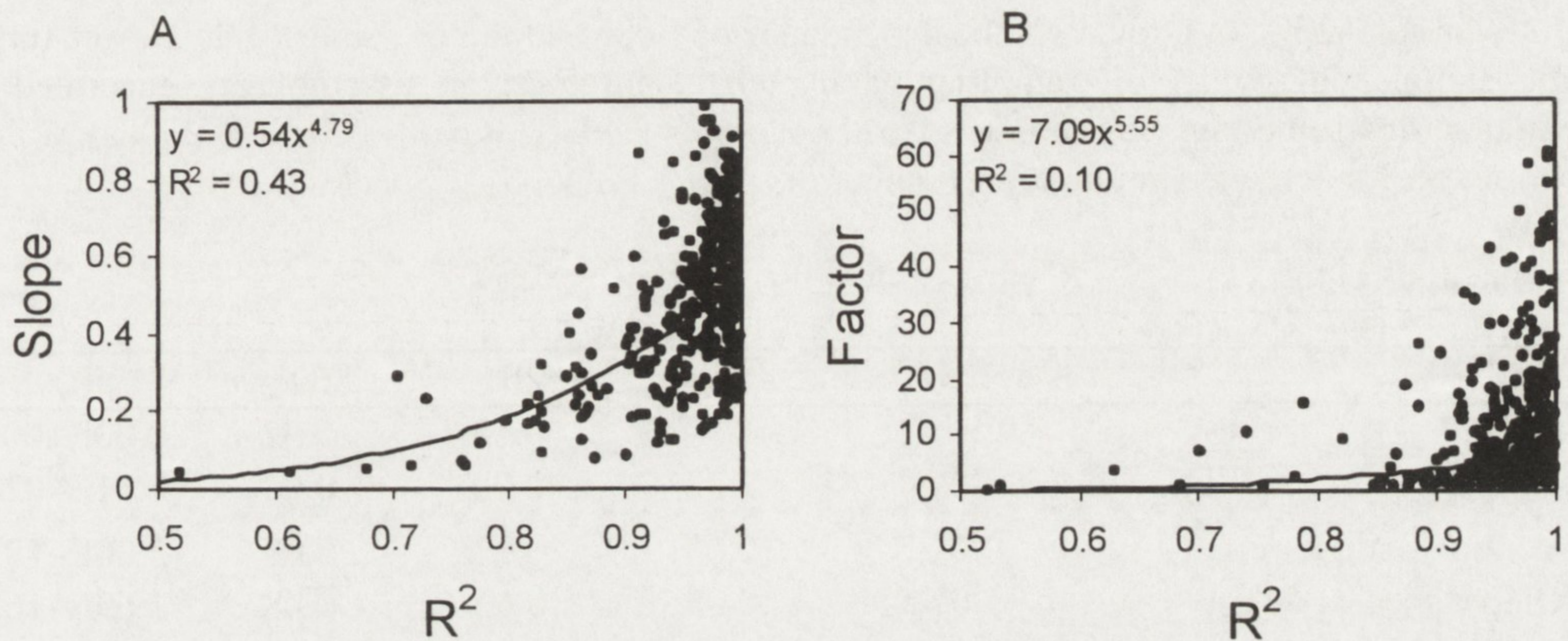


Fig. 1. Model fit (variance explanation  $R^2$ ) and (A) slope of the power function model or (B) factor of the exponential model of SPARs. Assemblages 1 and 4 to 6 of Table 1 in Ulrich (2000a). Sampling method: sequential adding.  $R^2$ : variance explanation

Sampling method and parameter values were also connected (Table 1). With the same types of assemblages sequential adding sampling resulted in the highest slope parameters, non-nested sampling in intermediate values, and a nested sampling design produced the lowest slope values. The variance of slope values also differed. The variance/mean ratio was highest in the sequential adding samplings (0.09), intermediate in non-nested samplings (0.06), and least in nested ones (0.04). In the exponential function no such dependencies could be detected.

species density ( $S_{unit}$ ), number of species sampled ( $S_s$ ), relative species density ( $S_{unit}/S_a$ ),  $S_{unit}/S_s$ , No. of singletons, fraction of singletons in the sample (singletons/ $S_s$ ), and fraction of species found ( $S_s/S_a$ ). Table 2 contains the results of a stepwise multiple regression of these variables with the slope and the factor and shows that the slope values are more connected with relative measures (fraction of singletons, rel. species density,  $S_{unit}/S_s$ , proportion of species found), whereas the factor is more connected with absolute measures (species density, No. of species found,

Table 1. Slope values (power function) and factors (exponential function) of SPAR models in dependence of sampling method. Data from assemblages No. 1, 2, 3 in Part I (Ulrich 2000a). These 120 assemblages each had the same underlying generating variable values. The differences in the slope values are significant at  $p(t) < 0.0001$ . The factor values do not differ significantly

Sampling method	Power function		Exponential function	
	Slope		Factor	
	Mean	StDev.	Mean	StDev.
Sequential adding	0.53	0.22	8.44	6.78
Non-nested design	0.35	0.14	9.00	6.09
Nested design	0.27	0.10	9.22	5.99

### 3.2. INFLUENCE OF SAMPLING-PARAMETERS

In a first step of analysis the variables connected with the sampling process were studied and the dependence of slope value  $z$  and factor on them. These variables were

No. of singletons). Figure 2 shows that especially the slope value is highly sensitive to the fraction of singletons in the sample. This view is emphasized by the partial correlation of the variables with  $z$  and factor (Table 3). Significant partial correlation with  $z$  have all

Table 2. Multiple regression to detect the dependence of slope and factor values of the power function and the exponential model on sampling parameters. Data from all assemblages generated with sequential adding sampling: Assemblages No. 1, 4 to 6 in Table 1 of part I (Ulrich 2000a).  $S_a$ : Total number of species in the assemblage,  $S_s$ : number of species found,  $S_{unit}$ : Species density.

Power function: Multiple  $r$ : 0.95,  $R^2$ : 0.90,  $F(5, 349)$ : 598,  $p(t) < 0.0001$

Variable	BETA	StDev.	B	StDev.	p(t)
Intercept			0.73	0.02	0.00E+00
$S_a$	-0.04	0.02	0.00	0.00	1.61E-02
Fraction of singletons	0.27	0.03	0.27	0.03	5.00E-18
Relative species density	0.45	0.05	0.66	0.08	1.69E-16
$S_{unit}/S_s$	-0.97	0.04	-1.30	0.06	0.00E+00
Proportion of species found	-0.25	0.04	-0.19	0.03	2.77E-10

Exponential model: Multiple  $r$ : 0.99,  $R^2$ : 0.99,  $F(3,351)$ : 54041,  $p(t) < 0.0001$

Variable	BETA	StDev.	B	StDev.	p(t)
Intercept			-0.06	0.04	1.49E-01
Species density	-0.65	0.01	-0.36	0.00	0.00E+00
$S_s$	1.39	0.01	0.32	0.00	0.00E+00
No. Singletons	0.06	0.00	0.06	0.00	5.44E-28

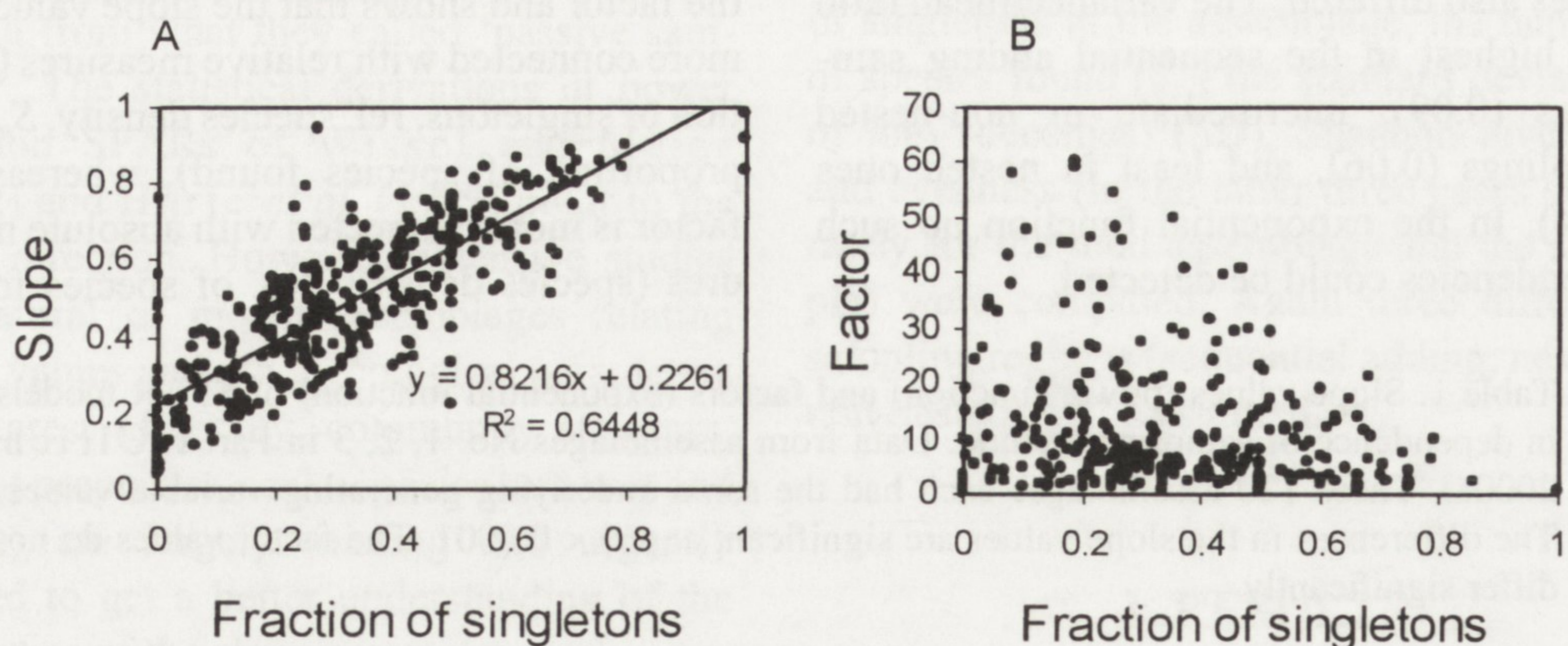


Fig. 2. Fraction of singletons in the sample and (A) slope of the power function model or (B) factor of the exponential model of SPARs. Assemblages 1 and 4 to 6 of Table 1 in Ulrich (2000a). Sampling method: sequential adding.  $R^2$ : variance explanation

