

THE IMPORTANCE OF NEUTRAL-MODELS IN DETECTING INTERSPECIFIC SPATIAL ASSOCIATIONS FROM 'TRAINSECT' DATA

S. Bartha and M. Kertész

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

Abstract. We compared two Monte-Carlo randomization tests as neutral models, for measuring interspecific spatial associations. Complete randomization randomly re-allocates the individuals of each species into the spatial sample with the same abundances as in the studied sample. Thus, this separates two components of association: (1) textural, i.e. related to the species abundance distribution, and (2) structural, i.e. related to the spatial distribution of the species. Random shifts keep the spatial patterns of individual species, but change their relative positions. Thus, this method separates (1) the interspecific components of associations from (2) the intraspecific ones, randomizing the first. We also applied a new algorithm for determining the signs of pairwise associations (i.e. positive or negative). This technique is based on the comparison of the observed and simulated joint distributions of a species pair. We tested the neutral models and sign determination algorithms on multispecies circular transect ('trainsect') data. All statistics were calculated across a range of scales, i.e. the calculations were repeated at a series of sampling unit lengths. One field transect and four simulated patterns were examined. The simulated patterns, created by spatially explicit individual based model, represent known assembly rules. We concluded that circular transects are appropriate in field studies for detecting fine scale community patterns because their topology allows us to apply more types of randomizations. The separation of textural, intraspecific and interspecific effects provided us more opportunity to detect and interpret the spatial associations. The new sign determination algorithm performed better than the traditional method. The need of developing neutral models based on biological mechanisms instead of simple statistical assumptions and the dynamical relevance of textural and structural constraints are also discussed.

Keywords: circular transect, computerized sampling, information theory, randomization tests, spatially explicit Monte-Carlo simulation, sandy grasslands, scale dependence, spatial autocorrelations, textural versus structural constraints

S. Bartha, M. Kertész, Institute of Ecology and Botany of the Hungarian Academy of Sciences, H-2163 Vácrátót, Hungary

Introduction

Interspecific association, as a basic coenological phenomenon (Juhász-Nagy, 1980), is one of the most important aspect of community structure. It measures the tendency of species to co-occur or to avoid each other (e.g. Kershaw 1964, Greig-Smith 1983). The spatial associations of species have important consequences on their interactions, dynamics, and coexistence (Czárán and Bartha, 1992, Czárán, 1998). Therefore spatial associations are

often interpreted in terms of assembly rules, i.e. pattern generating factors (Watkins and Wilson, 1992, Wilson, 1994, Wilson and Watkins, 1994, Gotelli and Graves, 1996). Despite its importance, case studies on interspecific associations are scarce (for a few exception see e.g. de Vries 1953, Agnew 1961, Fekete and Szujkó, 1973, Lepš and Buriánek, 1990, Matus and Tóthmérész, 1990, Bartha, 1992, Hauser, 1993, Margóczy, 1995). Reviewing this literature, it is difficult to generalize the results and create general rules or laws about community

structure and organization. We believe that the unsolved methodological problems associated to this field are responsible mainly for this 'state of the art'.

There are many attempts for developing new methods for detecting pairwise spatial associations (cf. Greig-Smith, 1983). However, they are often limited to a single scale (e.g. Yarranton, 1966) or they are biased by within-species autocorrelations and indirect effects of the other species (Juhász-Nagy 1980, Dale *et al.*, 1991). Recently, Palmer and van der Maarel (1995) reviewed these methodological problems and suggested to use certain types of randomizations to separate these effects. They suggested to use circular transect, called 'trainsect' to provide opportunity for several types of randomizations.

In the present study, we test their suggestions and compare the performance of different neutral models for detecting the significance of associations, and techniques for determining their positive or negative signs. For testing the performance of these methods, we use simulated data generated with known assembly rules (i.e. rules for the spatial relationships of individuals), and field data from perennial sand grasslands. We use Juhász-Nagy's information theory models (Juhász-Nagy 1967, 1980, 1984) for calculating pairwise associations. Our present study does not deal with the effect of other species on the pairwise association. However, we chose the models of Juhász-Nagy because they allow us to generalize from the pairwise relationships to multiple-, and partial associations, and to the multi-species spatial dependences in coalitions or in the whole community.

Our principal aim is to develop and test a simple technique for detecting spatial associations in the

field. There is a need for simple, fast, still confident sampling designs and receipts of analyses that support huge field campaigns of data collections and allow us to compare dozens of data sets collected and analysed with the same standard technique.

Methods

Simulated patterns

To test the performance of different versions of spatial pattern analyses, two-species patterns with contrasting intra-, and interspecific spatial dependence between individuals were generated with a spatially explicit simulation model (DIVGEN1). DIVGEN1 is a simplified, one dimensional version of PATPRO (Czárán, 1984, 1993). The model generates stochastic point processes along a continuously scaled circular line in discrete time units. In the first generation, colonizing individuals appear in random positions along the transect. In the second generation, species reproduce and their offsprings disperse around the parent individuals. Distance of offsprings from their parent individual is a stochastic variable with either a Gaussian or a Poisson distribution. Survival of an offspring individual of species i , $p_{surv}(i)$, depends on the number of individuals of species i (n_i) and of species j (n_j) in a neighbourhood d . $p_{surv}(i) = g_{i,i}^{n_i} \times g_{i,j}^{n_j}$, where $g_{i,i}$ and $g_{i,j}$ are indicators of the intensity of interaction. Competitive effects are multiplicative. Survivorship decreases exponentially with the increasing number of individuals in the neighbourhood. Using this model with different assembly rules we created simulated patterns with different intra-, and interspecific spatial dependence of individuals

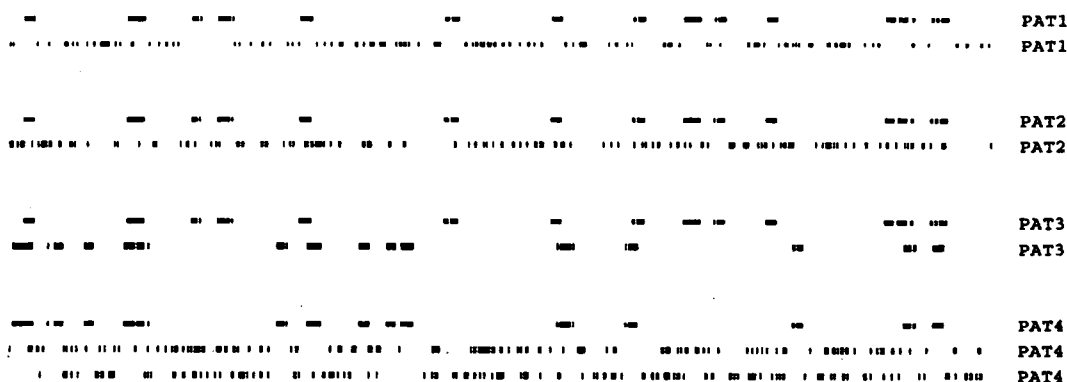


Fig. 1. Simulated spatial patterns of species along the transects (presences are marked by black dots). PAT1: species 1 is a dominant with limited dispersal, species 2 is subordinated, i.e. competitively inferior, with unlimited dispersal. PAT2: species 1 with limited dispersal, species 2 with unlimited dispersal, no competitive interaction between them. PAT3: both species with limited dispersal, no competitive interaction between them. PAT4: species 1 dominant with limited dispersal, species 2 and 3 are subordinated, i.e. competitively inferior, with unlimited dispersal. (In this paper, the association was analysed only between species 2 and 3.)

