

Exploring plant community dynamics in abstract coenostate spaces

S. Bartha¹, T. Czárán² & J. Podani³

1 Institute of Ecology and Botany of the Hungarian Academy of Sciences, H-2163 Vácrátót, Hungary. Email: sanyi@botanika.botanika.hu

2 Ecological Modelling Research Group of the Hungarian Academy of Sciences and Eötvös University, Ludovika tér 2, H-1083 Budapest, Hungary. Email: czaran@ludens.elte.hu

3 Department of Plant Taxonomy and Ecology, Eötvös University, Ludovika tér 2, H-1083 Budapest, Hungary. Email: podani@ludens.elte.hu

Keywords: Chronoserries, Diversity, Information theory, Quadrat size, Scale, Simulations, Spatiotemporal pattern, Species associations, Succession.

Abstract: The paper addresses a fundamental issue of plant ecology: how to find the links between the analysis of spatial pattern and community dynamics. It is demonstrated that temporal processes can be effectively illustrated in abstract spaces with coenostate descriptors as dimensions. If these descriptors are based on Juhász-Nagy's information theory methodology, then a new approach results which treats spatial and temporal aspects, spatial dependence (species associations) and spatial variability (diversity) simultaneously, providing a promising tool at the hand of plant ecologists. The new method is illustrated by actual and simulated data.

Abbreviations: AS - associatum, CA - characteristic area, FD - florula diversity, LD - local distinctiveness.

I. Introduction

A major objective of the extensive and sophisticated methodology of quantitative plant ecology is *simplification* and *summarization*. This implies that the data reflecting complex spatiotemporal phenomena of vegetation are brought into an easily comprehensible, yet meaningful format. The methods designed to achieve this universal goal range from simple statistical functions to multivariate analytical procedures, with results ranging from single numbers to elaborated diagrams such as biplots, consensus dendrograms and stereograms. The need for this information concentration stems in part from the limitations of the human brain: relationships among many objects are difficult, if not impossible to handle simultaneously without external help. For example, the spatial pattern of many individuals belonging to different species in a meadow involves too much information for us to process in 'real time'. Appropriate, problem-oriented, and almost inevitably computerized statistical meth-

ods are required to reduce the dimensionality of the system to a conceivable level. Furthermore, whilst experienced surveyors always develop for themselves a 'Gestalt'-like picture of the subject matter, verbal communication to others is difficult and insufficient in most cases. Again, some statistical methodology is necessary to forward our knowledge to the scientific audience in a concise manner.

The different stages of quantitative vegetation surveys are implicitly associated with a methodological sequence of *real* and *conceptual spaces* (Podani 1984a, 1992). Sampling is our main activity in the real → conceptual conversion through which data are gathered. In studies of community pattern, data express species performance of some sort in sampling units, thus giving the primary basis, the *data* space (e.g., species space with species as dimensions and sample plots as points) for further computations. The simplest statistical procedures provide variables that are *derived* from those observed and recorded in the field. There is an

impressive arsenal of parametric and non-parametric methods available to obtain such derived variables, ranging from simple scalar quantities to vector or matrix representations (e.g., Gauch 1982, Green 1979, Legendre & Legendre 1983, Orlóci 1978). Trivial examples are stand-level average cover values, higher moments of biomass or density distributions of the species examined, species/individual diversity, interspecific correlation and so on. Each of these statistics focuses upon a selected aspect of the actual *coenological state* ('coenostate') of the given community, site by site or over the entire study area. The *scalar* measures (parametric and non-parametric alike) may be understood as *coenostate descriptors* or *coenostate variables*, so that a new abstract space is introduced implicitly. A great number of such descriptors can be derived from the same data set to characterize the coenological state of the community (c.f. Feoli & Orlóci 1991). In the *coenostate-space*, with the relevant coenostate variables as dimensions, a community stand or a relevé is represented by a point. Although, in the most general sense, the species abundances themselves could be regarded as simple coenostate variables, in which case the dimensionality of the coenostate space would equal the number of species in the community, we shall restrict our definition to derived variables. This is reasonable because the species space by itself (if considered as a coenostate space) and raw data, in general, do not imply summarization nor simplification.