of the relative measures, but of the absolute ones only  $S_a$ , in the case of the factor of the exponential function this reverses.

In principle, it is possible to construct a species–area curve out of only two data points, because for every combination of  $S_{unit}$  and  $S_s$  the slope or the factor values of both models are fixed. Indeed a correlation be-

tween  $\ln(S_{unit}/S_s)$  and slope and  $(S_s - S_{unit})$  and factor (using 728 assemblages, assemblages 1 to 8 of part I, Ulrich 2000a) resulted in correlation coefficients of 0.982 and 0.995. Therefore, under a random placement model there are indeed only two data point necessary if both data points have no variances, that is if the area is measured exactly and if all species had been sampled.

Table 3. Partial correlations of slope value (power function) and factor (exponential model) of SPAR models with all sampling and structural parameters. Data from all assemblages generated with sequential adding sampling: Assemblages No. 1, 4 to 6 in Table 1 of part I (Ulrich 2000a).  $S_a$ : Total number of species in the assemblage,  $S_s$ : number of species found,  $S_{unit}$ : Species density, SD: Standard deviation of  $\log_2$  densities of the assemblage. Significant values ( $p(t) < 0.01$ ) are marked in bold type

Variable	Power function		Exponential function	
	Part. Corr.	p(t)	Part. Corr.	p(t)
$S_{unit}$	-0.03	5.31E-01	<b>-0.97</b>	0.00E+00
$S_a$	<b>-0.20</b>	1.34E-04	-0.09	9.94E-02
$S_s$	0.12	2.55E-02	<b>0.99</b>	0.00E+00
No. Singletons	0.13	1.47E-02	<b>0.51</b>	1.74E-24
Fraction singletons	<b>0.39</b>	1.17E-13	-0.10	6.28E-02
Relative species density	<b>0.38</b>	2.87E-13	0.05	3.27E-01
$S_{unit}/S_s$	<b>-0.70</b>	0.00E+00	-0.03	6.39E-01
$S_s/S_a$	<b>-0.42</b>	3.86E-16	-0.08	1.44E-01
SD	-0.07	2.26E-01	-0.04	4.42E-01
Diversity	-0.07	1.96E-01	-0.01	8.02E-01
Evenness	0.16	3.33E-03	-0.01	9.04E-01
DWD slope	-0.01	8.94E-01	-0.03	5.88E-01
Heterogeneity	0.11	4.32E-02	0.05	3.20E-01

More interesting, however, are comparisons of the residuals of the above regression with variables describing underlying assemblage or habitat structures. Figure 3 shows that in the power function  $\ln(S_{unit}/S_s)$  slightly overestimates the fitted slope. This trend is independent of species density, SD and habitat heterogeneity, but more pronounced at a higher evenness of the assemblage. At low relative species densities the variance of the residuals is much higher than at medium and high rel. species densities. In the case of the exponential model higher rel. species densities caused an overestimation of the factor value. At high levels of evenness the variance of the residuals was also enhanced.

Which underlying structural variables of the assemblages influence the variable  $S_{unit}/S_s$ ? A multiple regression with  $S_{unit}/S_s$  as dependent and SD, diversity, evenness,  $S_a$ , DWD slope, and heterogeneity as independent variables revealed a complicated pattern. No single variable turned out to be dominant. Positively correlated were the species density ( $\beta$ -value = 0.56), SD ( $\beta$ -value = 0.13), and the

evenness ( $\beta$ -value = 0.51), negatively correlated were the diversity ( $\beta$ -value = -0.82), the DWD slope ( $\beta$ -value = -0.32), and the heterogeneity ( $\beta$ -value = -0.37) (total multiple r: 0.82).

In the exponential SPAR model there is of course no such relation between factor and  $S_{unit}/S_s$ . The main influencing variable on the factor value was the number of species found (Table 2). The dominating factors influencing the number of species found were the diversity (partial r: 0.81) and the evenness (partial r: -0.51) of the assemblage. Figure 4 shows an exponential relation between  $S_s$  and diversity, which is again dependent on sampling method (Fig. 4C).

### 3.3. INFLUENCE OF STRUCTURAL PARAMETERS

Is it possible to connect the slope and the factor values of both SPAR models with variables describing the structure of the assemblages and the heterogeneity of the area sampled. Table 4 gives the results of a step-

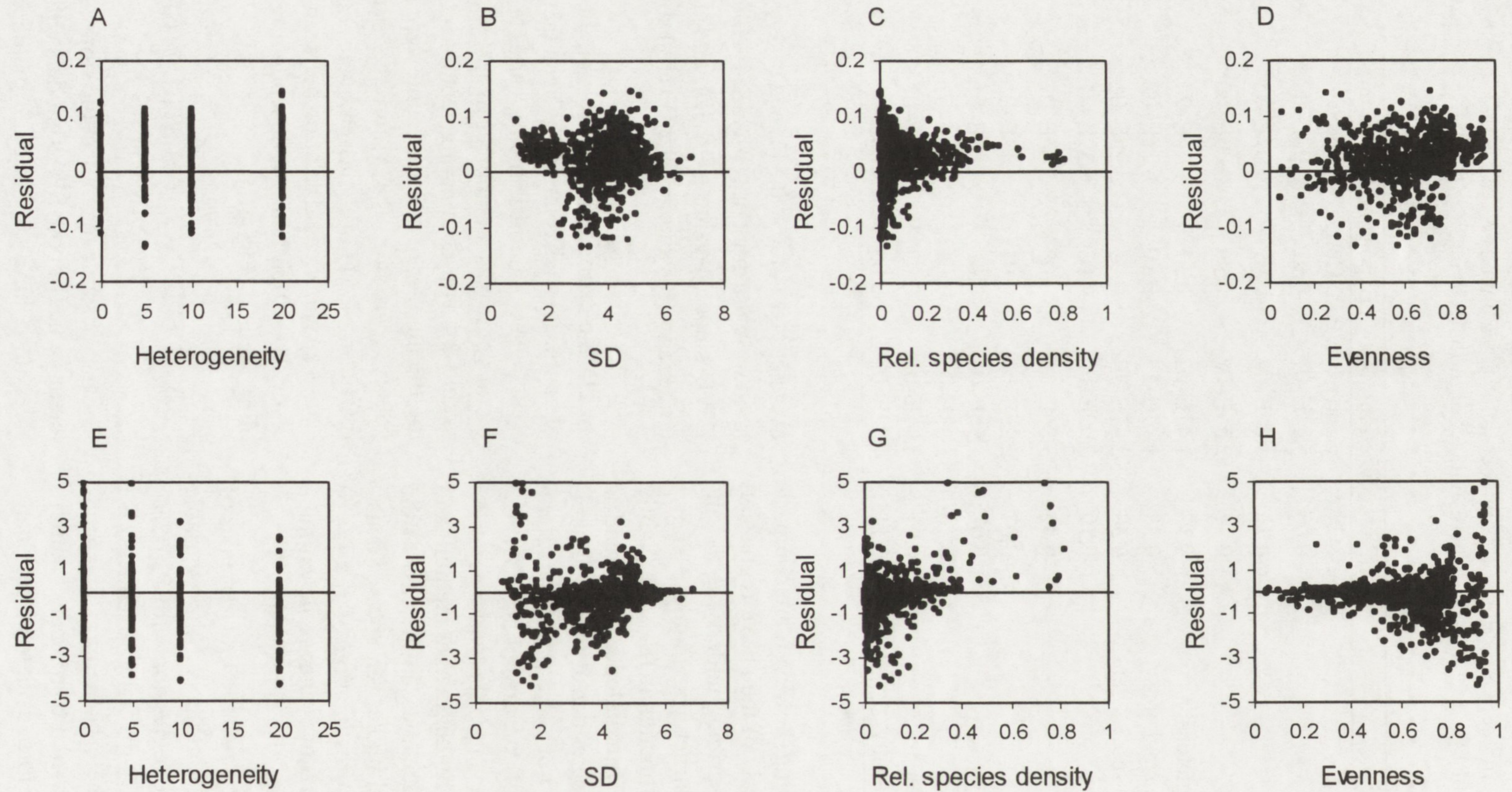


Fig. 3. Residuals (real-predicted values) of slope ( $z$ ) and factor ( $a$ ) values derived from the relations  $(1/\ln(\text{area}))(\ln(S_n/S_{\text{unit}})) = z$  and  $(S_n - S_{\text{unit}})/\ln(\text{area}) = a$  in dependence of area heterogeneity, standard deviation of  $\log_2$  (densities), relative species density, and evenness. A to D: power function, E to H: exponential model

