VARIABILITY OF SPATIAL DEPENDENCE WITHIN A PIONEER PLANT COMMUNITY

A. Horváth and L. Makrai

Abstract. Since it is important to know the degree of variability of spatial dependence among plant populations when comparing two communities, but only a very few literature data exist on it, the first aim of this study is to give some compensation for this deficiency. We applied information statistical models to estimate vegetation pattern organization in an early stage of primary succession, and computed associatum values and their variances along the spatial scale. The null model was the random shift, the validation procedure was also executed for it. We can summarize, that the sampling area with 50 m^2 was sufficiently large to gain interpretable associatum curves. The area of maximum spatial organization was 15 cm $\times$ 15 cm, which significantly differs from null model. When we divided the whole sampling area to 10 m^2 stripes, the associatum difference curves of every parcels had significant interval, and average of them was similar to the curve from the whole area. Therefore, a 10 m^2 area could represent the unit of entire community. When the area was decreased to 2.5 m^2, the variability of spatial dependence increased, and the appearance of significant associatum values became more unpredictable. At this size of area the unity of patterns organization seems to be collapsed, and the samples can not represent effectively the entire community. Six types of associatum measures were compared regarding to their variability in the spatial series steps.

Keywords: information theory models, syncretic functions, spatial dependence, spatial scaling, primary succession, randomization tests, validation of null models

A. Horváth, Department of Ecology, Szeged University, H-6701 Szeged, Pf. 51, Hungary
L. Makrai, Department of Botany, Szeged University, H-6701 Szeged, Pf. 657, Hungary

Introduction

Since the beginning of nineties, after the methodological foundation by Juhász-Nagy (1972, 1976, 1980, 1984), the application of the information statistical models started to get into the practice of the spatial pattern analysis. In addition to solving methodical problems (e.g. computation: Bartha *et al.* 1994, Erdei and Tóthmérész 1993, Podani 1993, Horváth 1998), some case studies were also carried out. These experiments produced some evidences about the pattern transformation during primary succession (cf. Bartha 1990, 1992, Margóczi 1995, Horváth 1997), or, e.g., spatial dependence of populations in different types of loess steppe (Hochstrasser 1995), dolomite grassland communities (Szollát and Bartha 1991, Bartha *et al.* 1998) or prairie vegetation (Bartha *et al.* 1995).

In these studies two or more plant communities were compared on the basis of main syncretic functions, characteristic scaling, or plexus graphs of pairwise associations. Generally, each community types (e.g. successional stage) were represented by only one sample. Though the field sampling procedures were very extensive in all cases, and the samples consisted of several hundreds or thousands of primary plots (microquadrats), the statistical evaluation of comparison between communities was impossible because of missing repetition of samples. For correct separation of different stands it is unavoidable to know the variation of adequate coenological characteristics.

There is only a few field experiment that yielded some facts about the variability of syncretic functions of a given community. Hochstrasser (1995) studied seven types of loess grasslands, four of which were

examined in *two different (consecutive) years*. She observed relatively large change between two years in maximum associatum and relative associatum (both from field and random difference values), which may be originated only partly from secondary successional dynamics. The temporal shift of associatum values makes the unambiguous distinction of community types impossible. So, in her study, the temporal dynamics (trend, periodicity or stochasticity) of spatial dependence could be greater than its difference between coenotaxa, whereas their floristic and physiognomic diversity was significant.

Bartha *et al.* (1998) compared two dry dolomite grassland types existing on contrasting slopes. One of them was represented by four, the other one by five samples collected from different dolomite hills. Mann-Whitney method was applied for testing the null hypothesis that the aspect had any effect on some basic syncretic functions. The results showed that significant differences between two grassland types could be detected when they were restricted to assemblages of only dominant species. Otherwise, analysing the entire communities, the variance of structural characteristics was quite large. In this study, however, the sample repetitions for a community types were collected from *different stands*, from different floristic situations, that could be one component of the great variability.

We can find some data about variance of structural dependence relating to *one stand* of one community only in the work of Bartha *et al.* (1995). The subject of this study was to compare three prairie communities with two samples for each. It has been pointed out that there was no clear difference between two stands regarding the relative associatum curves, only the third grassland type was separable from the others. The correlation between total transect lengths (each sample was divided to two part, and they were fused, respectively) and maximum relative associatum was not significant.

Some general, partly methical questions or problems emerge from the results cited above:

(1) What is the size of a *sufficiently large sample area* (grid or transect) that represents the entire (and more or less homomorfous) stand of a given community for comparison to other ones? The question is related to the problem of variability of vegetation pattern along spatial scale. In this effect, the scaling problem is independent of *characteristic scaling*, because now the aim is just to find the smallest part of a vegetation stand that gives constant characteristic values (if they are), independently its location in the stand.

(2) How large is the variance of spatial dependence in a community when its pattern is divided into some parts? How does this variance depend on the size of the partial patterns?

(3) Can the average of associatum functions of different parts approach the spatial dependence in entire community?

(4) Is there any significant difference between average (or median) of values coming from parcels of field sampling area and randomized patterns?

We will focus on the last three problems. We suppose that our relatively extensive field sample represents the studied community, therefore the main aim will be to get some evidences about the variability of mutual spatial dependences. We compare the variance of field values with random patterns even when randomizations are related not exactly to same parcel, but the other part of the given stand. We use the random shift null model to make random patterns, and, for correct statistical evaluation, we give the validation of the test.

Materials and methods

Study site

The study site was located on the plateau of a spoil bank consisting of sterile clayey bedrock, which originated from an opencast coal-mine at Hepworth (near Huddersfield), West Yorkshire, England. The field sampling was carried out in first half of May, 1993. Since the stand was only several years old, the development of vegetation could reach the stage of a pioneer community, so it could be considered as an early stage of primary successional process. The apparently primitive grassland had no complex physiognomical structure, it consisted of only two simple layers. The lower one was composed of mosses and lichens, while the upper one could be specified as the composition of grasses and some dicotyledons.

The population number was very low, there were only ten taxa in the studied area. Only two species (*Pogonatum urnigerum* and *Agrostis capillaris*) reached more than 10 % of frequency regarding the whole sampling area, otherwise they were present at least every third cell of the grid (Table 1). Only the half of the species were vascular plants, three of which were grasses. Dominant grass was *Agrostis capillaris*, subdominant was *Deschampsia flexuosa*, which have circumpolar area type, and they are common in England. Both species exhibited clumped pattern (Fig. 1); the tussocks of *Deschampsia* are smaller and more segregated.