(Fig.1). The size of the plot (circular transect) was 1024 units, same as of the field sample.

Pattern 1

The first (dominant) species has limited dispersal. The second (subordinate) species disperses randomly. The dominant species has a strong competitive effect on the subordinate species. As a result, negative interspecific association is expected between the two species. The dominant species should have an autocorrelated (aggregated) pattern due to its limited dispersal while the individuals of the subordinate species could appear either random or slightly aggregated because they are restricted to the gaps between the patches of the dominant.

Pattern 2

The first species has limited dispersal, the second disperses randomly. There is no competitive effect between the species. The first species should be patchy, the second randomly distributed, and they should be spatially independent.

Pattern 3

Both species has limited dispersal, but there is no competition between them. Consequently, both species should have an autocorrelated (aggregated) pattern with no interspecific association between them.

Pattern 4

Three species were simulated here. The first (dominant) species has limited dispersal and it has the same strong competitive effect on both subordinate species. The two subordinate species could spread randomly and have no competition between them but their distributions are similarly restricted by the dominant species. We are interested in the pattern of the two subordinate species. As an indirect effect of the dominant competitor, positive interspecific association is expected between the subordinate species. Their individual patterns could be either random or slightly aggregated as they were restricted into the gaps of the dominant.

The parameters for invasion rates, fecundity of individuals, competitive strengths, and the dispersal parameter for limited dispersal are the same for the species in all simulations. The only difference was that competitive effects and/or dispersal limitations were switched on or off in different combinations.

The simulated continuous point patterns were transformed into circular transects of 1024 contiguous, discrete units. This extent and resolution are very close to the scaling applied at the field data collection. Choosing appropriate parameter values

we could produce patterns with the same 12.5% frequency for each species. The species differ only in their individualistic (intraspecific) and collective (interspecific) spatial dependence.

Field data

We tested the performance of the analyses on field data from a perennial sand grassland community. The study site is a protected sand dune area near Csévharaszt, in the northern part of the Danube-Tisza Interfluve. The vegetation consists of a mosaic of woods (oak, poplar, juniper, and black locust) and grasslands (consisting of stands dominated by annuals, *Poa angustifolia*, cryptogams, *Festuca vaginata*, and *Stipa borysthena*). Mean annual temperature is 10.2 °C, annual precipitation is 545 mm. The soil is calcareous sandy soil with shallow humic layer (<10cm) of low humus content (< 1%). The sampled grassland was a ca. 20×30 m undisturbed opening in a Juniper-Poplar wood mixed with scattered *Robinia* trees. The grassland is dominated by *Festuca vaginata*, and has a well developed cryptogam layer (mainly *Cladonia magyarica* and *furcata*). *Stipa borysthena* is subdominant (its abundance was ca. the half of *Festuca*). Annuals (e.g., *Viola arvensis*, *Polygonum arenarium*, *Arenaria serpyllifolia*), and an ephemeral grass, *Poa bulbosa* are also abundant. Other perennial grasses (e.g. *Koeleria glauca*) and perennial forbs (e.g. *Potentilla arenaria*, *Cynoglossum hungaricum*, *Alkanna tinctoria*) appear only sporadically. Presences of plant species were recorded along a circular belt transect of 1040 5×5 cm contiguous sampling units (52m) in June, 1995.

Sampling from transects and data analysis

The field and simulated transects served as primary samples (cf. Horváth, 1998) or primary references (sensu Juhász-Nagy, 1980) for the subsequent computerized sampling (Podani, 1984). Complete sampling (Bartha *et al.*, 1995, Horváth, 1997) was performed across a range of scales from 0.05 m to 25 m, and the resulting *s*×*m* (species by sampling units) binary matrices were analysed using information theory models (Juhász-Nagy, 1980, 1984, Juhász-Nagy and Podani, 1983). We computed associations between two species from 2×2 contingency tables applying the mutual information between the species:

$$MI(A,B) = m \log m + a \log a + b \log b + c \log c + d \log d - (a+c) \log(a+c) - (b+d) \log(b+d) - (a+b) \log(a+b) - (c+d) \log(c+d)$$

where symbols a , b , c and d correspond to the notation of the 2×2 contingency table. a is the number of sampling units where both species, A and B , are present, d is the number of empty quadrats. b and c note the number of quadrats, where only one species is present, A or B respectively. $m = a + b + c + d$ is the sample size. If we use the unweighted form of this equation (i.e. not multiplied by m), then $I(A,B)$ ranges between $0 \leq I(A,B) \leq 1$. Thus, the equation is

$$I(A,B) = H(A) + H(B) - H(A,B)$$

i.e. the contingency information of the two-dimensional contingency table is the difference between the pooled entropy of marginals ($H(A) + H(B)$) and the entropy of species combinations within the table ($H(A,B)$) (Juhász-Nagy and Podani, 1983). This logic is easy to generalize to many species. Associatum refers to the overall spatial association in a community. If the community consists of s species, associatum is calculated as the contingency information of an s -dimensional contingency table, containing $2s$ cells (Juhász-Nagy, 1980, 1984).

$$I(A,B,\dots,S) = H(A) + H(B) + \dots + H(S) - H(A,B,\dots,S)$$

The significance of associations were detected by Monte-Carlo randomization tests. Two types of randomization were used as neutral models. (1) Complete randomization (Diggle, 1983) randomizes the positions of individuals (the presences in our case) along the transects, but the number of species and the frequency of species remains the same as in the field. (2) Random shifts (Palmer and van der Maarel, 1995, Horváth, 1998) keep the spatial pattern of each individual species, but randomize the relative (interspecific) positions of the species by shifting (rotating) it randomly along the circular transect. Thus, while complete randomization randomizes the pattern both within and between species, the random shifts randomize only the interspecific patterns.

Significance of any observed value ($I(A,B)$ or $I(A,B,\dots,S)$) was calculated by comparing it with the values obtained by the Monte-Carlo randomizations, i.e. representations of the null hypothesis. The significance was expressed as the relative frequency of the randomizations resulted in higher values than the observed one. 5000 randomizations were applied in each tests. The trend of the association, i.e. whether the association is positive or negative, is given by a comparison of the expected and observed values of the joint occurrence of species A and B . (Kershaw, 1964). We compared two methods for

determining the sign of association. We compared the sums of the frequencies of the diagonal cells in the 2×2 contingency table (comparing $a + d$ to $b + c$). For example, if $b + c$ is greater than $a + d$, species A and B occur together less than expected, so they are negatively associated. The other method follows the same logic but the expected frequency is obtained from the a Monte-Carlo randomization instead of calculating it from the marginal of the observed 2×2 contingency table. For example, species A and B are negatively associated if $b + c$ observed is higher than the average $b + c$ calculated from the random reference.

To avoid artefacts due to rarity and limited sample size, we analysed only the 10 most frequent species in the case of field data.

All statistics were calculated across a range of scales, i.e. the complete sampling and the calculations were repeated at a series of sampling unit lengths.