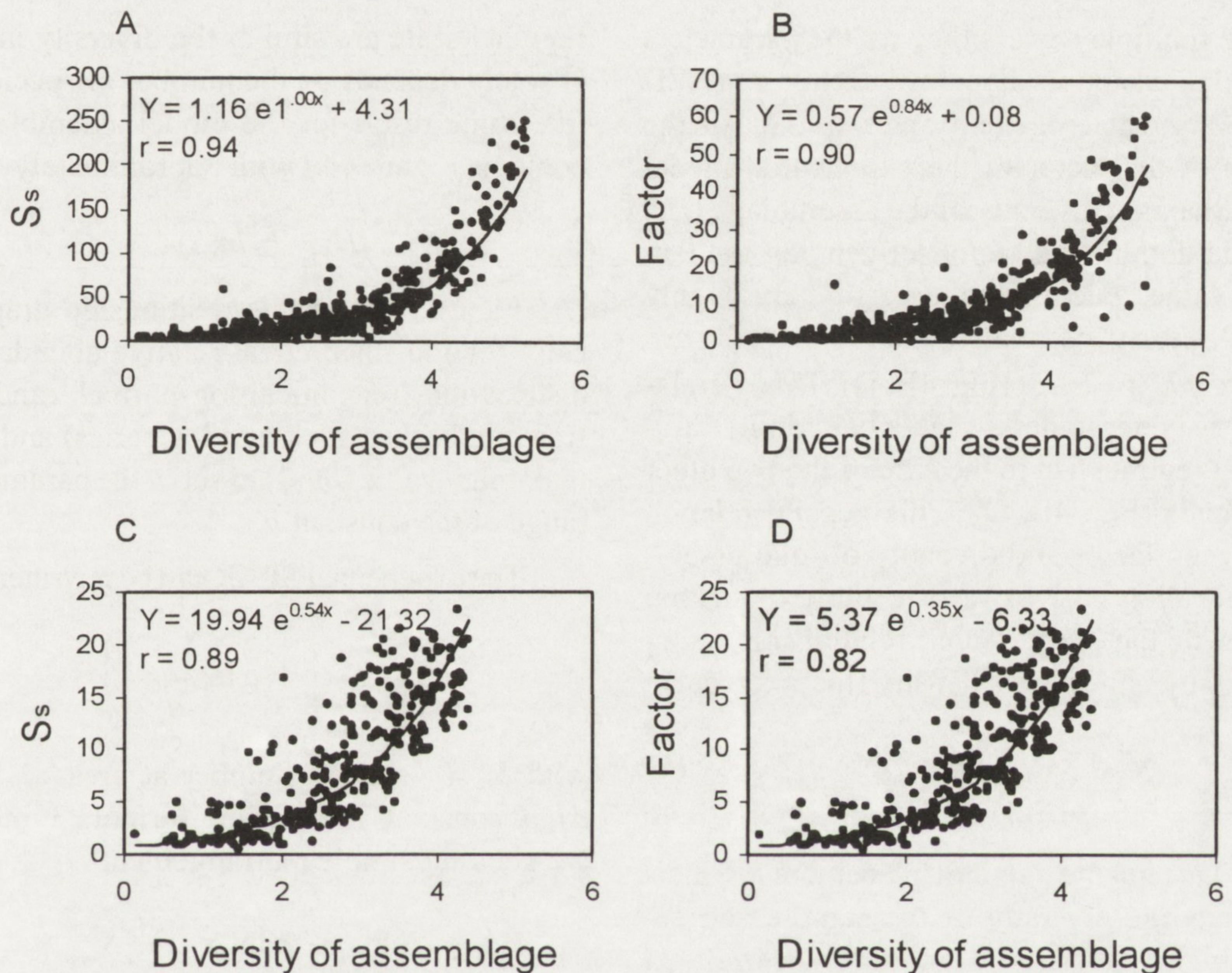


Fig. 4. Number of species sampled ( $S_s$ ) and factor of the exponential model in dependence of the Shannon diversity of the assemblage. A and B: sequential adding sampling (assemblages 1 and 4 to 6 of Table 1 in Ulrich 2000a); C and D: nested and non-nested sampling design combined (assemblages 2, 3, 7, and 8).  $r$ : Pearson's correlation coefficient

Table 4. Multiple regression to detect the dependence of Slope and factor values of the power function and the exponential model on the structure of the assemblages and area heterogeneity. Data from all assemblages generated with sequential adding sampling: Assemblages No. 1, 4 to 6 in Table 1 of part I (Ulrich 2000a).  $S_a$ : Total number of species in the assemblage,  $S_s$ : number of species found,  $S_{unit}$ : Species density, SD: Standard deviation of  $\log_2$  densities of the assemblage.

Power function: Multiple  $r$ : 0.74,  $R^2$ : 0.55,  $F(4, 386)$ : 116,  $p(t) < 0.0001$

Variable	BETA	StDev.	B	StDev.	p(t)
Intercept			0.13	0.03	9.85E-05
Diversity	0.36	0.07	0.07	0.01	1.15E-06
Evenness	-0.21	0.07	-0.21	0.07	3.96E-03
DWD slope	-0.53	0.04	-0.20	0.01	1.02E-37
Heterogeneity	0.61	0.03	0.02	0.00	0.00E+00

Exponential model: Multiple  $r$ : 0.85,  $R^2$ : 0.72,  $F(5,385)$ : 199,  $p(t) < 0.0001$

Variable	BETA	StDev.	B	StDev.	p(t)
Intercept			17.10	3.26	2.47E-07
SD	-0.33	0.04	-2.75	0.37	6.55E-13
Diversity	1.22	0.06	11.86	0.56	0.00E+00
Evenness	-0.77	0.07	-41.27	3.78	2.56E-24
DWD slope	0.19	0.03	3.92	0.70	4.84E-08
Heterogeneity	-0.07	0.03	-0.11	0.04	1.22E-02

wise multiple regression with the parameters  $S_a$ , SD, Shannon diversity, evenness, DWD slope, and heterogeneity and shows that in the case of the factor of the exponential model the Shannon diversity of the assemblage ( $H_a$ ) is the dominating factor influencing the factor value. Factor and diversity are highly non-linearly correlated [ $factor = 0.57 \exp(0.84 \times H_a) + 0.08$ ] (Fig. 4B, D). This correlation is better under a sequential adding sampling regime than in the case of the two other methods (Fig. 4B, D). This general relation between factor and diversity of total assemblage allows us to rewrite the exponential form of the species–area relationship using diversity and species density (the intercept)

$$S = [\alpha \exp(\beta H_a) + \gamma] \ln(\text{area}) + S_{unit} \quad (1)$$

Due to the random placement of individuals the diversity of the sample and assemblage diversity were highly correlated ( $H_a = 0.94 \times H_{sample}$ ;  $r = 0.96$ ), being the sample diversity a good estimator of latter, we can also denote

$$S = [\alpha \exp(\beta H_{sample}) + \gamma] \ln(\text{area}) + S_{unit} \quad (2)$$

The above relation also leads to a new interpretation of the factor value of the exponential model of the species–area relation. It is an estimate of regional diversity.

Formula 1 can also be obtained analytically. Under the assumption that the relative abundance distributions at the local and the

regional scale are similar the diversity index  $H$  solely depends on the number of species  $S$  (the same result for the model assemblages contains Figure 4A) with approximately

$$H \approx b \ln(S) \quad (3)$$

Figure 5 gives this relationship graphically for four theoretical relative abundance distributions (canonical log-normal, random fraction, broken stick and log series) and the maximum value [ $\ln(s)$ ] to show the parameter range of the constant  $b$ .

The exponential SPAR can be rewritten as

$$S_i - S_{unit} = a \ln\left(\frac{a_i}{a_{unit}}\right) \quad (4)$$

with  $S_i$ ,  $S_l$ : species number at areas  $a_i$  and  $a_l$ .  $a$ : constant. Introducing formula 3 into 4, we get after simple rearrangement

$$e^{\frac{H_i}{b}} \approx a \ln(a_i) - a \ln(a_{unit}) + S_{unit}$$

and after setting  $a_{unit}$  to 1

$$a \approx \frac{1}{\ln(a_i)} e^{\frac{H_i}{b}} - \frac{S_{unit}}{\ln(a_i)} \quad (5)$$

This formula equals the factor of the empirical fit of Figure 4 and formula 1.  $S_l$  is of course the species number per unit area,  $a_i$  the total area sampled. Formula 5 gives also an interpretation of the constants in formula 1:  $\alpha$  equals the inverse of the natural logarithm of total area sampled,  $\beta$  is the inverse of  $b$ , the constant of the diversity–species number relation, and  $\gamma$  is  $S_{unit}/\ln(a_i)$ .