Since we are concerned with spatial aspects, the notion of *scale* cannot be neglected. To be more straight to the point: *it is absolutely natural* that scale matters in both real and conceptual spaces, and a vast amount of evidence has accumulated in the literature of plant ecology (e.g., Greig-Smith 1983, Schneider 1992, Wiens 1989) to support this view¹. Therefore, derived variables that are determined at different scales (e.g., pairwise correlations of species along with increases of sampling unit size, Kershaw & Looney 1985) should deserve the greatest attention in studies of vegetation pattern. The primary aim of this paper is to describe some possibilities for such a coenostate-space representation of community structure. The procedures advocated are quite general, yet simple enough for the practitioner, and concise enough for efficient graphical illustration. We emphasize that their application is quite straightforward to explore the links between spatial and dynamical aspects of community organization.

Coenostate descriptors are extensively used in two, frequently isolated areas of ecological research: a) syntaxonomic classification of vegetation at the community or landscape level ('phytosociology', e.g., Braun-Blanquet 1964, Mueller-Dombois & Ellenberg 1974, Mucina et al. 1993), and b) synmorphological-syndynamical studies to evaluate supraindividual organization (Agnew et al. 1993, van der Maarel & Sykes 1993, Zobel & Zobel 1988). These approaches differ in their attitude towards the subject, especially in goals and methods. Syntaxonomic problems are usually solved by finding the most relevant original variables (i.e., species) that provide the best description of, and the strongest discrimination among, syntaxa. Once group description/separation is judged to be satisfactory, the other dimensions of the species-space become less important or even negligible. Coenostate variables are derived only subsequently, when the classification is completed. In synmorphology and syndynamics, on the other hand, the coenostate variables need to be chosen beforehand to ensure that the maximum amount of information is extracted from the data to reveal both structural and functional aspects and to minimize subjectivity in the analysis.

The latter requirement is relatively strict, but it does not necessarily call for the simultaneous use of *many* coenostate variables: there is no point in retaining the high dimensionality of the spaces. What it does require is a careful choice of variables that are essential in the given context. If one wishes, for example, to explore the effect of spatial pattern (structure) on the coexistence of the species that constitute the community (function), then the variables representing the pattern should be dynamically relevant features, such as the spatial proximity of individuals or species combinations appearing in small areas. This choice can only be appropriate when based on a solid and coherent theory that offers a complete formal description of coenological states at any spatial scale. This theory must also open the way for a direct syndynamical reference in order to link coenological structure to community dynamics. We believe that there is at least one such approach, the information theoretical methodology introduced by Juhász-Nagy (1967, 1980, 1984, 1993a, see also Juhász-Nagy & Podani 1983). In this paper, we propose a two-dimensional coenostate-space representation of vegetational entities based on Juhász-Nagy's propositions in order to detect pattern-related coenological phenomena and to infer evidence on their

¹ Nevertheless, an even vaster amount of work appears to ignore the issue completely.

dynamical consequences. We shall have the following specific suggestions in mind:

- Derived variables *florula diversity* and *associatum* (Subsections II.1-2) are effective coenostate descriptors of community structure. These are the two most important information statistical functions suggested by Juhász-Nagy, characterizing *spatial variability* and *spatial dependence*, respectively. The inherent property of these functions is that they are relevant only if determined for a series of increasing sampling unit sizes within a community stand.
- We propose taking the *maximum values* of these two functions for each community stand as further coenostate descriptors in a new, two-dimensional abstract space. Coenostate dynamics, i.e., spatiotemporal processes, can be visualized in this space by '*coenostate-trajectories*'. Although spatial variability and spatial dependence are not separable concepts, their relationships are far from being trivial. Depending on the actual dynamics that produced the community pattern, florula diversity and associatum may be positively or negatively correlated or even uncorrelated. Furthermore, as we shall see later, boundary relationships may be even more important than direct interactions of these variables.
- We suggest that possible events in the florula diversity/associatum state space are best studied using simulated community patterns produced via spatially explicit, individual-based models. Experience gained from such computerized experiments can then facilitate the evaluation of field data.