Results and discussion

Compared to other grasslands, e.g. the loess grasslands in Hungary, xeric perennial sand grasslands are considered as relatively simple systems with high stochasticity and relatively low degree of organization (Fekete, 1992). Their high diversity in life forms and ecophysiological strategies (Kalapos, 1994, Tuba *et al.*, 1998), their relatively low total biomass (Kovács-Láng, 1974), and the phenological differentiation of their species (Kárpáti and Kárpáti, 1954) suggest strong abiotic control on plant adaptation and community organization. Thus, it is supposed that there is considerable niche differentiation in this community, but the differentiation takes place rather in time (in the form of seasonal or inter-annual dynamics), than in space. We can expect low degree of spatial organization, i.e. low and non-significant spatial dependence between the species. In contrast with this expectation, we found significant and considerable departure from randomness within a wide range of scales (between 0.05m to 25m) in the case when complete randomization was used as a neutral model (Fig. 2a). This result corresponds with the large niche overlaps found by Fekete *et al.* (1995). Margóczy (1995) also obtained high spatial associations by investigating various types and successional stages of xeric sand grasslands. However, in our analyses, the other neutral model produced different results. Using random shifts, we found much less significant departure from randomness. The significant associations appeared only in a narrow range of spatial scales, between 0.1m and 0.4m (Fig. 2b). Table 1 shows the same con-

Table 1. Patterns of the signs of the pairwise spatial associations of the the seven most frequent species in a xeric perennial sand grassland in Csévharszt. The signs of the associations (+,-) were determined by the comparison of the sums of diagonals ($a+d$ vs. $b+c$) of the 2×2 contingency table.

Species codes:

CLAMAG *Cladonia magyarica*; CLAFUR *Cladonia furcata*; FESVAG *Festuca vaginata*; STIBOR *Stipa borysthenica*; POLARE *Polygonum aviculare*; POABUL *Poa bulbosa*; VIOARV *Viola arvensis*

critical value for significance $p < 0.01$

- . non-significant association
- +1 positive association
- 1 negative association

neutral model: complete randomization with fixed frequencies at the finest resolution														
length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275	495	1000
species pairs														
CLAMAG x CLAFUR	.	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	.	.	.
CLAMAG x FESVAG	.	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	.	.	.
CLAMAG x STIBOR	+1	+1	+1	+1	.	.	.
CLAMAG x POLARE	.	.	-1	-1	-1	-1	-1	+1	+1	+1	+1	+1	.	.
CLAMAG x POABUL	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	+1	.
CLAMAG x VIOARV	+1	-1	-1	-1	-1	-1	-1	-1	-1	+1	+1	+1	.	.
CLAFUR x FESVAG	+1	+1	+1	.	.	.
CLAFUR x STIBOR	+1	+1	+1	.	.	.
CLAFUR x POLARE	+1	+1	.	.	.
CLAFUR x POABUL	+1	+1	-1	-1	-1	-1	-1	-1	-1	-1	+1	.	.	.
CLAFUR x VIOARV	+1	.	.	.
FESVAG x STIBOR	.	+1	-1	-1	.	+1	.	.	.	+1	+1	.	.	.
FESVAG x POLARE	+1	+1	.	.	.
FESVAG x POABUL	+1	+1	-1	-1	.	.	-1	-1	-1	+1	+1	.	.	.
FESVAG x VIOARV	+1	+1	.	.	.
STIBOR x POLARE	+1	+1	.	.	.
STIBOR x POABUL	+1	.	.	.
STIBOR x VIOARV	-1	.	+1	.	.	.
POLARE x POABUL	.	.	+1	.	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
POLARE x VIOARV	.	.	+1	+1	+1	+1	+1	+1	.
POABUL x VIOARV	+1	.

neutral model: random shift of species along the circular transect														
length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275	495	1000
species pairs														
CLAMAG x CLAFUR	+1
CLAMAG x FESVAG	.	.	+1	+1	+1
CLAMAG x STIBOR
CLAMAG x POLARE
CLAMAG x POABUL
CLAMAG x VIOARV	+1	-1	-1
CLAFUR x FESVAG
CLAFUR x STIBOR
CLAFUR x POLARE
CLAFUR x POABUL	+1	+1	-1
CLAFUR x VIOARV
FESVAG x STIBOR	.	.	-1
FESVAG x POLARE
FESVAG x POABUL	.	+1
FESVAG x VIOARV
STIBOR x POLARE
STIBOR x POABUL
STIBOR x VIOARV
POLARE x POABUL	+1
POLARE x VIOARV
POABUL x VIOARV

trasting differences between the patterns of significant pairwise associations of the most frequent seven species determined by the two types of randomizations. Detecting by complete randomization, all pairs of the seven species show significant associations. Significant associations are scale dependent, with considerable variation. In other studies on community structure, significant departure from randomness were rarely found (e.g., Wilson *et al.*, 1987). We think that negative results do not necessarily prove the lack of spatial organization in plant communities. They rather indicate the limitations of the applied methodology.

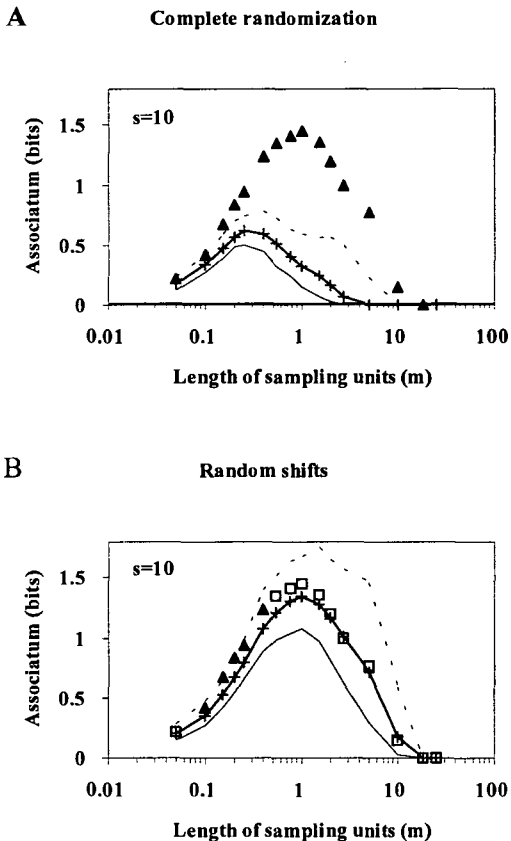


Fig. 2. Overall spatial associations (associatum) of the ten most frequent species in a xeric perennial sand grassland in Csévharaszt, detected by different neutral models: ▲ (filled triangle) significant departure of an observed value from the neutral model ($p < 0.01$), □ (empty square) non-significant observed values, + (plus) average of the random references based on 5000 Monte-Carlo randomizations, ----- maximum of the random references, — minimum of the random references.

In studies restricted to single (often the finest) scale, or restricted to a few arbitrary chosen scales only, the majority of information about community

structure may remain hidden. Scanning through a wide ranges of scales we found significant associations which do not occur at the finest scale (i.e. at the scale of the elementary 0.05 by 0.05 m sampling units). Beyond that, we had opportunity to study the pattern of associations as a function of scale.