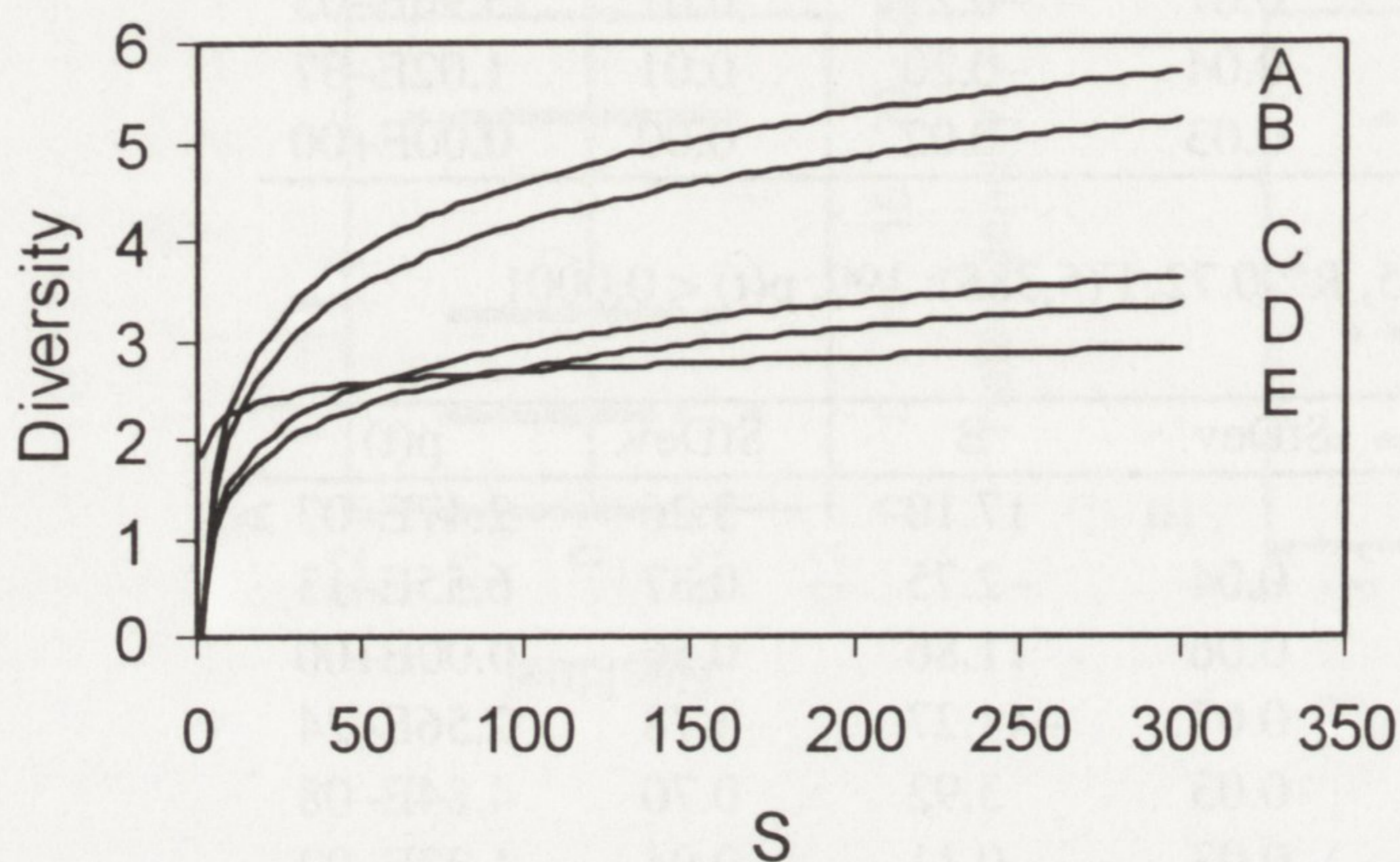


Fig. 5. Dependence of Shannon diversity on the number of species for four theoretical relative abundance distributions (B: broken stick, C: canonical log-normal, D: random fraction, E: log series with slope of  $-0.1$ ). A: plot of max. possible diversity [ $=\ln(S)$ ]. All curves are fitted by logarithmic functions (obtained from species numbers of 10, 25, 50, 75, 100, 200, 300) with variance explanations  $> 0.99$



Figure 6 shows that formula 5 is indeed a good descriptor of the relation between diversity and factor value. Plotted are predicted and real factor values of all assemblages with sequential adding sampling. Because formula 5 is the general solution independent of assemblage structure and all assemblages were taken with a mean value of the constant  $b$  the fit is not as good as the empirical fit in Figure 4B but sufficient to confirm the general applicability of formula 5.

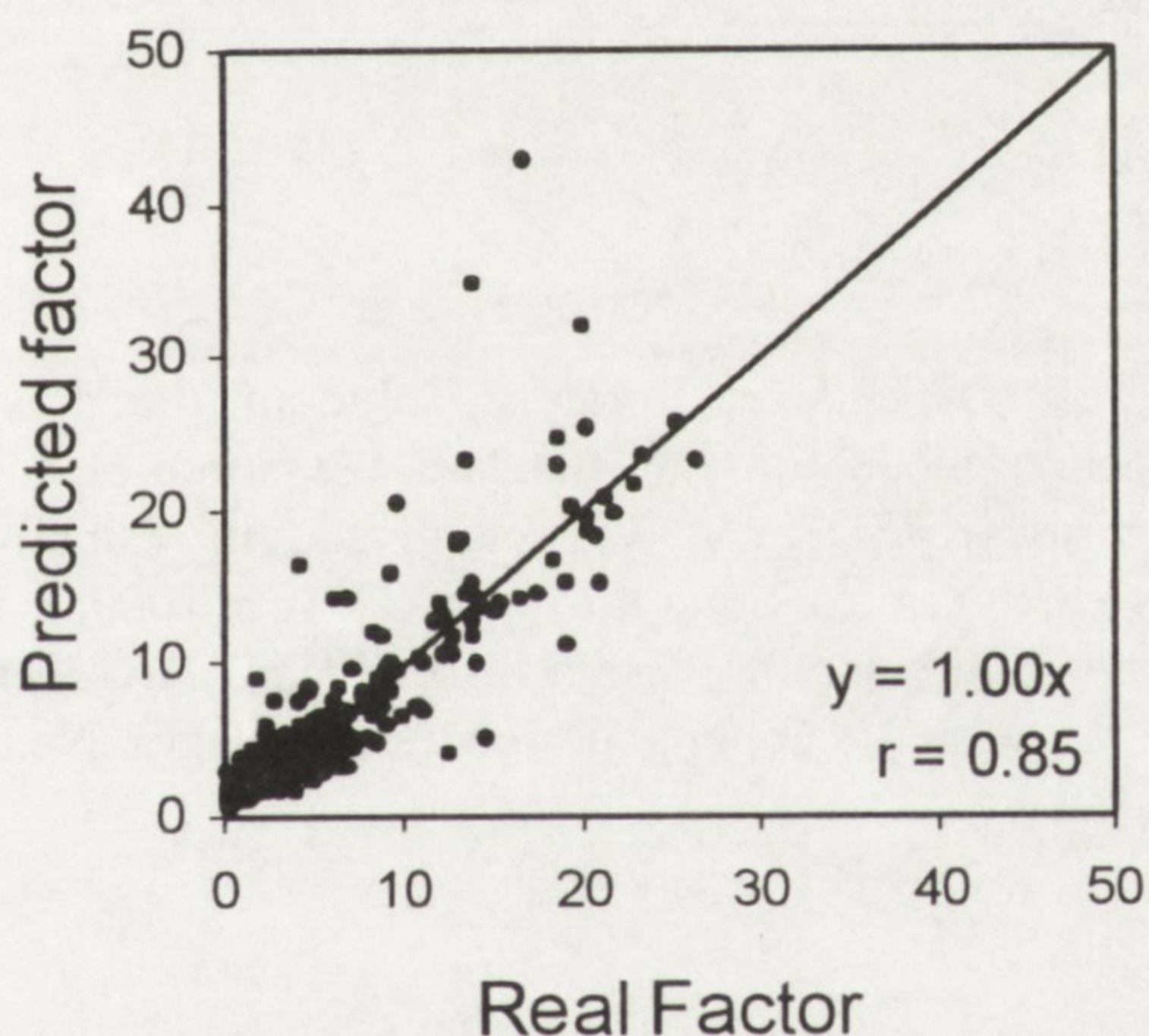


Fig. 6. Test of the model described by formula 5: predicted values and observed values of the factor of the exponential model for the assemblages No. 1 and 4 to 6 of Table 1 in Ulrich 2000a (sequential adding sampling). Given is also the regression function and Pearson's coefficient of correlation  $r$ . The constant  $b$  in formula 5 was set to 0.93

In the case of the slope of the power function model the multiple regression resulted in a more complicated picture. As expected DWD slope and evenness were negatively, diversity and heterogeneity positively correlated with the slope value  $z$ . The variance explanation, however, was relatively low (55%) and it was not possible to link the slope  $z$  directly with one of the above variables.