Table 1. Taxon list of the studied community. The frequency values are calculated from whole sampling area of 20000 plots.

Code	Taxon name	Frequency (%)
Pu	*Pogonatum urnigerum* (moss)	39.68
Ac	*Agrostis capillaris*	35.26
Gp	*Grimmia pulvinata* (moss)	7.24
Df	*Deschampsia flexuosa*	5.50
Dc	*Distichium capillaceum* (moss)	5.25
Ra	*Rumex acetosella*	4.06
Cp	*Ceratodon purpurascens* (moss)	3.26
Ep	*Epilobium* sp.	1.59
Li	*Lichenes*	0.86
Hl	*Holcus lanatus*	0.72

Table 2. The properties of spatial series steps regarding to analysis of whole sampling area, stripes and blocks.

Step #	Number of fused primary plots	Area of secondary plots (m^2)	Number of secondary plots		
			in whole area	in stripes	in blocks
1	1 × 1	0.0025	20000	4000	1000
2	2 × 1	0.0050	19900	3980	980
3	2 × 2	0.0100	19701	3781	931
4	3 × 2	0.0150	19602	3762	912
5	3 × 3	0.0225	19404	3564	864
6	4 × 4	0.0400	19109	3349	799
7	5 × 5	0.0625	18816	3136	736
8	6 × 6	0.0900	18525	2925	675
9	7 × 7	0.1225	18236	2716	616
10	8 × 6	0.1600	17949	2509	559
11	9 × 9	0.2025	17664	2304	504
12	10 × 10	0.2500	17381	2101	451
13	12 × 12	0.3600	16821	1701	351
14	14 × 14	0.4900	16269	1309	259
15	16 × 16	0.6400	15725	925	175
16	18 × 18	0.8100	15189	549	99
17	20 × 20	1.0000	14661	181	31
18	24 × 24	1.4400	13629	–	–
19	28 × 28	1.9600	12629	–	–
20	34 × 34	2.8900	11189	–	–
21	40 × 40	4.0000	9821	–	–
22	50 × 50	6.2500	7701	–	–
23	60 × 60	9.0000	5781	–	–
24	70 × 70	12.2500	4061	–	–
25	80 × 80	16.0000	2541	–	–

Primary and secondary sampling

The sampling area was a 10 m × 5 m grid (50 m^2) with 5 cm × 5 cm microquadrats (primary plots), total number of *primary plots* was 20000 (200 × 100). We recorded the presence or absence data of each species in all cells. For studying the variability of pattern characteristics at relatively large area, we established five uniform, non-overlapping *stripes* on the grid, which are marked with upper case letters (A-E). Then we divided each stripe into four *blocks* (signed lower case letters, a-d, cf. Fig. 1), which can still provide sufficiently large sample size. The area of each stripe is 10 m^2 (10 m × 1 m), while the blocks are 2.5 m^2 (2.5 m × 1 m).
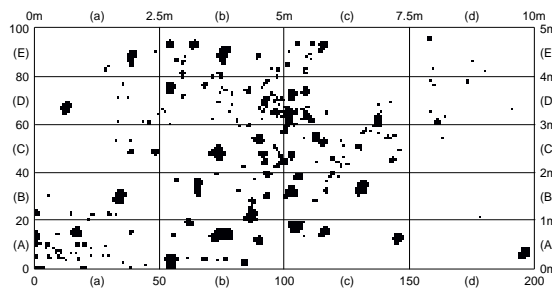

Fig. 1. The pattern of *Deschampsia flexuosa*. The numbers of cells on the grid are indicated at bottom and left, while the distance in meters at top and right. The letters in parentheses mark the stripes of sampling area (A-E) or the blocks of stripes (a-d).

For spatial scaling we performed *systematic secondary sampling* in several *spatial series steps* (Table 2). Given number of contiguous cells of the grid was fused, which formed (except second and fourth step) isodiametric secondary sampling units (cf. Table 2). The number of secondary plots (sample size) was the greatest possible at each steps, so we realized a *complete sampling*.

Random references

To compare the results with null hypothesis – to generate random patterns – we applied the *random shift* methods (Palmer and van der Maarel 1995), and in one case the *complete randomization* (i.e. *independent assignment* in Roxburgh and Matsuki 1999). Using random shift we can analyse the degree of spatial dependence between populations in such a manner that we randomize only the interspecific patterns (the autocorrelation of each population pattern is kept more or less constant), and the results are free from textural constraints (cf. Bartha and Kertész 1998). It is reasonable that in the case of non-stationary patterns (e.g. if the pattern consists of several large patches), the random shift modifies the value of autocorrelation (Palmer and van der Maarel 1995, Roxburgh and Chesson 1998), regarding to our relatively large sample area, however, this problem does not arise (cf. Fig. 1).

Executing Monte-Carlo simulation the number of randomizations was in general 99, but with smaller data sets it was 999. In the statistical evaluation we performed one-sided test using significance level as $p = (n - ND + 1) / (n + 1)$, where n is the number of randomization, ND is the number of positive or negative (the higher) differences between field and random values (cf. Manly 1997).

When we calculate significance level, we have no evidences about the real values of probability of type I error. It can be determined however, if we compare random patterns to other ones originated from same methods, and we calculate probability that a characteristic value of a random pattern significantly differs from others (Roxburgh and Matsuki 1999). In ideal cases, the significance level and probability of making a type I error is same value, but we have to evaluate *validation procedure* to settle whether it is true or not in a given situation (using a given randomization test, and calculating a given characteristic value). In the first part of Results we will give the detailed description of the validation test method and its results for using random shift null model on a grid, and analysing spatial dependence of species patterns.

Measure of spatial dependence and its variability

Using information statistical models, we were interested in the degree of mutual dependence of all populations within the entire community (Juhász-Nagy 1976, 1984). To estimate it, we calculated *associatum* for field patterns (*fAss*) and the *associatum difference* values (*dAss*) along the spatial scale. We can get the *dAss* values if we subtract *aAss* from *fAss*,

$$dAss = fAss - aAss,$$

where *aAss* is the *average of associatum* values regarding random references.