Applying random shift as a neutral model (random reference), we found much less significant associations between species pairs. However, there was no contradiction between the two types of randomization. Results given by the random shifts appear simply as a subset of results detected by the complete randomization.

Complete randomization separates textural and structural effects on associatum, while random shifts separate intraspecific spatial dependence (autocorrelation) from the interspecific spatial dependence (interspecific association). We would expect zero associatum in case of complete randomization. However, associatum would be zero only in an ideal community where individuals would be dimensionless points and the transect would be continuous and infinite. However, vegetation forms complex mosaics, and individuals (ramets and genets) cover discrete areas of various sizes. Due to the finite and discrete nature of individuals, some (rare) combinations never appear, while other (frequent) combinations are well represented ('textural constraints', cf. Bartha, 1992). Associatum is sensitive to differences between the frequency distributions of species combinations in ideal versus real plant communities. Applying spatially explicit individual based models, recent theoretical studies suggest that textural constraints have very important dynamical consequences on species coexistence, and consequently, on the estimations of diversity (for a review see Czárán, 1998). Intra-, and interspecific spatial dependence as 'structural constraints' modify further the number and frequency of realized species combinations, thus, they modify the interactions and dynamics of populations.

It is important to note that with applying several neutral models together, we could separate these effects step by step. With these neutral models we do not simply remove the artifacts from our results, but we can distinguish and measure the relative importance of different textural and structural constraints on the dynamics of populations and communities. We emphasize that the constraints have effects on the dynamics, and not on the patterns that appear in single-time samples, which represent only "snap-shots" from dynamic processes. As we can see on Fig.2a and Fig.2b, the average random associatum is not zero in the neutral models. However, it does not simply mean that associatum is

a wrong measure because it shows 'artefact association' in case of 'randomness'. What associatum shows is the deviation of the frequency distribution of realized species combinations from an expected frequency distribution in an ideal plant community, where species can coexist in a completely unconstrained way. Because the number, frequency, and identity of realized species combinations inform us about the coexistence relations of species in a community, any deviation from the ideal (unconstrained) coexistence should contain some important information about the organization of a community.

The term "species association" in its traditional form was introduced for expressing the tendency of species to co-occur or to avoid each other (e.g. Kershaw, 1964, Greig-Smith, 1983). The significant tendencies of co-occurrences are thought to form the discontinuities of the vegetation at a longer, even an evolutionary time scale. For example, these discontinuities are which can be classified as syntaxonomic units. This involves that significant associations are the most important measures of community organization. However, few studies analysed directly how effective is a certain interspecific association in determining species coexistence or community dynamics. Fig.3 shows the frequency distribution of the values of maximum associations found in the studied stand of xeric sand grassland. These values are very small comparing to the theoretical maximum of $I(A,B)$ which is 1 bit. Consequently, even if there are significant associations among species, their effects may be weak on species coexistence, i.e. the frequencies of realized species combinations differ only slightly from the random expectations (neutral models).

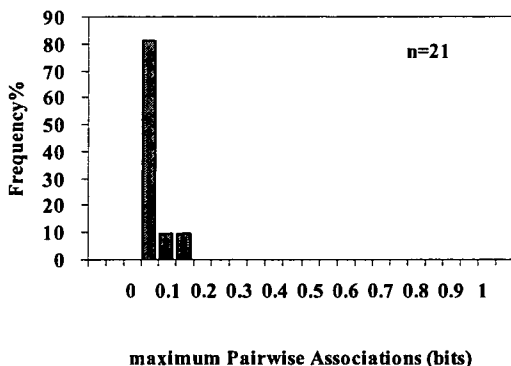


Fig. 3. Frequency distribution of the maxima of pairwise associations calculated from all the possible pairs of the seven most frequent species in a xeric perennial sand grassland in Csévharaszt. (Note the small absolute values of maxima relatively to the potential upper limit, $I(A,B)=1$.)

There are several attempts to remove the textural effects and the effects of autocorrelation as 'artefacts' from the estimated values of species associations (Dale et al, 1991, Hauser, 1993). They over-emphasize one type of structural constraint and tend to neglect the others. Although the actual abundance of species, and their intraspecific spatial patterns (autocorrelations) are individualistic properties, they have evolved in a community, so they are not independent characters. Textural and structural constraints are dynamically related, therefore the task of separating them is not simply a statistical issue.

Scale dependence and changes of the signs of associations in spatial series provide further opportunities in interpreting results on species coexistence. Kershaw (1964) and Greig-Smith (1983) expected a logical pattern of the change of the signs with increasing plot sizes. Associations start with negative values as plant individuals exclude each other at fine scales, i.e. at scales comparable to their sizes. With increasing sampling unit sizes, associations might become positive, that will disappear at even coarser scale, where both species are present in every sampling unit. From this basic logic the following variants can be expected with increasing sampling unit sizes: negative to positive to neutral, positive to neutral, negative to neutral, all neutral. There are case studies that supported these trends (e.g. Fekete and Szujkó, 1973, Bartha, 1983), while others reported more complex patterns (e.g. Hauser 1993, Podani and Czárán 1997). We found complex patterns too: e.g., neutral to positive to neutral, neutral to negative to positive to neutral, or positive to negative to positive to neutral. Comparing a large set of results (Bartha, unpublished), the most general pattern of the signs of significant associations appear to be positive to negative to positive with the increasing plot size.

Because field data are usually too complex and details about pattern generating factors (assembly rules) are unknown, we used the simulated data to test further the performance of the different types of neutral models. Similarly, we used the same sampling parameters (length and resolution of the circular transects) and the same type of data analyses because the scaling conventions of the sampling design itself could potentially introduce additional constraints on the detected frequency of species combinations.

Spatial associations between the simulated species are shown in Fig. 4. Recall that negative interspecific association is expected in case of Pattern 1 but positive interspecific association in case of Pattern 4 while species should be independent in Pattern 2 and 3. Results based on the neutral model of random shifts fit better to these