A plot of the  $z$  values against the SD values of the assemblages showed a more decisive pattern (Fig. 7). There was a bell shaped relation between SD and slope, independently of sampling method (Fig. 7B, C). SD

values below 2 and above 5 resulted in slope values of the SPARs below 0.2, SD values between 3 and 4 gave slopes between 0.6 and 1. This result can be compared with typical SD values of theoretical relative abundance distributions (see Fig. 9 of part I, Ulrich 2000a). SD values below 2 characterize the random assortment, the broken stick, the dominance decay, and the overlapping niche model. These models should all result in slope values below 0.3. High SD values above 6 have the geometric, the dominance preemption, and the log-series model. They also will have slopes below 0.3. Random fraction, log-normal and (dependent on the parameter settings) power fraction and Zipf-Mandelbrot distributions are characterized by SD values between 3 and 5 and it is concluded that their power function SPARs will frequently have slopes between 0.4 and 1.

Because of the dependence of the slope parameter on species density, the slope should also depend on the unit of area sampled. This was indeed the case (Fig. 8A). Higher units of areas reduced the slope value independently of sampling method. The factor of the exponential SPAR model, on the other hand, was not affected by the unit of area (Fig. 8B). The sample size did not significantly influence both parameters (Fig. 8C, D).

We can turn the relation between relative abundance distribution and slope value on the head. Although it is impossible to predict the slope value it is possible to predict the SD value from any given slope value (at fixed sampling unit) if we change the axes in Figure 6 (Fig. 9). For any slope value and any given unit of area two SD values exist which are unambiguously linked with them. The higher the slope is the more equal are these values and at slopes above 1 SD values of 3 are predicted. Both relations follow linear regressions:  $SD = -1.9 \text{ slope} + 5.5$  and  $SD = 1.5 \text{ slope} + 1.3$ . Changing the unit of area will result in a parallel shift of the regression lines. This is easily seen when rewriting the power

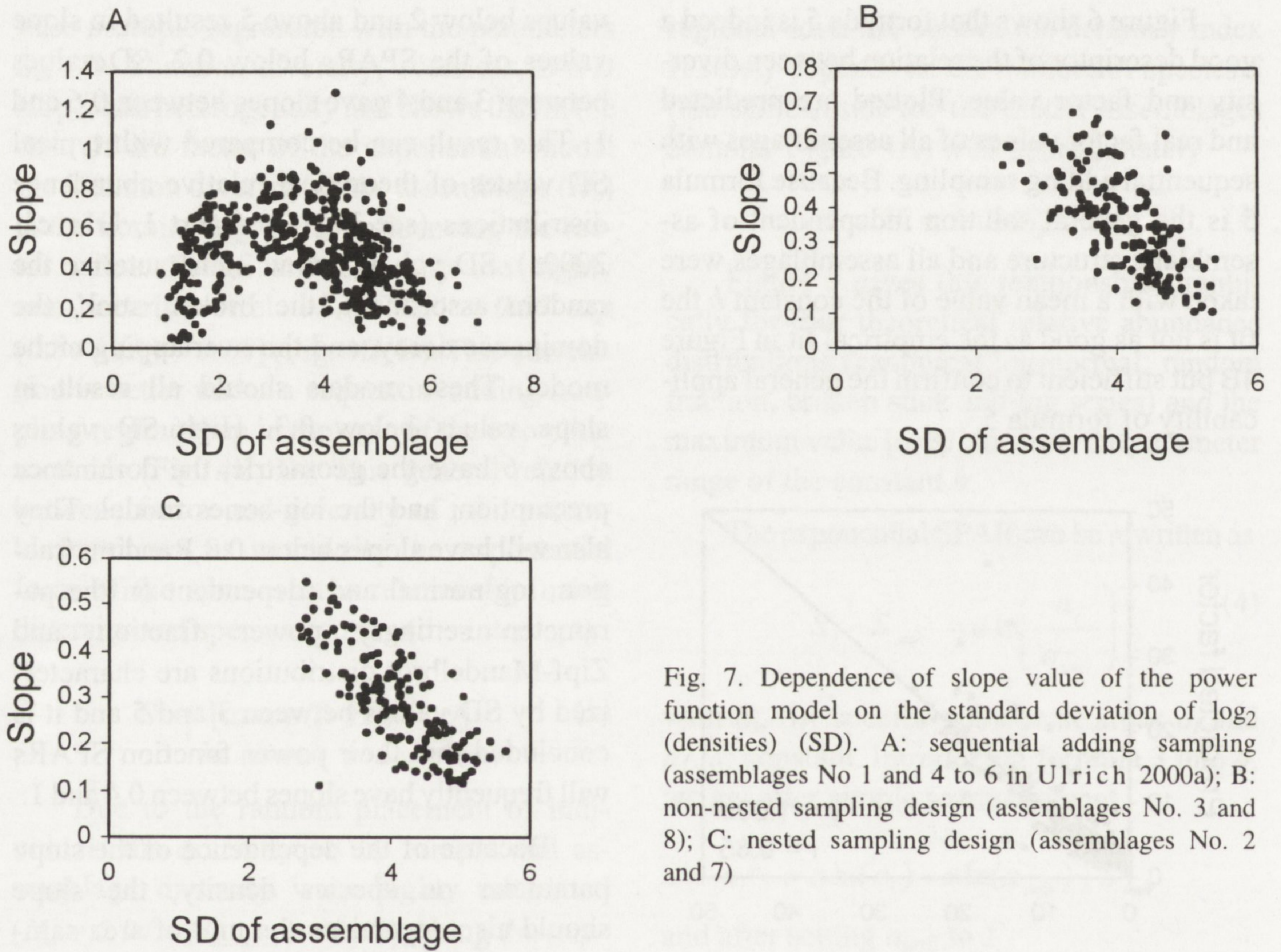


Fig. 7. Dependence of slope value of the power function model on the standard deviation of  $\log_2$  (densities) (SD). A: sequential adding sampling (assemblages No 1 and 4 to 6 in Ulrich 2000a); B: non-nested sampling design (assemblages No. 3 and 8); C: nested sampling design (assemblages No. 2 and 7)

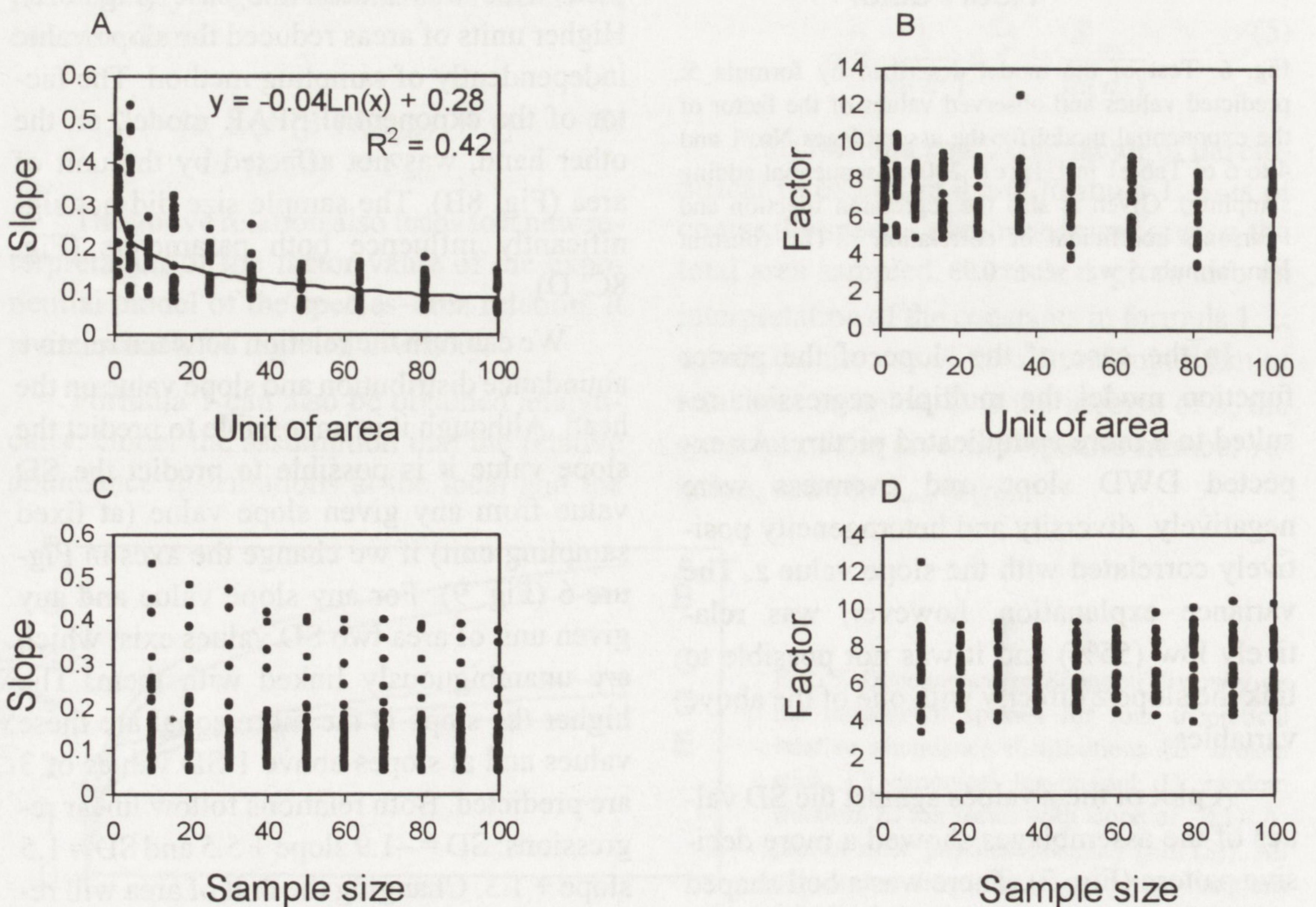


Fig. 8. Dependence of slope and factor values on unit of area and sample size. A, B: slope and factor in dependence of unit of area (number of cells); C, D: slope and factor in dependence of sample size (10 to 100). Givenis also the only significant ( $p(t) < 0.01$  regression,  $R^2$ : variance explanation