II. Juhász-Nagy's information theory methods

In the context of pattern analysis, two fundamental questions deserve particular attention:

- The analysis of species coexistence in a community stand; capturing coexistence relations through species combinations and measuring the influence of scale upon these relations (which have to do with *spatial diversity* or *variability*).
- The measurement of the deviation of coexistence relationships of plant populations from the ran-

dom expectation at different spatial scales (a case of *spatial dependence*).

It very often happens that either of the above issues is emphasized and the other remains forgotten. An optimal community study, however, considers both of them by defining appropriate coenostate variables. If the first axis in the coenostate space relates to spatial diversity, and the other axis reflects dependence, then we obtain a graphical tool optimally suited to the problem. In this abstract space, a community stand sampled at a given time is represented by a point, and different processes of community dynamics, for example successional changes, may be depicted by connecting temporally adjacent points, i.e., *trajectories*. Obviously, it is most crucial how the two coenostate variables are defined.

Juhász-Nagy (1967, 1980, 1993a) was the first to develop a coherent conceptual framework to address these two fundamental issues simultaneously. He proposed a family of coenostate variables to measure various aspects of spatial diversity and dependence in plant communities. In addition, he was among the first to recognize the extreme importance of spatial scale in vegetation studies. In his information statistical approach, the change of scale, realized as a series of sampling unit sizes (see Subsections II.1-2), is an integral part – without considering scale effects these functions are meaningless. Although the mathematical details have been described in several papers in the literature, we feel that some repetitions are necessary before we put Juhász-Nagy's methodology into a new context.

II.1. Florula diversity: a measure of spatial variation

One of Juhász-Nagy's major innovation is the concept of '*florula diversity*' or '*compositional diversity*' (Juhász-Nagy 1993a, Juhász-Nagy & Podani 1983), a function referring to one of the three basic aspects of coenological behaviour (the other two will be formalized in Subsection II.2), namely *coenological diversity*. Assume that the community in question is composed of s species, A, B, ..., S, and is sampled with plots of size j . The species composition, that is the realized combination of species (a small flora, i.e., *florula*) is recorded in each of the m sampling units. The essence of the approach is that we count the number of times a given species combination occurs in the sample². Let f_{kj} be the number of occurrences of the k th species combination and $p_{kj} = f_{kj}/m$ is the relative fre-

² Terms *sample* and *sample plot* are often confused in the literature. In accordance with the statistical usage, sample is understood here as a collection of all sampling units (sample plots) taken in the community.

quency of this combination. $z = 2^s$ is the number of possible combinations of s species (including the empty florula). Then, the florula diversity (FD) of the community at this scale point is estimated as the Shannon-entropy for the set of species combinations detected all over the study area:

$$FD_j = - \sum_{k=1}^z p_{kj} \log p_{kj} \quad (1)$$

FD is thus a special application of Shannon-diversity, otherwise extensively used to express species/individual diversity of communities (Pielou 1975, Magurran 1988). This time, however, it is defined on the frequency distribution of observed (realized) species combinations within the sample, rather than on the frequency distribution of species abundances themselves. FD is a scalar estimate of the florula diversity of the community pattern at sampling unit size j . The larger the sample size m (i.e., the number of sampling units per sample), the more accurate this estimate. By changing quadrat size j , florula diversity can be expressed as a function of spatial scale (i.e., calculated for samples recorded with a series of different sampling unit sizes). For very small sample plots, florula diversity tends to be very small, because the units are either empty, or contain only a single species. These combinations are not 'attractive' ecologically, and there is very little uncertainty regarding the contents of a random plot. Extremely large plot sizes, on the other hand, will tend to include all or almost all species. Again, such combinations are not very interesting, and the uncertainty is low as well. Between these extremes, florula diversity takes at least one maximum value; at a sample plot size where the most diverse species combinations were captured by the sampling procedure.