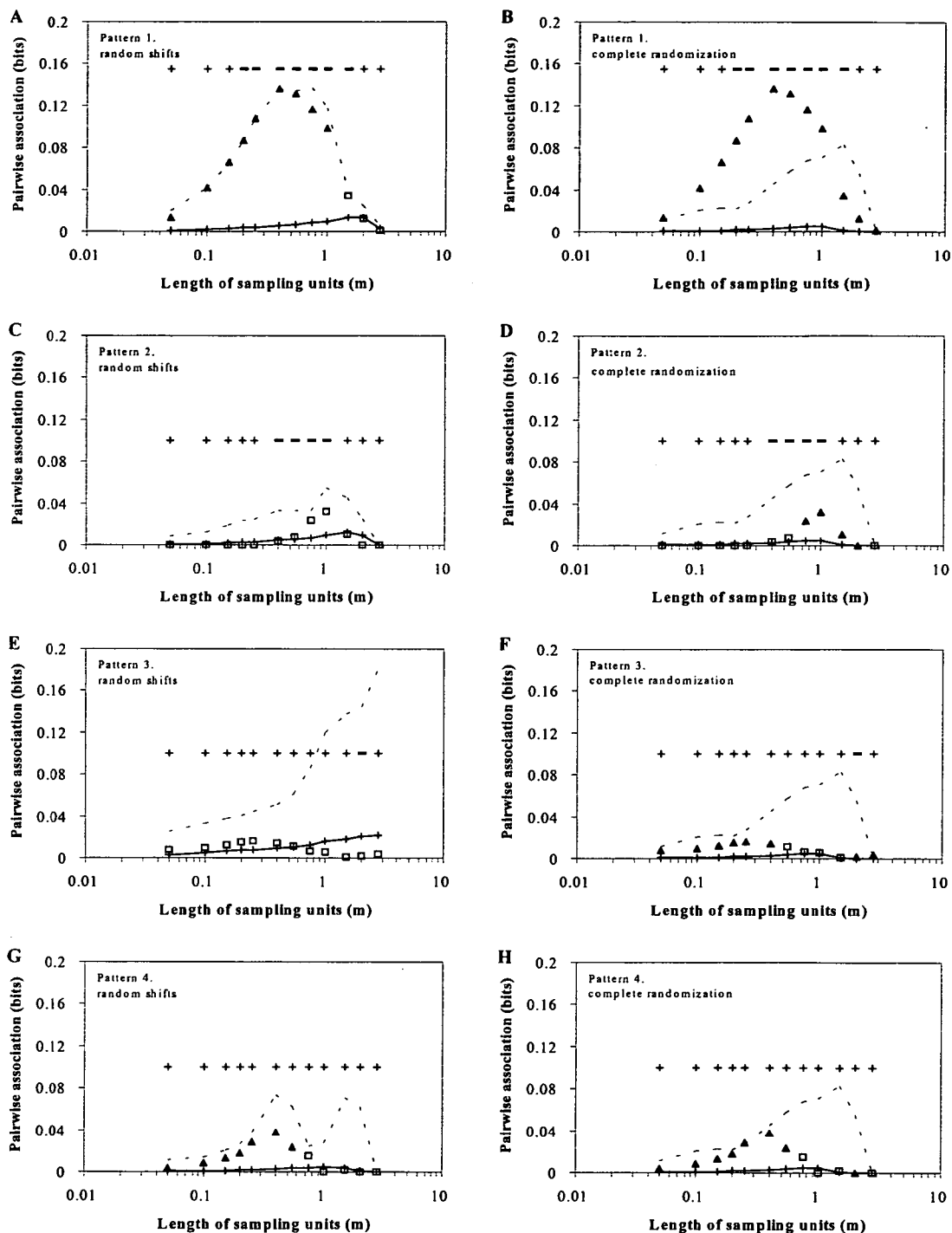


Fig. 4. Pairwise spatial associations between the simulated species. The signs of the associations (+,-) were determined by the comparison of the sums of diagonals ($a+d$ vs. $b+c$) of the 2×2 contingency table. ▲ (filled triangle) significant departure of an observed value from the neutral model ($p < 0.05$), □ (empty square) non-significant observed values, + average of the random references based on 5000 Monte-Carlo randomizations, ----- maximum of the random references, (minimum of the random references is always 0).

Table 2. Type of the patterns of individual species (intraspecific spatial dependence). Significance was based on 5000 complete randomizations.

Simulation	Species	Status of the species	Type of pattern $p < 0.01$	Intraspecific spatial dependence
PAT1	Species 1	Dominant	aggregated	yes
PAT1	Species 2	Subordinated	random	no
PAT2	Species 1	Dominant	aggregated	yes
PAT2	Species 2	Co-dominant	random	no
PAT3	Species 1	Dominant	aggregated	yes
PAT3	Species 2	Co-dominant	aggregated	yes
PAT4	Species 2	Subordinated 1.	random	no
PAT4	Species 3	Subordinated 2.	random	no

expectations. There are no significant spatial dependence in Pattern 2 and 3 (Fig. 4c and Fig. 4e) while significant positive associations are found in Pattern 4 (Fig. 4g). There are positive associations in Pattern 1 at smaller plot sizes and negative associations at medium plot sizes (Fig. 4a). The majority of results corresponds to the assembly rules of the simulations, however the positive associations found at fine scales in Pattern 1 contradict to our expectation. The results based on the complete randomization are more complex and more difficult to interpret. For Pattern 4, the results are the same as the results by the random shifts, even the random envelopes (minimum and maximum values of randomizations) produced by the two types of neutral models are different (Fig. 4h). However, significant associations are found in all the other simulated patterns and the signs are both positive and negative, depending on the scale (Fig. 4b, Fig. 4d, and Fig. 4f). Difference of the results calculated with the two types of neutral models can be attributed to the fact that complete randomization also refers to the intra-specific spatial dependence (autocorrelation) of species. As it was shown in Fig. 1, the dominant species was patchy (aggregated) in all patterns, the subordinate species was patchy only in Pattern 3. In the other cases, the subordinate species was distributed by unlimited dispersal, which could produce either random or slightly segregated patterns, depending on the competitive effect of dominant.

To test the above expectations about the individualistic patterns (i.e. intraspecific spatial dependence) of species, we calculated the probability of finding a species in a quadrat as a function of quadrat sizes. If a species is aggregated, it appears in the sample with lower probability than in case of randomness (Fig. 5). Table 2 summarizes the results. These correspond to our expectations

according to the assembly rules of the simulation. In the cases when competition occurs, segregation of the subordinate species from the dominant do not produce patchy distribution, because the frequencies of the subordinate species are low (12.5 %). However, limited dispersal always produces significant aggregations. Comparing the significant associations detected by the two neutral models, there are additional significant associations in the results obtained by complete randomizations. If only one of the species is aggregated (Fig.4b and Fig.4d), then the additional significant associations appear only at coarse scales. But, if both species are aggregated, the associations are significant at almost all scales (Fig. 4f). It is clear that intra-specific spatial dependence changes the probability of the species to appear in a sampling unit (cf. Fig. 5). However, the consequence of this change appears in the probability of co-occurrence of species, too. If at least one of the species becomes aggregated, the probability of co-occurrences will decrease, as it was shown by applying the neutral model of complete randomization.

Are the significant associations artefacts in Fig. 4d and Fig. 4f? We think they are not. We extend the concept of association to the intraspecific level, thus, the individuals are associated if their probability to appear in a sampling unit are significantly different from a random expectation. We emphasize that association between individuals changes their interactions. A second question is that how individuals can be classified into groups (e.g. taxons or functional types etc.). To solve this task we can separate the within-, and between-group associations with appropriate neutral models.