In comparison of different communities with each other, it is worthy to consider the application of *relative associatum* that is defined by the ratio of *associatum* and *florula diversity* from field data:

$$relAss = fAss / fFD$$

(Bartha *et al.* 1995, 1998). This measure can be derived from the so called *simple relative associatum* introduced by Juhász-Nagy (1984: p. 371), as a redundancy estimate given by $1 - fAss / fFD$. In the cited work (p. 385), Juhász-Nagy obtained another formula to calculate redundancy-like relative associatum, better than the previous one. It is signed as *redAss* here, and given by:

$$redAss = 1 - fAss / locEv,$$

where *locEv* is the *local evenness* (see Juhász-Nagy, 1984 for more details).

We will introduce two other relative associatum measures, in relation to null model. Let the first one called *calibrated associatum* (*calAss*). It is „calibrated" between its actually possible maximum and minimum values, which are the *local distinctiveness* from field (*fLD*) and the *average associatum* from random references (*aAss*). The formula is defined as:

$$calAss = (fAss - aAss) / (fLD - aAss),$$

or more simple $calAss = dAss / (fLD - aAss)$. The other measure is the „*sensitive associatum*" (*senAss*), which is more responsive to spatial dependence, because it is defined by the ratio of *associatum difference* (*dAss*) and *dissociatum* from field data (*fDiss*):

$$senAss = dAss / fDiss.$$

The sensitivity can be seen easily, if we consider that *associatum* and *dissociatum* are complementary measures.

For analysing spatial patterns (calculating information statistic functions, performing spatial scaling and making random references) we applied the *INFOTHEM* program (Horváth 1998). All the measures are *standardized by the number of secondary plots* (listed in Table 2).

For estimating the variability we used the *variation coefficient*, which was calculated as a ratio of the standard deviation and the mean of the data. In comparison of associatum values of field and random patterns, since the normality of our data can be rejected only in several cases (*Kolmogorov–Smirnov* and *Shapiro–Wilks' W test*), generally the *two sample t-test* can be applied. The results of both *t-test* and *Mann–Whitney U test* were similar, so we will present only the significance values of the latter one, listed along the spatial scale.

Results

Validating random shift null model

Though Palmer and van der Maarel (1995) referred to the validation process of random models applied to analyse spatial dependence, and Roxburgh and Chesson (1998) also showed the results of this method, the most detailed and correct description can be found in the paper of Roxburgh and Matsuki (1999). All three works applied validation test for *pairwise association*, so it was necessary to extend it to *associatum difference*, as well.

Regarding *random shift* procedure on a grid, we can find some facts about validation in only two works. Palmer and van der Maarel (1995) pointed out that at $p = 0.05$ the probability of type I error is exactly 5 %, so the null model can be validated. Afterwards, Roxburgh and Matsuki (1999) demonstrated, that when spatial autocorrelation

increases, the chance of making type I error also increases. It means that random shift model can be validated, if the degree of autocorrelation is low (Morans $I < 0.4$). We can not fully accept this result, because they confused two different aspects of randomization test. One of them is the question of keeping autocorrelation value constant. From this point of view, the random shift is indeed not a perfect model in case of non-stationary patterns (as it was noted above). However, the question of reliability is the other aspect. It depends on the realization of randomization process and the calculation of the significance level. Thus, if the test produces real probability of type I error, it can be validated even when its applicability is questionable (because of e.g. too few and large patches). For this purpose it is clear that to make *pseudo-observed* patterns (comparing it to random references), they should have to use also the random shift model, which was applied for randomization test, but they did it in another way (Roxburgh and Matsuki 1999).
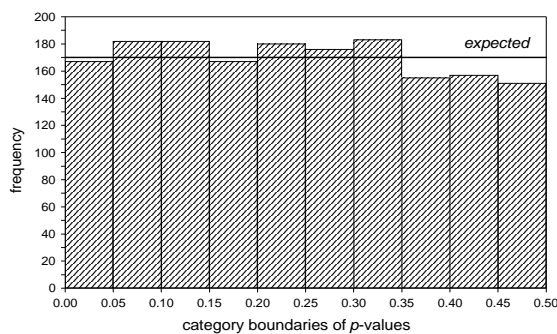

Fig. 2. Frequency histogram of the distribution of *p*-values referring to *dAss*, resulting from validation test of random shift model. Since the number of *p*-values is 1700, the expected frequency at each category is 170. Note, that x-axis ranges to 0.5, because of the applied formula of significance level.

We accomplished the validation test for random shift model with keeping the above notes in mind. We chose the stripe A from the whole grid, and executed random shift procedure on its field pattern making random patterns. The number of such *pseudo-observed patterns* were 100 (each contained 4000 cells). For each pattern the *dAss* curves were calculated using 99 randomizations (also with random shift), and then the significance level was computed. Since there is no real spatial dependence in pseudo-observed patterns, a significant *dAss* value can be detected as a result of random coincidence. Therefore we expected the probability for a significance level to fall constantly within the

different *p*-value categories. The result must be a *rectangular distribution* of the frequency histogram, as we can see on Fig. 2, which is made by accounting all 17 spatial series steps from 100 pseudo-observed patterns). To detect any departure from rectangular distribution, a *Kolmogorov-Smirnov one-sample test* and χ^2*-test* were applied, which gave non significant values ($d_{max} = 0.0276$, $p < 0.2$; $\chi^2 = 8.0353$, $p = 0.43$). When we calculated d_{max} and χ^2 for each spatial series steps separately, we got also non significant deviation in all steps.

For completing the validation test the calculation of the probability of type I error is also necessary. We chose two significance level ($p = 0.01$ and $p = 0.05$), and counted the number of significant positive or negative *dAss* values at each spatial series step. The frequencies, given as relative values, are exactly the rate of type I error, because no significant differences are expected between pseudo-observed (actually random) patterns and their random references. It can be seen in Table 3, that the probability of type I error corresponds to both significance levels (considering both negative and positive differences). A *Wilcoxon matched pairs test* was applied to detect statistically any difference between observed and expected relative frequencies of significant *dAss* values along spatial scale. The test confirmed that the type I error rate approximated the chosen *p* level (Table 3).