function SPAR and introducing into the above regression:

$$SD = cz + d = \frac{c}{\ln(n)} \ln \frac{S_{unit}}{S_n} + d \quad (6)$$

where  $n$  is the area sampled,  $z$  the slope and  $d$  the constant. Changing the unit of area changes  $S_{unit}$  by a certain factor  $c_1$ . Therefore

$$\begin{aligned} SD &= \frac{c}{\ln(n)} \ln \frac{c_1 S_{unit}}{S_n} + d = \\ &= \frac{c}{\ln(n)} \ln \frac{S_{unit}}{S_n} + \\ &\quad + \left( \frac{c \ln(c_1)}{\ln(n)} + d \right) \end{aligned} \quad (7)$$

Formula 7 is the same as formula 6 except of a different intercept and is the general solution of the relation between SD and slope of the power function SPAR. The factors  $c$  (−1.9 and 1.5) seem to be constants, the intercepts depend on assemblage structure, dispersion and on the habitat. They have to be determined empirically for each habitat.

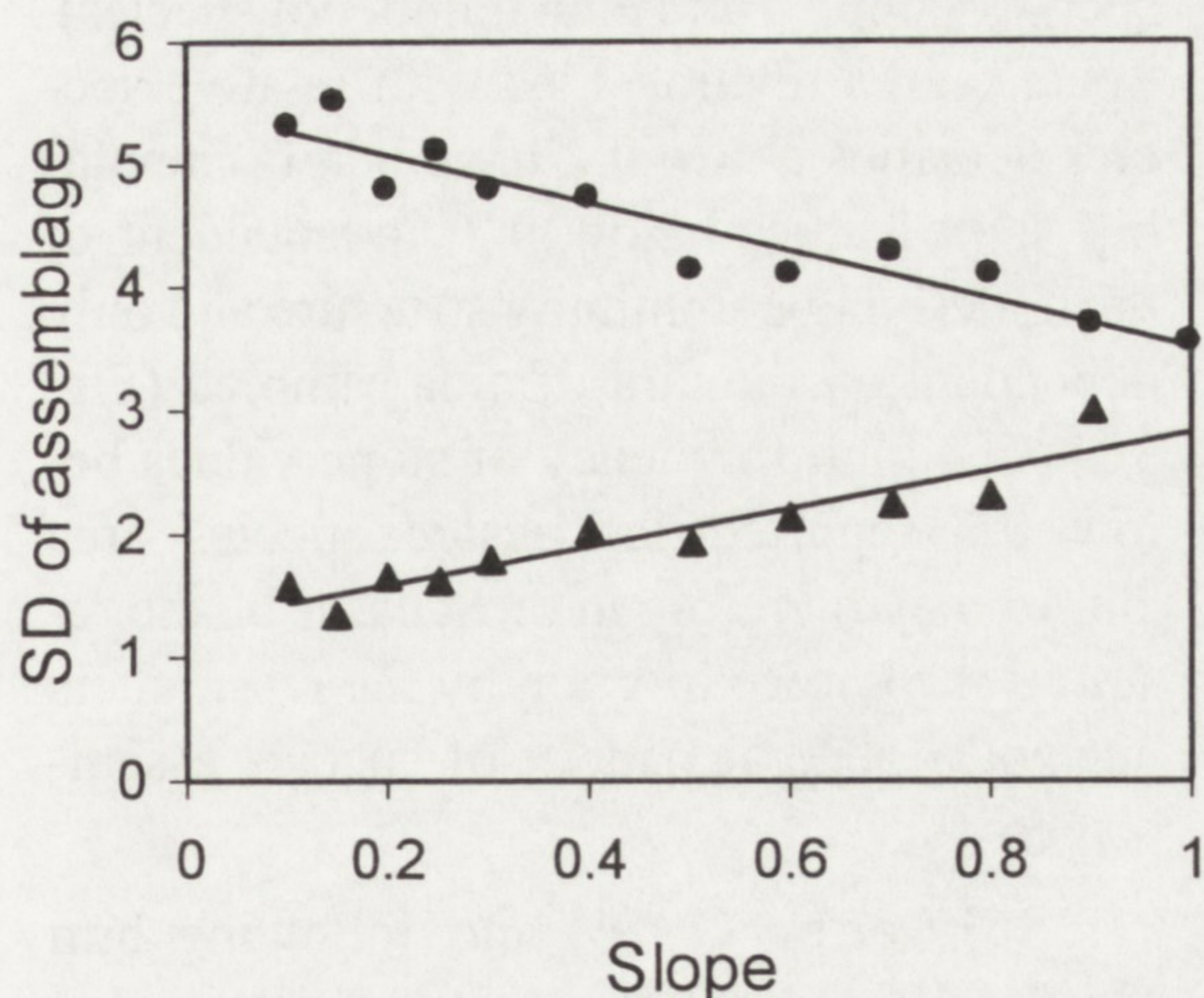


Fig. 9. Linkage between slope of the power function and standard deviation of  $\log_2$  (densities) (SD) at constant unit of area (1 cell = 0.0011% total area). Plotted are the associated SD values at given slope values. The SDs are means for slope values ( $x \pm 0.02$ ) of assemblages 1 and 4 to 6 of Table 1 of Ulrich (2000b). The standard deviations of the SD-means are in the upper regression always less than 10% mean in the lower regression less than 36% mean ( $n = 6$  to 32). Upper regression:  $SD = -1.9 \text{ slope} + 5.5$ ; lower regression:  $SD = 1.5 \text{ slope} + 1.3$ . Both regression are significant at  $P(t) < 0.01$

## 4. DISCUSSION

The main goal of the present study was to link the slope and the factor of the power function and exponential model of species–area relationships with other variables. A first point, however, which deserves attention is the correlation of model fit and slope value in the power function model. Because  $R^2$  values below 0.9 can frequently hardly be called good fits, low slope values may be the result of the inappropriateness of the power function model. The huge compilation of SPARs in Connor and McCoy allow to test whether the correlation between  $R^2$  and slope also occurs in real SPARs. Figure 10 shows that this is indeed the case. The regression of Figures 1A and 10 are very similar pointing to the interpretation that low slope values are connected with a worse fit of the power function model. This seems especially to be the case at slope values below 0.1.

The study showed that the slope of the power function model was more related to relative measures and that the factor was more influenced by absolute measures. Additionally, the slope of the power function turned out to be more effected by the sampling procedure (especially the unit of area) than the factor of the exponential model. The slope was also dependent on sampling procedure. Sequential adding sampling resulted in significantly higher slopes than a nested or a