In Fig. 4 we depict the signs of $I(A,B)$ at all scales, disregarding if they show significant associations or not. Signs were determined by comparing the $a+d$ vs $b+c$ sums calculated from the

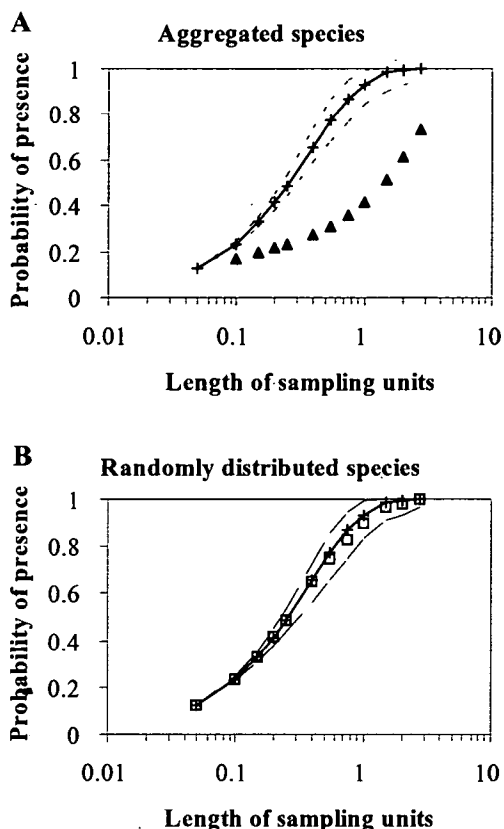


Fig.5. Detection of the spatial pattern of individuals species. An example with the two simulated species in Pattern 1. A, aggregated pattern of dominant species, b, random pattern of subordinated species. \blacktriangle (filled triangle) significant departure of an observed value from the neutral model ($p < 0.01$), \square (empty square) non-significant observed values, + average of the random references based on 5000 Monte-Carlo randomizations, ----- maximum and minimum of the random references.

diagonals of the 2×2 contingency table of observed data. The signs follow the same positive to negative to positive pattern with the increasing plot sizes in each simulated transects. This pattern is similar to the one found in the field data (see Table 1), and might be simply a direct consequence of increasing the plot sizes in the spatial series. We expect that $a+d > b+c$ at very small plot sizes, because d is very frequent, i.e. there are many empty plots. At large plot sizes, close to the minimum area of the two-species assemblage, a is very frequent, thus $a+d > b+c$ again. In the contrary, at intermediate plot sizes, $b+c$ might be greater than $a+d$. The intervals of scales where signs change vary from pattern to pattern, but the logic is the same. This pattern of positive to negative to positive signs is often invisible, because significant associations may appear only in a limited interval of scales. The

effects of changing the frequencies of species combinations with increasing plot sizes might mask the "real" trends of positive and negative associations. Textural effects might override the real tendencies of species to co-occur. It may result in seemingly illogical patterns of the signs of associations (e.g. Fig. 4a) that are difficult to interpret, or at least, there is a danger of misinterpretation.

To avoid these effects, we suggest a second type of algorithm for determining the signs of the associations. Here the observed sums of diagonals were compared to the average sums of diagonals calculated from the Monte-Carlo random references (neutral models). Table 3 shows the signs of significant associations detected by this method. This result is easier to interpret than the result obtained by the first method. The positive signs disappeared from the simulated Pattern 1. Using random shift as a neutral model, Pattern 4 shows only positive associations, which corresponds to our expectation. There is a change of the signs from positive to negative in case of complete randomization, when within-species spatial dependence is also detected. It is reasonable, because aggregation in the original species pattern reduces the probability of finding a species in a sampling unit. For example, if both species are aggregated but they are independent from each other (Pattern 3), the relative frequency in a (both species are present) will be smaller, the frequency in d (both are absent) will be higher than the random expectation created by complete randomization (Fig. 6a and Fig. 6d). Relative frequencies in b and c (only one species is present) are lower than the random expectation at finer scale and then become higher at coarser scale (Fig. 6b and Fig. 6c). All these effects together result in the positive to negative change of the signs of the significant associations.

Recalculating the signs of associations with the second algorithm, we found more reasonable and less controversial patterns in case of the field data too (Table 4). When we apply random shift as neutral model, there are either only positive or only negative associations along the spatial series. Applying complete randomization, there is the same positive to negative change of the signs as found in the case of simulated pattern. Exceptionally, negative to positive to negative pattern appears if there are other components of associations beyond the intraspecific. Applying random shifts as neutral models significant associations show a coalition structure. There are negative associations between the dominant grasses *Festuca vaginata* and *Stipa borysthenica*, as well as between the matrix-forming *Festuca* and the gap-species *Poa bulbosa*. Annuals

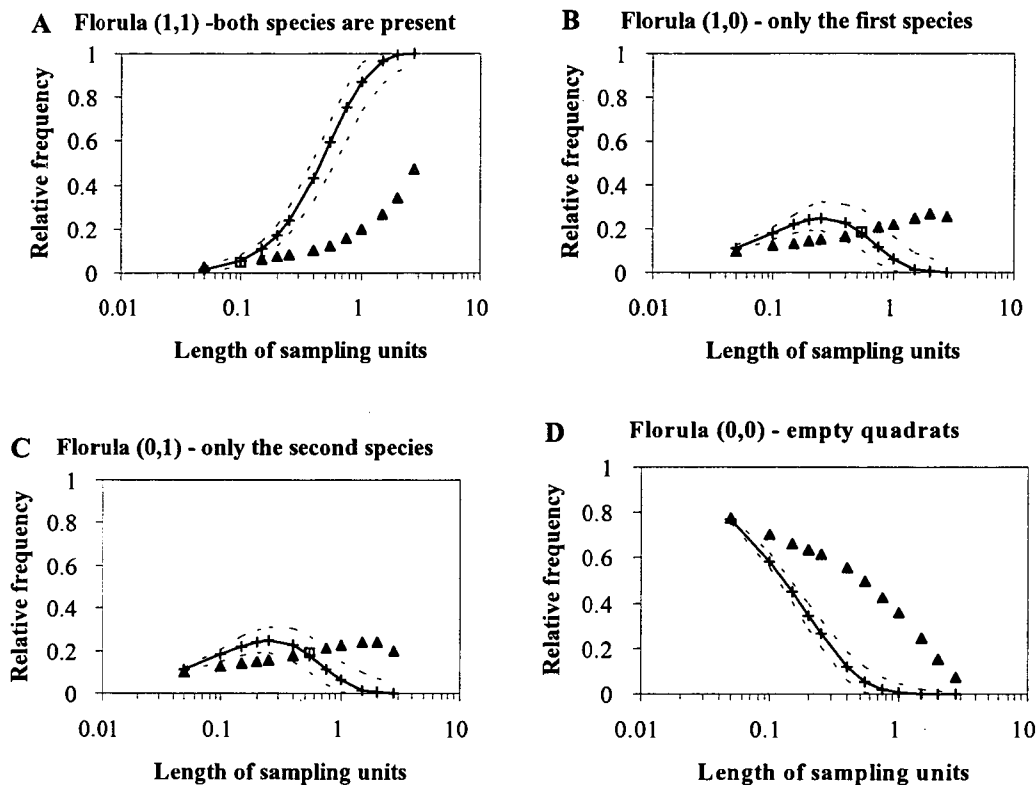


Fig. 6. The effect of within-species spatial dependence (autocorrelation) on the spatial patterns of species combinations. An example with Pattern 3, where both species are aggregated due to their limited dispersal, however, there are no interaction between them. \blacktriangle (filled triangle) significant departure of an observed value from the neutral model ($p < 0.05$), \square (empty square) non-significant observed values, + average of the neutral model (5000 complete randomizations), ----- maximum and minimum of the random references.