Table 3. The relative frequencies of significant negative (*rFSD* –) or positive (*rFSD* +) *dAss* values from 100 pseudo-observed patterns, at two significance levels. The Wilcoxon *T* for differences between observed (as probability of type I error) and expected *rFSD* values (given by *p*) are also listed.

Area (m^2)	$p = 0.01$		$p = 0.05$	
	rFSD –	rFSD +	rFSD –	rFSD +
0.0025	0.01	0.01	0.05	0.02
0.0050	0.03	0.01	0.06	0.02
0.0100	0.00	0.00	0.05	0.05
0.0150	0.01	0.00	0.07	0.04
0.0220	0.01	0.01	0.07	0.05
0.0400	0.01	0.01	0.06	0.04
0.0630	0.00	0.01	0.04	0.03
0.0900	0.01	0.02	0.04	0.03
0.1200	0.00	0.02	0.04	0.05
0.1600	0.01	0.00	0.05	0.04
0.2000	0.01	0.00	0.07	0.03
0.2500	0.00	0.00	0.03	0.05
0.3600	0.02	0.02	0.04	0.09
0.4900	0.01	0.03	0.04	0.05
0.6400	0.03	0.03	0.06	0.06
0.8100	0.00	0.01	0.08	0.09
1.0000	0.01	0.01	0.02	0.06
average	0.0100	0.0112	0.0512	0.0471
Wicoxon T	17.5	27.5	57.0	30.0
p for Wilcoxon T	0.944	0.625	0.865	0.279

This and the previous results show, that applying the random shift model in a statistical test, it is neither too liberal, nor too conservative. So, regarding to *dAss* values, the model can be validated. Note, that the validation process was successful even when only 99 randomizations were executed.

Characteristics of whole sampling area

We have calculated the basic syncretic functions along the spatial scale to characterize entire community regarding to whole sampling area of $10\,m \times 5\,m$ grid (Fig. 3a). We have also computed the difference values using random references (Fig. 3b-c). It can be seen on Fig. 3a, that the *fAss* values are very small compared to other functions, but not zero. Fig. 3a shows, that only low degree of spatial dependence appears agreeing with pioneer community, but it can be detected even if only ten taxa are present. Characteristic area of *fLD* (as compensatory area, A_{comp}), *fFD* (as A_{flor}) and *fDiss* is $0.16\,m^2$ equally, however for *fAss*, $A_{ass} = 0.49\,m^2$. The minimum area (in the sense of Juhász-Nagy and Podani 1983), $A_{min} = 12.25\,m^2$. The ordering of maximum areas gives the relation as

$$A_{flor} = A_{comp} < A_{ass}\,.$$

The *characteristic interval* is located between $0.01\,m^2$ and $0.64\,m^2$.

Analysing the *dAss* curve coming from complete randomization procedure, it can be almost the same as *fAss* values (Table 4, Fig. 3b). This phenomenon is due to the given null model which make the spatial dependence among species completely impossible, and indicates that the textural constrains are negligible in our sample. Whereas, the influence of patchiness of species pattern can not be neglected, because the *dAss* curve arising from random shift model is sufficiently different (cf. Table 4 and Fig. 3c). Significant *dAss* values appear at smaller areas, from $0.0025\,m^2$ to $0.09\,m^2$, and the maximum values of them is located at $0.0225\,m^2$. It means, that the organization of community pattern reaches its maximum value in a $15\,cm \times 15\,cm$ area.

Table 4. The characteristic associatum values and their areas for whole sample area.

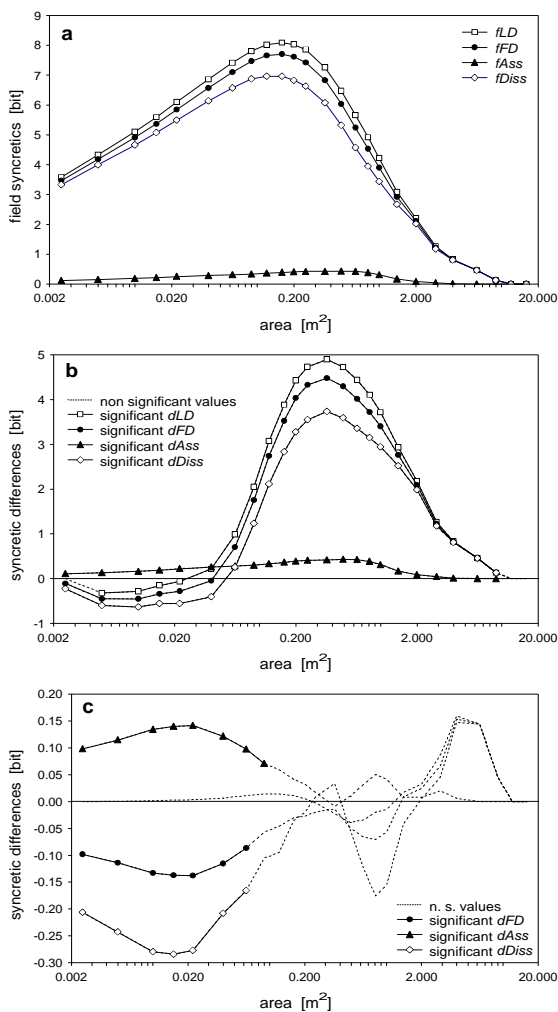| Function	maximum value (bit)	maximum area (m²)
FAss	0.435	0.4900
dAss from complete randomization	0.430	0.4900
dAss from random shift	0.142	0.0225

Fig. 3. Syncretic functions (a) and their difference values (b-c) from whole sample area. Random reference was represented by 99 compete randomizations (b), and random shifts (c) respectively. The significance level, $p = 0.01$. There are no significant *dLD* values in case of random shift, because this null model does not change the autocorrelation of species patterns.

Variability of associatum of stripes in relation to whole sampling area

Before analysing spatial dependencies among populations in the five stripes, we studied some textural characteristics of patterns. It can be seen on Table 5, that the relative frequency of a species can varies among stripes to a relatively high (cf. *Grimmia pulvinata*) or less degree (*Deschampsia flexuosa*). It is remarkable, that the *variation coefficient* of taxon saturation (i.e. sum of relative frequencies) within the stripes is very low, so the density of presences is rather similar in each stripe. Relative frequency of

the plots containing given number of taxa varies within the stripes to relatively low degree (Fig. 4). The most frequent are those microquadrats in which just one taxon appears (50-60 %), while frequency of empty plots is equal to those that contain two populations (about 20 %).