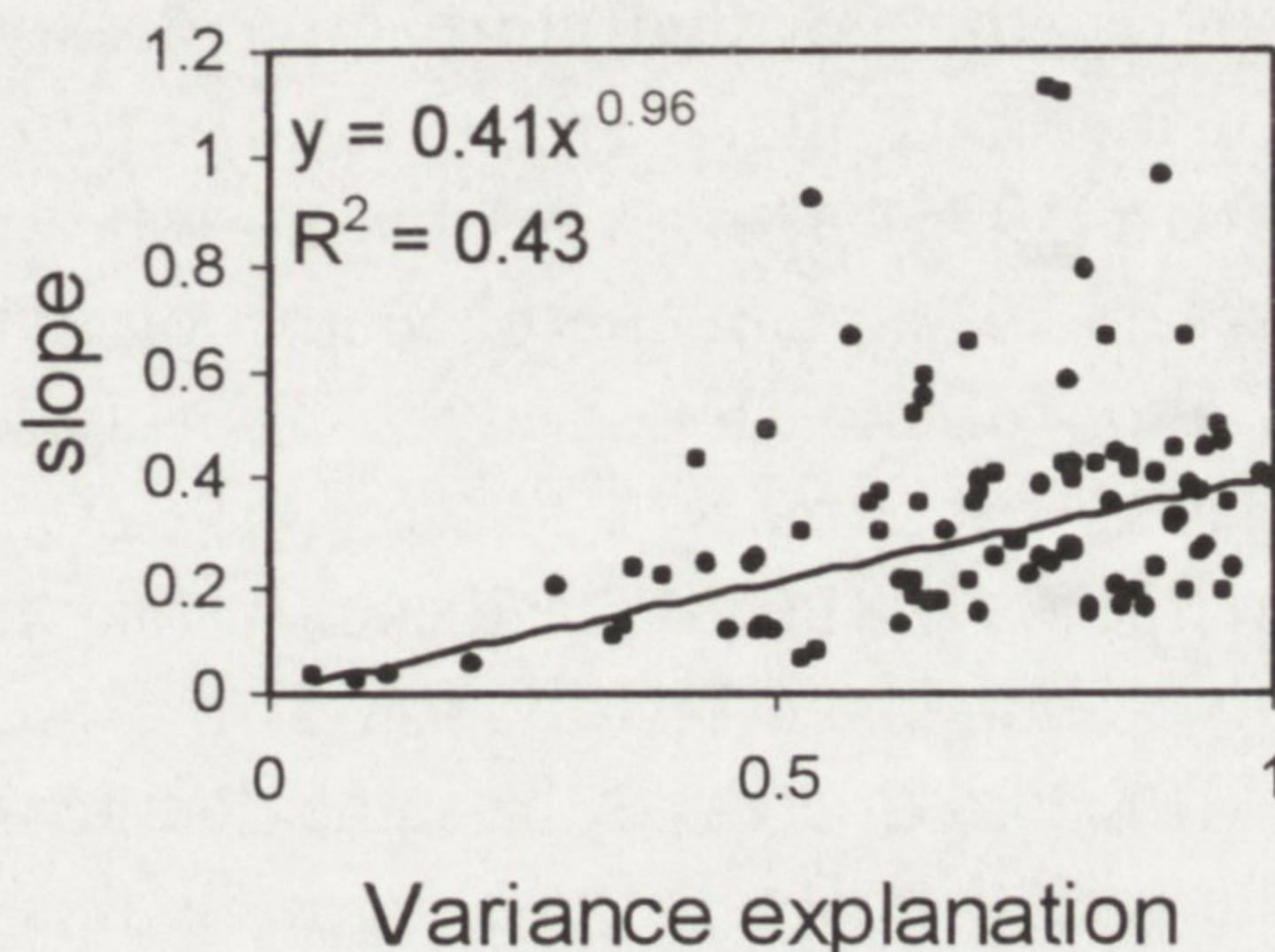


Fig. 10. Dependence of slope value on variance explanation for 100 species–area relationships given in Connor and McCoy (1979).  $R^2$ : variance explanation,  $P(t) < 0.01$

non-nested design. Low slope values for nested species area relationships had previously been reported by Rosenzweig (1995) who stated that such curves have frequently slopes below 0.2. The mean values reported in Table 1 are larger and range around 0.27. This may however stem from the selection of assemblage structures.

Slope values have long been thought to range around 0.25 (Preston 1962, May 1975) if the underlying community structure is log-normal. In a recent study, however, Leitner and Rosenzweig (1997) showed that this is not the case when dealing with nested species area curves. From their model assemblages they deduced slope values around 0.77 and were urged to add the assumption of a positive correlation of range size and abundance to the log-normal model to derive slope values below 0.3. The data presented here confirm and extend these results. High slope values generally characterized assemblages with SD values between 2 and 4, independently of sampling method. The finding of high slope values in the mid range of SD values is therefore not restricted to nested SPARs. The range between 3 and 4 is the range for log-normal, but also for power fraction, or random fraction distributions (see Fig. 9, Ulrich 2000a) and the mean slope value of my assemblages in this range was 0.61 which is in good agreement with the findings of Leitner and Rosenzweig. These authors introduced an additional assumption, the positive correlation between abundance and range size (Hanski 1982, Gaston 1996). However, this may be not necessary to explain the phenomenon: The derivation of the slope values from a log-normal distribution by Preston (1962) and Sugihara (1981) assumed tacitly that the relative abundance distribution at the local scale (the unit of area) is the same than in the whole area. For small sampling units this is not necessarily the case (Ulrich in prep.) and the strong logarithmic dependence of slope on unit of area (Fig. 8) indicates that the unit of sampling may be more fundamental than

abundance range size distributions. At units of area above 16 cells the slopes were nearly always inside the expected range. This indicates that the unit of area has to be large enough to ensure that local and regional assemblage structure are comparable. Additionally, in my model assemblages the dependence of slope on unit of area was more strong for assemblages with SD values between 2 and 4 than for assemblages with SD values above 4 (in the latter case units of 9 cells or more produced always slopes below 0.25).

Again, a comparison with the data in Connor and McCoy helps to explain the pattern. Of the 100 SPARs they compiled 28 contain data on total species numbers. In only two of these cases the relative species density was below 20%. This can be compared with the relation of relative species density and slope (see Part III of this paper, Ulrich 2000b) which is nearly the same than the relation between  $S_{unit}/S_s$  and slope (of course due to the high correlation between number of species sampled and total number of species). Species area relationships with relative species densities above 0.2 have nearly inevitably slope values below 0.2, independent of the underlying community structure and only dependent on the unit of area sampled (Fig. 3C). The high frequency of slope values below 0.2 reported for nested species area curves may therefore in most cases be the result of the choice of grain by the researchers but not a general pattern of species assemblages.

Schoener (1976) and following him Martin (1981) predicted that the slope value will vary between 0 and 0.5, based on the model  $z = 1 - [1/(2 - S/P)]$  where  $S$  is the number of species found and  $P$  the source size pool. In the case of our model assemblages  $S/P$  equals the proportion of species sampled  $S_s/S_a$ . A comparison of the values from this model with the 'real' values (of the model assemblages) proofed that Schoener's model is indeed able to predict real slopes under the special case of very equal distributed

assemblages (SD values below 2.5 according to a broken stick, a random assortment, or a dominance decay model) (regression between predicted and real  $z$  values:  $z_{real} = 0.63 z_{pred} - 0.09$ ;  $r = 0.88$ ). In general, however, the model failed. From the model assemblages studied here it was also impossible to derive a good relation between the proportion of species sampled and  $z$ .

The only author's relating the slope value and SD as a variable describing assemblage structure were Hanski and Gyllenberg (1997). Based on metapopulation dynamics they united species–area curves and distribution–abundance curves into a general theory and predicted a sigmoidal dependence of slope upon SD with decreasing  $z$  when SD increases. Even when leaving the problem aside that their theory does not discriminate between different sampling methods and subsumes the various aspects of grain and extent into a single parameter the plot of SD against slope in Figure 7 shows a more complicated bell shaped but not sigmoidal pattern. A fit of the predicted slopes with the real ones (assemblages 1 to 8 of Table 1) resulted in every case in only weak (although at  $p(t) < 0.05$  significant) correlations ( $r$  below 0.2) with nearly all predicted values ranging either much too low or much too high.

If  $z$  is not an ecological constant the question of an ecological interpretation of the slope parameter remains, or more precisely of the connection between  $z$  and other parameters describing assemblages and environments. The high dependence of  $z$  on sampling method, grain and extent indicates that this seems possible only to a limited degree (under standardized sampling methods). Then the slope can be linked with the fraction of rare species in a community (Fig. 2) and the relative abundance distribution. The bell shaped curves of Figure 7, however, tell that this connection is only one-sided. It will not be possible to predict assemblage structure from the slope value alone. The crucial factor determining the slope proved to be the grain and because in most studies this variable dif-

fers considerably it seems nearly impossible to compare slope values of different assemblages and habitats. This casts also doubts on some reported globally differences in slope value, especially the latitude dependence, and differences between mainland and island slopes (Rosenzweig 1995).

The factor of the exponential SPAR model has largely been ignored by ecologists but the results presented in Figure 4 and the derivation of formula 5 show that the factor deserves much more attention. The factor is closely linked with the Shannon measure of diversity and may be used as an estimate of regional diversity. The factor has therefore a much more founded ecological interpretation than the slope value of the power function. Of course, the relation between factor of the exponential function and Shannon diversity is not the same for all types of communities and in some even no such relation is expected. Imagine, for instance, an assemblage with maximum diversity =  $\ln(S)$  and high densities of the species. In this case, nearly all species will already be found in one unit of sample area and the slope of the logarithmic fit will be very low. According to formula 5 this will also result in a much too low estimate of Shannon diversity. However, at the same time, the fit of the logarithmic model will also be worse. This relation between model fit and applicability of formula 5 can be seen in Fig. 11. Plotted are the quotient of predicted and real Shannon diversity and goodness of fit of the exponential function ( $R^2$ ). If  $R^2$  exceeds 0.95 71.6% of all the estimates range between 75% and 125% of the real diversities, a very good agreement between prediction and real value. If the  $R^2$  is less than 0.90 this is the case in only 34.7%.