Table 3. Patterns of the signs of the pairwise spatial associations of the species in the simulated patterns. The signs of the associations (+, -) were determined by the comparison of the observed vs. the average random sums of diagonals of the 2×2 contingency table.

critical value for significance $p < 0.05$

. non-significant association

+1 positive association

-1 negative association

neutral model: complete randomization with fixed frequencies at the finest resolution

length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275
simulations												
Pattern 1		-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
Pattern 2		-1	-1	-1	-1
Pattern 3		+1	+1	+1	+1	+1	+1	.	.	.	-1	-1
Pattern 4		+1	+1	+1	+1	+1	+1	+1	.	.	.	-1

neutral model: random shift of species along the circular transect

length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275
simulations												
Pattern 1		-1	-1	-1	-1	-1	-1	-1	-1	.	.	.
Pattern 2	
Pattern 3	
Pattern 4		+1	+1	+1	+1	+1	+1	+1

Table 4. Patterns of the signs of the pairwise spatial associations of the the seven most frequent species in a xeric perennial sand grassland in Csévharaszt. The signs of the associations (+,-) were determined by the comparison of the observed vs. the average random sums of diagonals of the 2x2 contingency table.

critical value for significance $p < 0.01$
 . non-significant association
 +1 positive association
 -1 negative association

neutral model: complete randomization with fixed frequencies at the finest resolution

length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275	495	1000
species pairs														
CLAMAG x CLAFUR	.	+1	+1	-1	-1	-1	-1	-1	-1	-1	-1	.	.	.
CLAMAG x FESVAG	.	+1	+1	-1	-1	-1	-1	-1	-1	-1	-1	.	.	.
CLAMAG x STIBOR	-1	-1	-1	-1	.	.	.
CLAMAG x POLARE	.	.	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	.	.
CLAMAG x POABUL	-1	+1	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	.	.
CLAMAG x VIOARV	-1	+1	+1	+1	+1	-1	-1	-1	-1	-1	-1	-1	.	.
CLAFUR x FESVAG	-1	-1	-1	.	.
CLAFUR x STIBOR	-1	-1	-1	.	.
CLAFUR x POLARE	-1	-1	-1	.	.
CLAFUR x POABUL	-1	+1	+1	+1	+1	-1	-1	-1	-1	-1	-1	-1	.	.
CLAFUR x VIOARV	-1	.	.
FESVAG x STIBOR	.	-1	-1	-1	.	-1	.	.	.	-1	-1	.	.	.
FESVAG x POLARE	-1	-1	.	.	.
FESVAG x POABUL	-1	+1	+1	+1	.	.	-1	-1	-1	-1	-1	-1	.	.
FESVAG x VIOARV	-1	-1	.	.	.
STIBOR x POLARE	-1	-1	.	.	.
STIBOR x POABUL	-1	.	.	.
STIBOR x VIOARV	-1	-1	-1	.	.	.
POLARE x POABUL	.	.	+1	.	+1	+1	+1	+1	+1	-1	-1	-1	-1	-1
POLARE x VIOARV	.	.	+1	+1	-1	-1	-1	-1	.
POABUL x VIOARV	-1	.

neutral model: random shift of species along the circular transect

length of units (cm)	5	10	15	20	25	40	55	75	100	150	200	275	495	1000
species pairs														
CLAMAG x CLAFUR	+1
CLAMAG x FESVAG	.	.	+1	+1	+1
CLAMAG x STIBOR
CLAMAG x POLARE
CLAMAG x POABUL
CLAMAG x VIOARV	-1	-1	-1
CLAFUR x FESVAG
CLAFUR x STIBOR
CLAFUR x POLARE
CLAFUR x POABUL	-1	-1	-1
CLAFUR x VIOARV
FESVAG x STIBOR	.	.	-1
FESVAG x POLARE
FESVAG x POABUL	.	-1
FESVAG x VIOARV
STIBOR x POLARE
STIBOR x POABUL
STIBOR x VIOARV
POLARE x POABUL	+1
POLARE x VIOARV
POABUL x VIOARV

are positively associated to *Poa bulbosa*, while negatively associated to lichens. Lichens are positively associated to *Festuca*. The whole structure corresponds to our previous experiences about the ecophysiology, population biology, and larger scale coenological preferences of species in xeric perennial sand grasslands.

Conclusions

Pairwise spatial association of species is a measure to express the difference between frequency distributions of species combinations in an observed and in an ideal community. The traditionally used ideal distribution represents a community where all species can completely coexist, i.e. the probabilities of species combinations are determined only by the abundance of species.

There are several factors that might cause a deviation from this ideal case.

1, *Textural constraints* appear due to the fact that stands of plant communities and individual plants (or ramets) are *finite and discrete entities*. The sampling operations emphasize this finite and discrete character of natural communities when the limited extent and resolution of the sample reduce further the freedom of species combinations. Consequently, some rare species combinations may never appear in a limited, finite sample, while the combinations of the abundant species might be over-represented. Similarly, if the number of species is relatively high, the number of potential species combinations might exceed the size of the sample. In this case, the limited size of the sample puts constraints on the number and frequency of realized species combinations, that result in deviation from the ideal case. Important to emphasize that textural constraints are not simply sampling artefacts due to our limited sampling efforts. In nature, frequency of realized species combinations will also differ from an ideal (continuous and infinite) case due to natural scaling parameters (limited extent of vegetation stands and given natural scales defined by plant morphology) in a community.

Students of community structure usually pay little attention to textural constraints or they regard them as some unwelcome noise or bias from the real biological patterns and mechanisms. Considerable attempts have been made to remove these effects of the textural constraints from the analyses to get really „pure, biologically reasonable” results. However, because-textural constraints effect species interactions and community dynamics as well, neglecting or disregarding them might cause serious biases in our understanding of community organization.

2, *Structural constraints* appear because of the spatial dependence between individuals. *Within-species spatial dependence*, i.e. non-random spatial distribution of individuals of the same species might appear due to population level mechanisms (e.g. limited dispersal of propagules, patterns of vegetative growth and reproduction or preference of individuals to certain patches of a heterogeneous habitat). *Inter-specific spatial dependence* can be the consequence of direct interaction of species (e.g. competition, parasitism, mutualisms) or due to indirect effect, i.e. the common reaction to other species or limiting environmental factors. In each case, spatial dependence changes the frequency of species combinations either directly, by changing the probability of the occurrence of a given species combination, or indirectly, by changing the individualistic patterns, i.e. the individualistic probability of the occurrence of a species in a sampling unit.

3, We use the term '*constraints*' instead of 'effects' or 'contributions' to emphasize the *dynamic aspects* and dynamic consequences. From a dynamic point of view, zero association corresponds to the conditions that are assumed in a mass-based equilibrium model (e.g. Lotka-Volterra models) (cf. Czárán, 1998). In any more realistic model of plant community dynamics, association means that the behaviour of the species differ from what we expect directly from their overall abundances. The associated species co-occur or interact more or less frequently due to several textural and structural effects that constrain their behaviour. We emphasize the need of developing a well articulated set of neutral models that help to separate the different type of biologically relevant constraints. As a first attempt, we used here two different types of neutral models in parallel.

4, We propose to use (1) the complete randomization model to separate textural and structural constraints, and (2) the random shifts to distinguish intra-, vs. interspecific spatial dependence within structural constraints. We also suggest that detecting signs of significant associations, we should compare the relative frequencies of relevant species combinations between the observed data and the actual neutral model. In all cases, the application of spatial series is important and inherent in the analyses.