II.2. *Associatum: a measure of spatial dependence*

Whereas florula diversity is suitable to express overall spatial diversity in a community, it is not informative on interactions, that is, on the *spatial dependence* of the patterns of species. As usual in other approaches to the measurement of interspecific association, Juhász-Nagy's measure of spatial dependence will also rely upon the contingency table. Imagine, first, an idealized plant community of infinite spatial extent, with all plant individuals as dimensionless points dispersed over the two-dimensional, environmentally homogeneous plane ('habitat'). Assume, further, that this imaginary community is sampled by an infinite number of randomly allocated sampling units of finite (non-zero) size j . Under such circumstances, each of the 2^s possible species combinations may be realized

		species B		
		present	absent	
species A	present	$p(A,B)$	$p(A,\bar{B})$	$p(A)$
	absent	$p(\bar{A},B)$	$p(\bar{A},\bar{B})$	$p(\bar{A})$
		$p(B)$	$p(\bar{B})$	1

Figure 1. Probabilities in a 2×2 contingency table. A: species A is present in the plot, B: species B is present in the plot, \bar{A} : species A is absent from the plot, \bar{B} : species B is absent from the plot. Marginals: individualistic behaviour of species, inner cells: collective behaviour of species.

in a fraction of the sample (for certain combinations, this fraction might be 0). For each k , this fraction corresponds to the probability p_{kj} of hitting this particular combination within the infinite stand by a randomly allocated sampling unit of size j . These p_{kj} probabilities may be arranged in the cells of a 2^s contingency table (Fig. 1 gives an example for $s = 2$). The marginal totals of the table provide the probability of presence and absence of species within a random sampling unit of size j . For species A, these probabilities are denoted by $p(A)$ and $p(\bar{A})$, giving the marginal vector for that species. Note that the marginal vectors characterize only the 'individualistic behaviour' of the species within the sample. They are uninformative on the combinations a species contributes to, that is, on associations. From the marginal vector it is easy to calculate the Shannon-entropy of species A at quadrat size j :

$$H_j(A) = - p(A) \log p(A) - p(\bar{A}) \log p(\bar{A}) \quad (2)$$

which is a measure of the uncertainty as to the presence of species A in a randomly located plot. It is the maximum if species A is present in exactly half of the plots (for logarithms of base 2, $\max H(A)=1$). The sum of species entropies gives the pooled entropy of the species collection, termed by Juhász-Nagy as the *local distinctiveness (LD)* of the community stand:

$$LD_j = H_j(A) + H_j(B) + \dots + H_j(S) \quad (3)$$

LD is also called the *marginal entropy* of the 2^s contingency table. Juhász-Nagy (1976) used local distinctiveness as a measure of *local preference*, another basic aspect of coenological behaviour. It refers to preference in two different ways: a species shows

preference for the habitat as a whole if it is present in many quadrats, and it might prefer parts of the habitat only by being present in those parts and absent elsewhere. It is obvious that *LD* is low at both small and large sampling unit sizes, i.e., its change over spatial scales is similar to that of *FD*.

Suppose that the imaginary 'community' is generated such that each individual takes a random position within the habitat. That is, the resulting plant patterns are *spatially independent*. Independence ensures that the probabilities in the cells of the contingency table can be obtained by multiplying the corresponding marginal values. In the example of Fig. 1, the probability $p(A,B)$ that two species A and B will co-occur in a random quadrat can be directly calculated by multiplying the marginal probabilities of presence, $p(A)$ and $p(B)$:

$$p(A,B) = p(A) p(B).$$

Likewise, the probability that a random quadrat will lack both species can be calculated from the probabilities of absence:

$$p(\bar{A},\bar{B}) = p(\bar{A}) p(\bar{B}).$$

For the cases with either species present we have

$$p(A,\bar{B}) = p(A) p(\bar{B}), \text{ and}$$

$$p(\bar{A},B) = p(\bar{A}) p(B).$$

Given the above relationships, with some rearrangements of the formulae we find that the Shannon (joint) entropy for the inner cells (florula diversity) and the entropy for the marginal vectors (local distinctiveness) are equal:

$$LD - FD = 0.$$

That is, in a random community the probabilities of species presence and species absence will provide directly the probabilities of species combinations. Therefore, the expectation of florula diversity in a random community equals the pooled entropy of species. If randomness is not satisfied because of spatial aggregation or segregation among the individuals of species, the above equality will no longer hold. As a consequence of spatial dependence, many of the potential species combinations will not appear in the sample, whereas others will be encountered more frequently than expected under complete spatial randomness. As a result, florula diversity becomes smaller than local distinctiveness. The deviation of the pattern from spatial randomness can therefore be measured conveniently by the difference between *LD* and *FD*:

$$AS_j = LD_j - FD_j \quad (4)$$

This is the overall association for the collection of *s* species, termed the *associatum* (*AS*) by Juhász-Nagy (1980). Mathematically, it is the information of the 2^s contingency table, or mutual information of *s* binary variables. In the general sense, *AS* reflects the spatial *similarity* and *dissimilarity* structure of the community, another aspect of coenological behaviour.

Experience suggests that *AS* also has at least one maximum for a sufficiently wide range of plot sizes in actual communities, indicating the scale point where interspecific dependence is the highest. However, the explanation behind positive associatum is manifold. If an infinitely large sample is taken from a community of infinite spatial extension, *AS* will be an unbiased measure of spatial dependence. Deviation from a random reference may be caused by a multitude of ecological factors, such as intra- and interspecific competition, habitat heterogeneity or dispersal limitations. Actual community stands have finite spatial extension, however, introducing additional constraints on the local coexistence of species. For example, the individuals of rare species almost always have rare ('accidental') species combinations around them. In a small sample with non-overlapping units, the plot that includes the rare species will overrepresent a rare species combination. We cannot eliminate this effect by allowing overlaps of sample plots, because the same rare individual (and its neighbourhood) will appear over and over again. Due to these effects, the associatum in randomly generated artificial communities is more or less higher than zero, especially if the number of rare species is high (Podani 1984b, Szollát & Bartha 1991).

II.3 Maximum values, characteristic areas and orders

As mentioned above, all measures are meaningless if available for a single plot size only: they have to be determined over a sufficiently wide range of sizes. The change of these functions over plot size can be illustrated graphically. The shape of the curve and the location of minima and maxima, as determined by plant pattern, are characteristic of the spatial variation of the community under study (Juhász-Nagy & Podani 1983). The maxima and the scale points (plot sizes) where the maxima were reached carry the information most relevant to us.

The actual maximum values of the functions are abbreviated as FD_{max} , LD_{max} and AS_{max} , respectively. Juhász-Nagy (1967, 1980, 1993b) proposed using these maxima to simplify representation of community pattern and community dynamics. We shall further explore these possibilities in Section IV.

The scale points where the maxima occur are termed *characteristic areas* or, more specifically, maximum areas of florula diversity, local distinctiveness and associatum, and will be denoted by CA_{FD} , CA_{LD} and CA_{AS} , respectively. Using several examples, Juhász-Nagy & Podani (1983) pointed out that the order of these values may be as meaningful as the statistics themselves, and this order changes in a typical manner over succession. If the sizes of characteristic areas are used as special coenostate variables, we can make more careful statements than when simple orderings are examined, as we shall see in Section V.

III. State-space representations of vegetation structure

We are seeking a method to visualize vegetation structure which

- 1) refers to the three fundamental aspects of coenological relations (preference, diversity and similarity; see Section II.);
- 2) is based on spatial scaling due to the inherent scale-dependence of coenological phenomena (without doing so, coexistence relations cannot be detected, Juhász-Nagy 1967, 1980, 1984, Podani et al. 1993);
- 3) is useful to characterize community level patterns; but at the same time
- 4) is decomposable into coalition- and population-level patterns as well;
- 5) produces results in the form of relatively simple two-dimensional graphs for easy communication of results; and finally
- 6) represents temporal aspects, i.e., transformations of coenological states over time.