Of course, the above results and especially the new interpretation of the factor of the exponential SPAR model have to be checked using natural communities. To do this the species–area curve, species numbers, the local and regional diversity, and the community structure have to be reported simultaneously. Such compilations are up to now

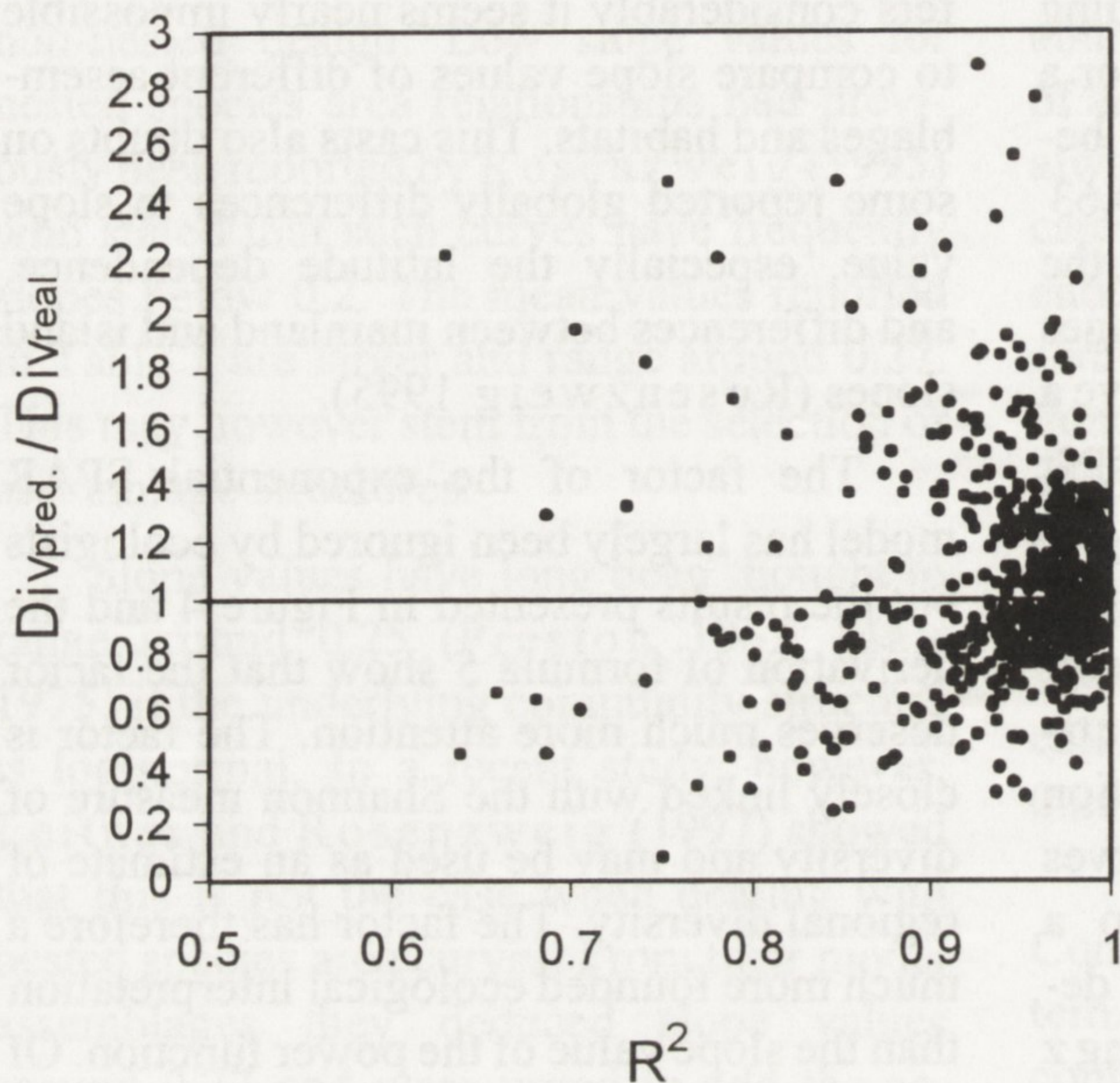


Fig. 11. Goodness of prediction of Shannon diversity in relation to fit of the exponential model of SPAR. The goodness is given as the quotient of predicted and real diversity on all 768 model assemblages (assemblages No. 1 to 10 in Table 1 of Ulrich 2000a).  $R^2$ : variance explanation of the exponential model

missing but it would be a valuable task and a further step to uncover the secrets of species–area relationships.

**ACKNOWLEDGMENTS:** I thank Prof. J. Buszko and Dr. Kartanas for critical and valuable suggestions on the manuscript. Miss H. Pearson kindly improved my English.

## 5. SUMMARY

Using model assemblages generated by a FORTRAN program the parameter values (slope and factor) of the power function and the exponential model of species–area relationships have been studied. It appeared that the slope value is not a constant independent of area and sampling method but depends strongly on grain (Fig. 8), sampling method (Tables 1 and 3), and model fit (Figs 1 and 10). The relationship between slope and sampling parameters and community structure was studied by multiple regression (Tables 2 and 4) and residual analysis (Fig. 3). Of the ecological parameters the slope of the density–weight relationship and the fraction of singletons in the sample (Fig. 2) proved to be of major importance.

A plot of slope against assemblage structure (estimated by the standard deviation of  $\log_2$  (densities) was bell shaped (Fig. 7) with the highest slope values at intermediate SD values (2 to 4) (Fig. 9). A comparison of this plot with SD values from theoretical relative abundance distributions (see Fig. 9 in Ulrich 2000a) showed that log-normal distributed

assemblages should have slope values that are higher than previously reported in the literature.

The factor of the exponential model was more independent of sampling methods but connected with sampling efficacy. A high non-linear correlation between factor and Shannon diversity was detected (Fig. 4) and a general function developed and tested (Figs 5, 6 and 11). The factor of the exponential species–area relation may serve as an estimate of regional diversity.

## 6. REFERENCES

- Arrhenius O. 1923 – Statistical investigation in the constitution of plant associations – *Ecology* 4: 68–73.
- Connor E. F., McCoy E. D. 1979 – The statistics and biology of the species –area relationship – *Am. Nat.* 113: 791–833.
- Begon M., Harper J. L., Townsend C. R. 1986 – *Ecology* – Oxford, pp. 1–876.
- Gaston K. J. 1996 – The multiple forms of the interspecific abundance – distribution relationship – *Oikos*, 76: 211–220.
- Hanski I. 1982 – Dynamics of regional distribution: the core and satellite hypothesis – *Oikos*, 38: 210–221.
- Hanski I., Gyllenberg M. 1997 – Uniting two general patterns in the distribution of species – *Science*, 275: 397–400.
- Harte J., Kinzig A. P. 1997 – On the implications of species – area relationships for endemism, spatial turnover, and food web patterns – *Oikos*, 80: 417–427.

- Harte J, Kinzig A. P., Green J. 1999 – Self-similarity in the distribution and abundance of species – *Science*, 284: 334–336.
- He F., Legendre P. 1996 – On species area relations – *Am. Nat.* 148: 719–737.
- Leitner W. A., Rosenzweig M. L. 1997 – Nested species-area curves and stochastic sampling: a new theory – *Oikos*, 79: 503–512.
- MacArthur R. H. 1965 – Patterns of species diversity – *Biol. Rev.* 40: 510–533.
- MacArthur R. H., Wilson E. O. 1967 – *The Theory of Island Biogeography* – Princeton, Univ. Press.
- Martin T. E. 1981 – Species–area slopes and coefficients: a caution on their interpretation – *Am. Nat.* 118: 823–837.
- May R. M. 1975 – Patterns of species abundance and diversity – (In: *Ecology and evolution of communities* Eds. M. L. Cody, J. M. Diamond) – Belknap (Cambridge), pp. 81–120.
- Ney-Nifle M, Mangel M. 1999 – Species-area curves based on geographic range and occupancy – *J. Theor. Biol.* 196: 327–342.
- Palmer M. W., White P. S. 1994 – Scale dependence and the species – area relationship – *Am. Nat.* 144: 717–740.
- Preston F. W. 1962 – The canonical distribution of commonness and rarity. Part I and II – *Ecology*, 43: 185–215, 410–432.
- Rosenzweig M. L. 1995 – *Species diversity in space and time* – Cambridge, Univ. Press.
- Rosenzweig M. L. 1999 – Heeding the warning in biodiversity's basic law – *Science*, 284: 276–277.
- Rosenzweig M. L., Sandlin E. A. 1997 – Species diversity and latitude: listening to area's signal – *Oikos*, 80: 172–176.
- Schoener T. W. 1976 – The species-area relation within archipelagos: models and evidence from island land birds (In: *Proc. 16<sup>th</sup> Int. Ornith. Congr.* 1976 Eds. H. J. Frith, J. H. Calaby) – Canberra, pp. 629–642.
- Strong D. R. 1974 – Nonasymptotic species richness models and the insects of British trees – *Proc. Nat. Acad. Sci. USA* 71: 2766–2769.
- Sugihara G. 1981 –  $S = CA^z$ ,  $z \approx 1/4$ ; a reply to Connor and McCoy – *Am. Nat.* 117: 790–793.
- Ulrich W. 1999 – The numbers of species of Hymenoptera in Europe and assessment of the total number of Hymenoptera in the world – *Pol. J. Entomol.* 68: 151–164.
- Ulrich W. 2000a – On species–area relationships I: fit of power function and exponential model – *Pol. J. Ecol.* 48: 3–20
- Ulrich W. 2000b – On species–area relationships III: the intercept – *Pol. J. Ecol.* 48: 37–48
- Whitehead D. R., Jones C. E. 1969 – Small islands and the equilibrium theory of insular biogeography – *Evolution*, 23: 171–179.
- Wilson E. O., Taylor R. W. 1967 – An estimate of the potential evolutionary increase in species density in the Polynesian ant fauna – *Evolution*, 21: 1–10.
- Wissel Ch., Maier B. 1992 – A stochastic model for the species-area relationship – *J. Biogeogr.* 19: 355–362.

(Received after revising October 1999)