Table 5. The relative frequency values of each taxon and all of them within the five stripes. The mean, the standard deviation and the variation coefficient of frequencies are also listed. For taxon names refer Table 1.

| Tax.	Stripe					Mean	Std. Dev.	Var. Coeff.
	A	B	C	D	E			
Pu	48.68	51.75	37.53	25.35	35.10	39.68	10.695	0.270
Ac	33.63	29.30	33.48	36.80	43.10	35.26	5.129	0.145
Gp	1.10	7.10	16.15	3.13	8.73	7.24	5.836	0.806
Df	7.10	5.03	5.15	6.35	3.88	5.50	1.252	0.228
Dc	3.18	4.10	3.20	4.63	11.15	5.25	3.355	0.639
Ra	2.10	3.75	4.78	6.30	3.35	4.06	1.579	0.389
Cp	2.53	2.13	3.63	5.20	2.83	3.26	1.216	0.373
Ep	1.98	1.05	0.98	1.90	2.03	1.59	0.525	0.331
Li	0.78	0.43	0.53	1.15	1.40	0.86	0.414	0.484
Hl	0.18	1.95	1.35	0.10	0.03	0.72	0.877	1.218
all	101.2	106.6	106.8	90.90	111.6	103.4	7.892	0.076
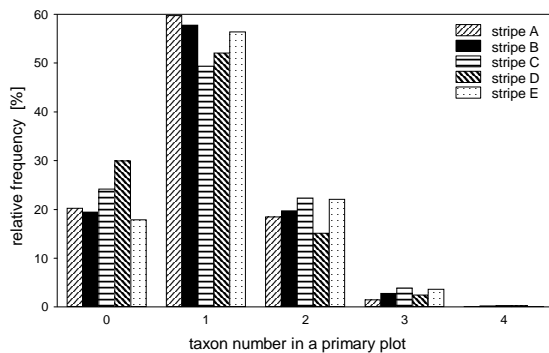
Fig. 4. Relative frequency histogram of the primary plots (microquadrats) containing given number of taxa in the five stripes.

The *fAss* curves of each stripes and whole sampling area are depicted on Fig. 5. An essential difference can be seen between the whole area and stripes. One reason of this deviation is that the *sample size were different* (cf. Table 2). (The effect of sample size can not be neglected even when the *fAss* values are standardized by it.) Larger sample size generates higher values for stripes in this case, but – because it was always the possible largest due to complete sampling procedure – it can not change the relative shape of an *fAss* curve. We can recognize however, that the curve for whole area is more

flattened without any well expressed maximum value, compared to the average curve. Otherwise, the area of maximum *fAss* of average curve is rather similar to A_{ass} of whole area (0.36 m^2 and 0.49 m^2).
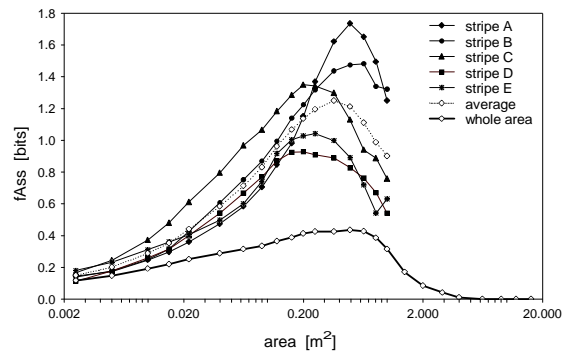

Fig. 5. The *fAss* curves of five stripes, their average, and whole sampling area.

Regarding to all *fAss* curves of stripes, both maximum value and its area are sufficiently variable. While *fAss* maximum is 0.928 bit in stripe D, in case of stripe A its value is nearly doubled (1.736 bit). Moreover, the A_{ass} varies between 0.2 m^2 and 0.64 m^2.

The *dAss* curves of stripes are also various, but theirs significant sections are more similar (Fig. 6). The smallest significant *dAss* maximum is 0.132 bit (in stripe D), and the largest one is 0.203 bit (in stripe C). The interval, in which the areas of *dAss* maximum of stripes are located, can be outlined between 0.01 m^2 and 0.0625 m^2. It is remarkable, that the average *dAss* curve of stripes is so much similar to curve of whole area, that the *dAss* maximum values are equal (0.14 bit), and the A_{ass} values, respectively (0.0225 m^2, cf. Fig. 6).
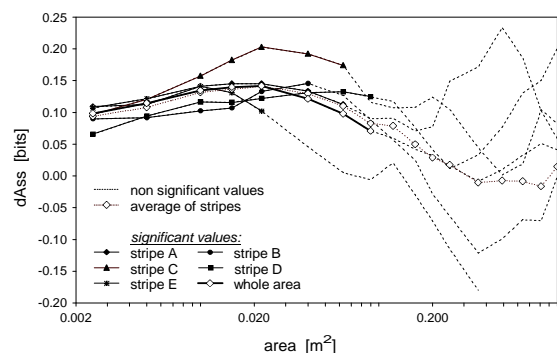

Fig. 6. The *dAss* curves of five stripes, their average, and whole sampling area. 99 random shifts were executed, $p = 0.01$.

For comparing different subtypes of associatum measures, we calculated the variation coefficient of functions along spatial scale, regarding to five stripes (Table 6). There are no considerable differences between *dAss* and ca*l*Ass, while the *senAss* has a little bit greater values. Generally they have larger variation coefficient than the other three measures. The *fAss* and *relAss* are similar to one another, but the variation of *relAss* is smaller, *redAss* is represented by smaller coefficients in first several spatial series steps, but larger values in second half of spatial scale. (The curves are depicted on Fig. 8.)

Table 6. Variation coefficients of different associatum functions along spatial scale from five stripes. The coefficient has not been calculated when at least one associatum value was negative.