Here we give a few arguments showing that Juhász-Nagy's methodology conforms these criteria and they apply to state-space representations of plant community dynamics (cf. Juhász-Nagy 1993b):

- Ad 1) The three basic aspects of coenological behaviour, i.e., 'preference', 'diversity' and 'similarity' can be expressed quantitatively by local distinctiveness (LD), florula diversity (FD) and associatum (AS), respectively (Juhász-Nagy 1980), that is, by three arithmetically related functions.
- Ad 2) These information statistical measures are determined along the scale of sampling unit size (j); and their maxima are used to characterize the overall pattern of communities.

Ad 3) Preference, diversity and similarity can be quantified at the community level if the information statistics are applied to the complete species list, but

Ad 4) any subset of the species list can also be specified, and the same functions can be calculated for the reduced species set (i.e., to any species coalition). Calculations with the subsets are extremely simple algebraically, because community level statistics can be decomposed into additive components.

Ad 5) Since the three basic aspects of coenological structure are canonical through the $AS = LD - FD$ relation (see II.2), two of them (in our case, FD and AS) are sufficient to characterize coenological states. A practical consequence is that the state space to be used is two-dimensional and is easily visualized. (The question whether the maxima occur at the same scale or not is examined later.)

Ad 6) The characteristic values of the coenostate-descriptor functions can be depicted diagrammatically for a series of permanent plots from the same community or for a chronosequence assumed to represent various stages of the same successional sere.

III.1. An example

The construction of state-space and its interpretation are shown on the example of Fig. 2 for FD and AS in a succession study. A data set obtained for primary succession processes on a strip coal mine in northern Hungary is used. By removing the top-soil and covering large areas with a mineral substrate, open-cast mining activity creates very suitable conditions for a primary succession of pioneer vegetation (Szegei et al. 1988). Micromaps of species distributions were made on such 'soil' dumps of a mining area near Visonta. Vegetation succession starts with an approximately random pattern of the first colonizing individuals of *Tussilago farfara*, *Lactuca serriola* and *Matricaria inodora*, to mention just a few names. After 3-4 years, the vegetation develops into a patchwork of annual grasses (e.g., *Bromus japonicus*) and short-lived perennial herbs (mostly legumes, e.g., *Medicago lupulina* and *Melilotus officinalis*). After 10 years, the area is covered by a closed, homogeneous grassland dominated by *Agropyron repens* (for more details, see Bartha 1990, 1992, 1993). Six stands of age 1, 2, 3, 4, 7 and 10 years were used as a chronoserics and were mapped for the topographic positions of individuals. From the micromaps thus obtained, computerized sampling procedures (Podani 1987) produced a series