Our simulation study proves that with these methods one can detect much details than with the traditionally used statistical method. These methods have the advantage of simplicity and easy use. However, neither complete randomization nor random shifts are perfect from a biological point of view. By artificial setting of the resolution of complete randomization we might cut a plant

individual into pieces and randomize along the transects. On the other hand, random shift keeps the individual pattern constant during the randomization involving the unrealistic assumption that pattern of one species was developing independently from the other species. The major problem lies in the tradition of developing static neutral models with statistical approaches disregarding that the majority of biological processes and mechanisms that are relevant only in dynamical context (Geoffrey M. Henebry, personal communication). Future neutral models for understanding plant community structure should be based on thorough knowledge about the dynamics of the morphogenesis of plants and the synmorphogenesis of vegetation.

Acknowledgements

We are grateful to Beáta Oborny and Gábor Fekete for helpful comments on the manuscript. This study was supported by the Hungarian Scientific Research Fund (OTKA T016225 to S. Bartha, and T021166 to M. Kertész).

References

- Agnew, A.D.Q. (1961): The ecology of *Juncus effusus* in North Wales. - *Journal of Ecology* 49,83-102.
- Bartha, S. (1983): Plexus gráfok térfolyamatban. - *Szakdolgozat (Kézirat)*, Vácrátót
- Bartha, S. (1992): Preliminary scaling for multi-species coalitions in primary succession. - *Abstracta Botanica* 16,31-41.
- Bartha, S., Czárán, T. and Oborny, B. (1995): Spatial constraints masking community assembly rules: a simulation study. - *Folia Geobot. Phytotax.*, Praha 30,471-482.
- Czárán, T. (1984): A simulation model for generating patterns of sessile populations. - *Abstracta Botanica* 8,1-13.
- Czárán, T. (1993): PATPRO: A Monte-Carlo simulation program for multispecies neighborhood competition. - *Abstracta Botanica* 17,275-281.
- Czárán, T. (1998): Spatiotemporal Models of Population and Community Dynamics. Chapman and Hall, New York
- Czárán, T. and Bartha, S. (1992): Spatiotemporal dynamics models of plant populations and communities. - *Trends in Ecology and Evolution* 7,38-42.
- Dale, M.R.T., Blundon, D.J., MacIsaac, D.A. and Thomas, A.G. (1991): Multiple species effects and spatial autocorrelation in detecting species associations. - *Journal of Vegetation Science* 2,635-642.
- de Vries, D.M. (1953): Objective combinations of species. - *Acta Bot. Neerl.* 1,497-499.
- Diggle, P.J. (1983): *Statistical Analysis of Spatial Point Patterns*. - Academic Press, London
- Gotelli, N.J. and Graves, G.R. (1996): *Null models in ecology*. - Smithsonian Institution Press, Washington
- Greig-Smith, P. (1983): *Quantitative Plant Ecology*. - 3rd ed. University of California Press, Berkeley, CA
- Fekete, G. and Szujkó-Lacza, J. (1973): Interspecific correlation of plants in oakwood at increasing block sizes. - *Acta Biol. Acad. Sci. Hung.* 24,31-42.
- Fekete, G. (1992): The holistic view of succession reconsidered. - *Coenoses* 7,21-29.
- Fekete, G., Tuba, Z. and Précsényi, I. (1995): Application in three approaches to evaluate abundance and rarity in a sand grassland community. - *Coenoses* 10,29-38.
- Hausser, M. (1993): Graphs and multigraphs depict the change of community structure by scale. - *Abstracta Botanica* 17,103-114.
- Horváth, A. (1998): INFOTHEM program: New possibilities of spatial series analysis based on information theory methods. - *Tiscia* 31, 71-84.
- Juhász-Nagy, P. (1967): On association among plant populations. Part 1. Multiple and partial association: a new approach. - *Acta Biol. Debr.* 5,43-56.
- Juhász-Nagy, P. (1980): A cönológia koegzisztenciális szerkezetének modellezése. Akad. Dokt. Ért. (Kézirat), 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.
- Kalapos, T. (1994): Homokpusztagyepi növények fotoszintézisének és vízforgalmának vizsgálata, különös tekintettel az időszakos szárazság hatására. (Studies on the photosynthesis and waterrelations of sand grassland species with particular attention on temporal water stress.) PhD thesis, Budapest.
- Kárpáti, I. and Kárpáti, V. (1954): The aspects of the calciphilous turf (*Festucetum vaginatae danubiale*) in the environs of Vácrátót in 1952. - *Acta Bot. Acad. Sci. Hung.* 1,129-157.
- Kershaw, K.A. (1964): *Quantitative and dynamic ecology*. - E. Arnold, London.
- Kovács-Láng E. (1974): Examination of dynamics of organic matter in a perennial open sandy steppe-meadow (*Festucetum vaginatae danubiale*) at the Csévharaszt IBP sample area (Hungary). - *Acta Bot. Acad. Sci. Hung.* 20,309-326.
- Lepš, J. and Buriánek, V. (1990): Interspecific associations in old-field succession. - In F. Krahulec, A.D.Q. Agnew, and J.H. Willems (eds): *Spatial processes in plant communities*. pp. 13-22. Academia Press, Prague.
- Margóczy, K. (1995): Interspecific associations in different stages of the vegetation in a Hungarian sandy area. - *Tiscia* 29, 19-26.
- Matus, G. and Tóthmérész, B. (1990): The effect of grazing on the structure of a sandy grassland. - In F. Krahulec, A.D.Q. Agnew, and J.H. Willems (eds): *Spatial processes in plant communities*. pp. 23-30. Academia Press, Prague.
- Palmer, M.W. and van der Maarel, E. (1995): Variance in species richness, species association, and niche limitation. - *Oikos* 73, 203-213.
- Podani, J. (1984): Analysis of mapped and simulated vegetation patterns by means of computerized sampling techniques. - *Acta Bot. Acad. Sci. Hung.* 30,403-425.
- Podani, J., Czárán, T. (1997): Individual-centered analysis of mapped point patterns representing multi-species assemblages. - *Journal of Vegetation Science* 8,259-270.
- Tuba, Z., Csintalan, Zs., Nagy, Z., Szente, K., Kemény, G., Takács, Z., Koch, J., Badacsonyi, A., Murakeözy, P., Palicz, G., Kóbor, Sz., Ötvös, E., Bartha, S. 1998. Szünfiziológia: alapozó gondolatok és exploratív vizsgálatok egy születő növényökológiai tudományterülethez. In Fekete, G. (szerk.): *A közösségi ökológia frontvonalai*. — Scientia Kiadó, Budapest, pp. 171-196.
- Watkins, A.J. and Wilson, J.B. (1992): Fine-scale community structure of lawns. - *Journal of Ecology* 80,15-24.
- Wilson, J.B. (1994): Who makes the assembly rules? - *Journal of Vegetation Science* 5,275-278.
- Wilson, J.B., Gitay, H. and Agnew, A.D.Q. (1987): Does niche limitation exist? - *Functional Ecology* 1,391-397.
- Wilson, J.B. and Watkins, A.J. (1994): Guilds and assembly rules in lawn communities. - *Journal of Vegetation Science* 5,591-600.
- Yarranton, G.A. (1966): A plotless method of sampling vegetation. - *Journal of Ecology* 54,229-237.