Area (m^2)	fAss	redAss	relAss	dAss	calAss	senAss
0.0025	0.1768	0.0285	0.1500	0.1852	0.2096	0.2201
0.0050	0.1722	0.0385	0.1282	0.1317	0.1389	0.1452
0.0100	0.1871	0.0527	0.1452	0.1666	0.1822	0.1899
0.0150	0.2120	0.0643	0.1739	0.2178	0.2220	0.2365
0.0220	0.2232	0.0847	0.1890	0.2698	0.2650	0.2869
0.0400	0.2215	0.1147	0.2002	0.4078	0.4023	0.4239
0.0630	0.2187	0.1557	0.2002	0.5704	0.5523	0.5728
0.0900	0.1747	0.1844	0.1534			
0.1200	0.1409	0.2307	0.1136			
0.1600	0.1360	0.2975	0.1117			
0.2000	0.1450	0.3656	0.1222			
0.2500	0.1737	0.4226	0.1608			
0.3600	0.2431	0.5544	0.2232			
0.4900	0.3198	0.5880	0.2932			
0.6400	0.3852	0.5389	0.3059			
0.8100	0.4210	0.4536	0.2986			
1.0000	0.4016	0.3475	0.2233			

Table 7. Variation coefficients of different associatum functions along spatial scale from 20 blocks. The coefficient has not been calculated when at least one associatum value was negative.

| Area (m^2)	fAss	redAss	relAss	dAss	calAss	senAss
0.0025	0.3689	0.0622	0.2960	0.5055	0.4572	0.4794
0.0050	0.3368	0.0691	0.2724	0.4933	0.4587	0.4761
0.0100	0.3308	0.0922	0.2717	0.5758	0.5359	0.5662
0.0150	0.3330	0.1137	0.2768	0.6878	0.6420	0.6844
0.0220	0.3259	0.1467	0.2659			
0.0400	0.3199	0.2040	0.2669			
0.0630	0.3450	0.2890	0.3000			
0.0900	0.3514	0.3695	0.3082			
0.1200	0.3491	0.4469	0.3100			
0.1600	0.3614	0.5198	0.3262			
0.2000	0.3800	0.5495	0.3349			
0.2500	0.3932	0.4955	0.3275			
0.3600	0.4824	0.3419	0.3596			
0.4900	0.6711	0.3766	0.4968			
0.6400	0.9034	0.3358	0.7263			
0.8100	1.1310	0.2985	1.0293			
1.0000	1.1453	0.2304	1.0586			
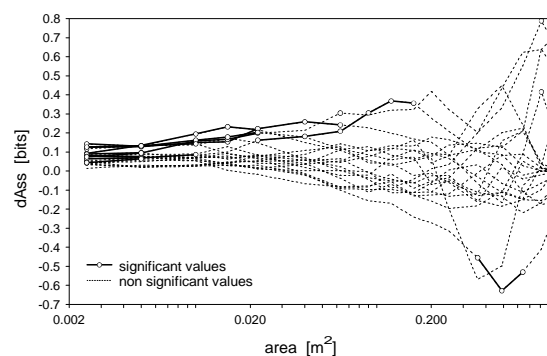
Fig. 7. The *dAss* curves of 20 blocks from field. Random reference was represented by 99 random shifts, $p = 0.01$.

Variability of associatum of blocks

As we continued the division of the stripes we have got altogether 20 blocks with 2.5 m^2 from the whole sampling area. Calculating *dAss* curve for each blocks, we can experience that the variability of spatial dependence increases if the sampling area decreases, as it was expected (Table 7). The relations among the different associatum measures in degree of variance are similar as they were noticeable in case of stripes.

Analysing the curves of *dAss* of blocks (Fig. 7), it can be seen that significant *dAss* values will appear more accidentally along the spatial scale. Some curves exhibit significant differences in neither spatial series step. Moreover, only 11 of the 20 curves have more than one significant values along the spatial scale. We can conclude the area of blocks being so small, that it can not represent perfectly the total sampling area, and the effects of a local pattern to the pattern organization can obscure the characteristics of entire community.

We have revealed that each *dAss* curve of stripes and half of blocks have significant domains. They appear when we compare *fAss* curve of a parcel to its „own" random references. The next question is whether *dAss* (or other associatum) curves of all stripes (or blocks) can be significantly different from those which are originated from stripes of random patterns. In this case we have to study the overlap between a number of associatum curves from observed and random patterns, as we will explain in the next chapter.

Comparison of associatum curves from field data to those from random patterns

If the community of the whole sampling area is considered as a unit of vegetation (examining its spatial dependencies), and this unity does not damage when the grid is divided into parts, then we should have to find that associatum curves
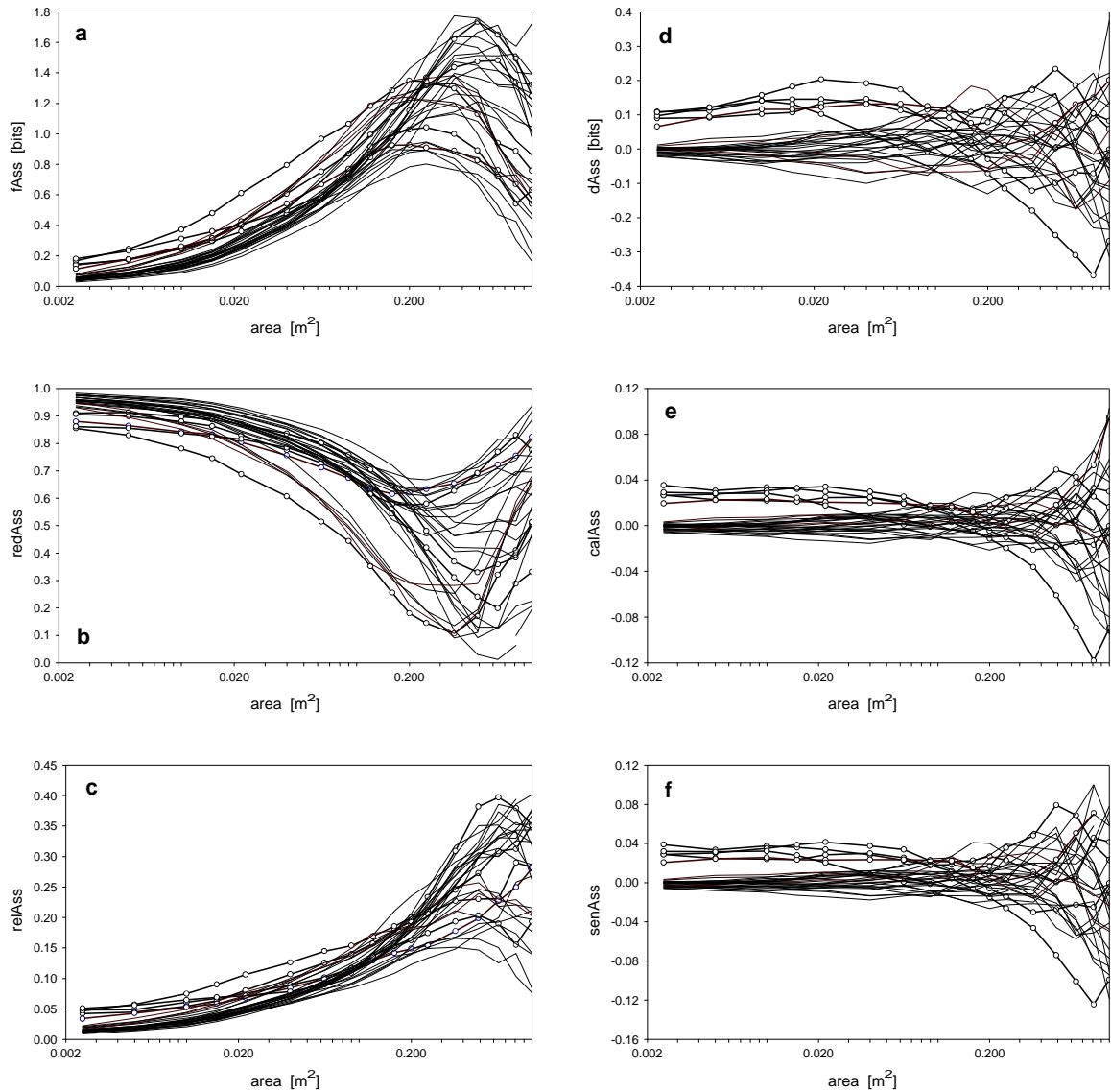
Fig. 8. The different associatum curves of 5 stripes from field pattern (thick lines with circles), and 25 stripes from 5 random patterns made by random shift (thin solid lines). For *dAss*, *calAss* and *senAss* 99 randomization were executed.