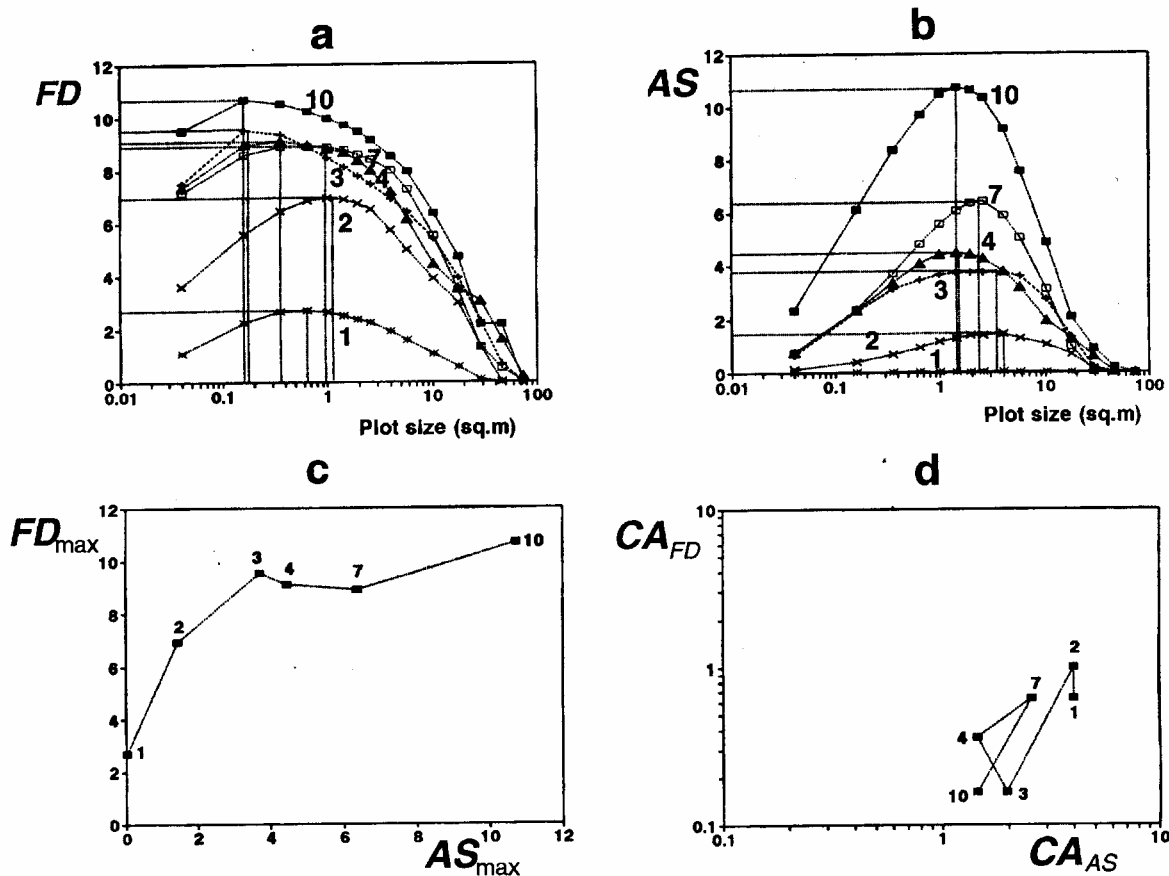


Figure 2. Creating state-space representations based on the maximum values of florula diversity and associatum. An example with a primary successional data set collected in an open-cast coal mining area in Hungary. Numbers within each diagram refer to years of succession.

of 2^s contingency tables for each stand. These were analyzed to estimate florula diversity (FD) and associatum (AS) as functions of sampling unit size.

Juhász-Nagy's measures indicate clear changes of the coexistence relations over time. Both FD and AS tend to increase with age at any particular quadrat size (Figs 2a-b), and the characteristic areas also change as succession proceeds. Fig. 2c depicts the 'trajectory' of the community within the FD_{max} - AS_{max} state plane over the first 10 years of primary succession, whereas the corresponding trajectory for the plane of CA_{FD} - CA_{AS} is depicted in Fig. 2d. The graph in Fig. 2c indicates 'community development' in the sense that both diversity (FD) and spatial dependence (AS) increase over time. The trajectory for the characteristic areas is more complex, showing a shift from larger to smaller quadrat sizes due to increasing species density and vegetation cover and a tendency of increasing uniformity of vegetation pattern over time.