of parts significantly differ from all curves calculated from parts of random patterns. The results explained above do not support the concept of unity, because both textural and structural measures exhibit considerable variability (cf. Table 5 and 6), and we have not had any information about associatum values of a null model yet. For solving this problem we made five random patterns applying random shift procedure, then designated the boundaries of five stripes in each random pattern (as we did it with field sample), and finally compared the five associatum curves from field to 25 other ones from random

patterns (Fig. 8a-f). To detect any differences in median values, we applied *Mann-Whitney U-test* at each spatial series step (Table 8).

We can conclude that each applied measure of spatial dependence reveals significant difference from random patterns, but only in first half of spatial series steps. The significant interval of spatial scale however is shorter in case of *fAss*, *redAss* and *relAss*, than associatum functions using random references. Moreover, while regarding to the second three functions (*dAss*, *calAss*, *senAss*), this interval corresponds to those in which the *dAss* itself also has

the significant values (among them the maximum value, cf. Fig. 6), but in view of *fAss*, *redAss* and *relAss*, the maximum values appear outside of this significant interval (cf. Fig. 5). It means, that at characteristic associatum area we can find differences between field and random patterns if we use just the associatum difference (*dAss*) and its derived functions (*calAss* and *senAss*). The results show that when we divided the whole area into five stripes, the unity of community remained uncorrupted. So, each stripe with 10 m² can represent the 50 m² area of stand.

Table 8. Significance levels for Mann-Whitney test. The null hypothesis is that the median of associatum values from 5 field stripes is equal to the 25 ones calculated from strips of random patterns. The sign of significance levels: *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, while no marker indicates the non significant difference in a spatial series step.

Area (m²)	fAss	redAss	relAss	dAss	calAss	senAss
0.0025	***	***	***	***	***	***
0.0050	***	***	***	***	***	***
0.0100	***	**	***	***	***	***
0.0150	**	**	***	***	***	***
0.0220	**	*	**	***	***	***
0.0400	*		*	**	**	**
0.0630				**	**	**
0.0900				*	*	*
0.1200				*	*	*
0.1600						
⋮						
1.0000						

To prove that this result is not an artifact, we compared the associatum curves of 5 stripes from one random pattern to other 4×5 ones (using *Mann-Whitney test*, as well). Since there are 5 random patterns and 17 spatial series steps, we applied the test 85 times at each associatum measure. From 85 comparisons, there were no significant differences regarding to *fAss*, *redAss*, *relAss*, and there were only two departures in case of *dAss*, *calAss* and *senAss* at $p \leq 0.05$.

Examining the consequences of further division, we compared the 20 field blocks to the other 20 blocks originated from one random pattern made by random shift. The results are listed in Table 9. Now, the block represent less the unity of community, and they are more different from each other, since significant deviation of associatum measures can be found only in first 2 – 5 spatial series steps.

Table 9. Significance levels for Mann-Whitney test. The null hypothesis is that the median of associatum values from 20 field blocks is equal to the 20 ones calculated from blocks of a random pattern. The sign of significance levels same as in Table 7.

Area (m²)	fAss	redAss	relAss	dAss	calAss	senAss
0.0025	***	**	***	***	***	***
0.0050	***	*	***	***	***	***
0.0100	*		**	***	***	***
0.0150			*	***	***	***
0.0220				***	***	***
0.0400						
⋮						
1.0000						

Conclusions

(1) The random shift null model applied to *associatum difference* measure of large grid data have been validated. The probability of making the type I error is neither larger nor smaller than a chosen level of significance. In the validation procedure we used random shift not only for random patterns but to make pseudo-observed ones; this is in contradiction with Roxburgh and Matsuki (1999).

(2) The studied pioneer community exhibits low degree of mutual spatial dependence among populations. This pattern organization appears even when it is compared to random shift null model. It is noteworthy if we consider that the community consists of only ten taxa, and half of them are non vascular plants. Associatum values of the same order of magnitude were also detected in early stages of primary succession on dumps from open-cast coal mining in Hungary (Bartha 1990).

(3) The area of maximum value is 0.49 m² for *fAss*, but 0.0225 m² for *dAss* using random shift. This rather small (15 cm × 15 cm) area indicates, that we can find the community organizing effects among the interspecific interactions, and it is presumable, that the significant positive associatum is not a result of only spatial heterogeneity of abiotic conditions. The ordering of maximum areas gives the relation as $A_{flor} = A_{comp} < A_{ass}$. Regarding to *characteristic ordering*, *characteristic interval* and the concrete values of characteristic areas, the result is most similar to ten years old stand of study cited above (Bartha 1990). It seems that the pattern organization process takes place similarly in the different primary succession on clayey spoil banks, independently of floristic composition.

(4) The average of *dAss* curves from stripes is more similar to associatum curve from the whole sampling area than average of *fAss* values. It means that the *dAss* originated from sufficiently large part

of a stand can be more effective estimator of the characteristics of the „whole" community.

(5) The variability of the examined associatum curves can not be neglected, and it increases when the sampling area decreases. While the variation coefficients of *fAss*, *redAss* and *relAss* have slightly lower values than *dAss*, *calAss* and *senAss*, the last three measures can be regarded as the „better" ones, because in spite of this larger variability they can exhibit significantly different values from random patterns in broader interval of spatial scale. Moreover, in case of *dAss*, this significant interval contains the characteristic (maximum) value.

(6) Though the degree of variability is considerable, there is an interval of spatial scale at which the median of associatum values from particular field pattern significantly differs from that originated from random patterns. This fact, which was observable in stripes, can indicate a unity of pattern formation of vegetation. We can conclude that the 10 m^2 of sampling area represents the entire community. On the other hand, the unity begins to collapse if the sampling area is only 2.5 m^2. In the blocks, the significant *dAss* values appear rather accidentally corresponding to the local pattern. So, the sampling area with 2.5 m^2 can not be considered as an adequate size of pattern organization unit of the examined community.

Acknowledgment

References

Bartha S. (1990): Spatial processes in developing plant communities: pattern formation detected using information theory. - In: Krahulec F., Agnew, A. D. Q., Agnew, S. and Willems J. H. (eds.): Spatial processes in plant communities. pp. 31-47. Academic, Prague.

Bartha, S. (1992): Preliminary scaling for multi-species coalitions in primary succession. - Abstracta Botanica *16*, 31-41.

Bartha, S., Collins, S. L., Glenn, S. M. and Kertész, M. (1995): Fine-scale spatial organization of tallgrass prairie vegetation along a topographic gradient. - Folia Geobotanica et Phytotaxonomica, Praha 30: 169-184

Bartha, S., Czárán, T., Oborny, B., Podani, J. and Kertész, M. (1994): JNP-modellek 1.0 - Számítógépes programcsomag a cönológia koegzisztenciális mintázatainak detektálására Juhász-Nagy Pál információ-statisztikai modellcsaládjával. (Computer program package to detect coexistence patterns of cenology by Juhász-Nagy's information statistical model-

family.) - III. Magyar Ökológus Kongresszus, Szeged. Elõadások és poszterek összefoglalói, p.17.

Bartha, S. and Kertész, M. (1998): The importance of neutral-models in detecting interspecific spatial associations from 'trainsect' data. -Tiscia *31*, 85-98.

Bartha, S., Rédei, T., Szollát, Gy., Bódis, J. and Mucina, L. (1998): Északi és déli kitettségû dolomitsziklagyepek térbeli mintázatainak összehasonlítása. (Compositional diversity and fine-scale spatial patterns of dolomite grasslands on contrasting slopes.) - In: Sziklagyepek szünbotanikai kutatása (Ed. Csontos, P.) Scientia Kiadó; Budapest, p. 159-182

Erdei, Zs. and Tóthmérész, B. (1993): MULTI-PATTERN 1.00. Program package to analyze and simulate community-wide patterns. - Tiscia *27*, 45-48.

Hochstrasser, T. (1995): The structure of different loess grassland types in Hungary. Diploma work; Vácrátót.

Horváth, A. (1997): Löszgyep-típusok términtázatának informá-cióstatisztikai vizsgálata. (Information statistical analysis of loess steppe stages) - IV. Magyar Ökológus Kongresszus, Pécs. Elõadások és poszterek összefoglalói, p.85.

Horváth, A. (1998): INFOTHEM program: new possibilities of spatial series analysis based on information theory methods. - Tiscia *31*, 71-84

Juhász-Nagy, P. (1972): Elemi preferenciális folyamatok információelméleti modellezése szünbotanikai objektumo-kon. (Information theory models of elementary preferential processes on synbotanical objects.) Kandidátusi értekezés. Budapest.

Juhász-Nagy, P. (1976): Spatial dependence of plant populations. Part 1. Equivalence analysis (an outline for a new model). - Acta Bot. Acad. Sci. Hung. *22*, 61-78.

Juhász-Nagy, P. (1980): A cönológia koegzisztenciális szerkezeteinek modellezése. (Models of the cenological coexistence structures.) Akad. Dokt. Ért. Budapest.

Juhász-Nagy, P. (1984): Spatial dependence of plant populations. Part 2. A family of new models. - Acta Bot. Acad. Sci. Hung. *30*, 363-402.

Juhász-Nagy, P. and Podani, J. (1983): Information theory methods for the study of spatial processes and succession. - Vegetatio *51*, 129-140.

Manly, B. F. J. (1997): Randomization, bootstrap and Monte-Carlo methods in biology. (2nd edn.) Chapman & Hall, London.

Margóczi, K. (1995): Interspecific associations in different successional stages of the vegetation in a Hungarian sandy area. - Tiscia *29*, 19-26.

Palmer, M. W. and van der Maarel, E. (1995): Variance in species richness, species association, and niche limitation. - Oikos *73*, 203-213.

Podani, J. (1993): SYN-TAX-5.0: Computer programs for multivariate data analysis in ecology and systematics. - Abstracta Botanica *17*, 289-302.

Roxburgh, S. H. and Chesson, P. L. (1998): A new method for detecting species associations in the presence of spatial autocorrelation. - Ecology *79*, 2180-2192.

Roxburgh, S. H. and Matsuki, M. (1999): The statistical validation of null models used in spatial association analyses. - Oikos *85*, 68-78.

Szollát, Gy. and Bartha, S. (1991): Pattern analysis of dolomite grassland communities using information theory models. - Abstracta Botanica *15*, 47-60.