Vegetation succession is represented frequently as a bundle of trajectories in the s -dimensional species space (e.g., Austin 1977, Matthews 1979, Collins & Adams 1983, Christensen & Peet 1984, Hogeweg et al. 1985), which may also be understood as a simple type of state-space representation as mentioned earlier. On the axes of this space species abundances are measured, and one trajectory shows the changing abundances of species over time in a given quadrat. It is possible to calculate pairwise dissimilarities between quadrats, i.e. the convergence (or divergence) of species composition between sampling units during succession, but such an approach suffers from at least three disadvantages. The patterns thus detected are strongly dependent on sampling unit size, which is chosen arbitrarily. Only dissimilarity relations of quadrat pairs can be evaluated, and all compared quadrats should have a common flora. Communities of different species composition or at different spatio-temporal scales are impossible to compare in this way. This is why Juhász-Nagy's models rely upon the pattern of *coexistence relations* rather than on species

composition, allowing comparison of the spatial organization of distant communities with distinct floras. Communities of different continents or those of extremely different habitats and spatiotemporal scales (e.g., phytoplankton vs. terrestrial vegetation, cf. Padišák 1992, 1994, Reynolds 1995) can even be represented and compared within the same state-space. The use of scaled coenostate spaces warrants the comparability of very different community patterns and community dynamics, mainly because the patterns 'scale themselves': no arbitrary choices of sampling unit size are involved in the analysis. By now, numerous case studies (e.g., Juhász-Nagy 1980, Juhász-Nagy & Podani 1983, Szollát & Bartha 1991, Bartha 1990, 1992, 1993, Tóthmérész & Erdei 1992, Podani et al. 1993, Tóthmérész 1994b, Oborny 1994, Hochstrasser 1995, Bartha et al. 1995b, 1997) have demonstrated the applicability of Juhász-Nagy's models to terrestrial plant community patterns.

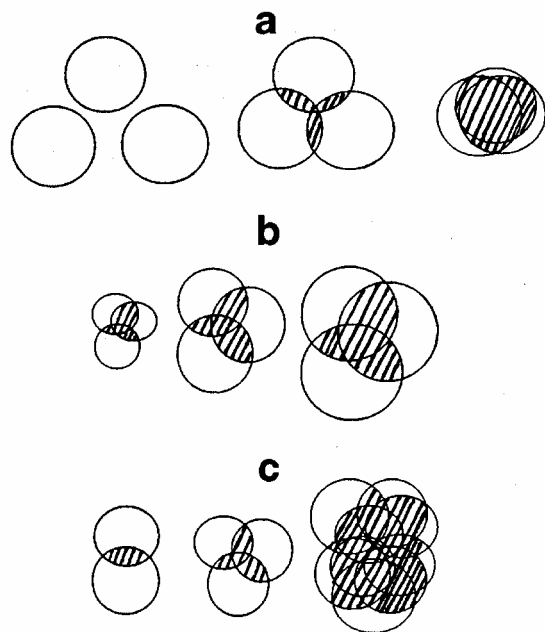


Figure 3. Expected relationships between florula diversity [FD] and associatum [AS] represented by Venn-diagrams. (The sizes of the circles are proportional to the entropy of species. Pooled entropy is the local distinctiveness [LD], union of circles represents FD , and the intersection of circles represents AS). a. Negative relationship between FD and AS is expected if local distinctiveness [LD] is constant. b. Positive relationship between FD and AS is expected if LD increases along with the increase of local entropies. c. Positive relationship between FD and AS is expected if LD increases as the number of species also increases.

III.2. The relationships of coenostate descriptors

Equation [4] expresses the fact that, at a certain sampling unit size j , the main coenostate descriptors are canonical, i.e., any two of them determine the actual value of the third one. In principle at least, this relationship alone would allow the entire FD_{max} - AS_{max} state plane to be 'filled' with coenostates, given a large enough set of samples from many different communities, but this turns out not to be the case for real data. Coenostates and coenostate trajectories are constrained definitely to certain parts of the state plane. This is in part an inherent feature of the method, but to a large extent it is a result of the spatiotemporal organization of community patterns. We shall highlight a few methodological aspects of the relationship between spatial diversity and spatial dependence, i.e., those of the interdependence of the FD_{max} - AS_{max} state plane axes.

Equation 4 itself suggests a negative relationship between FD and AS if the marginals of the 2^s contingency table remain constant (i.e., the cells change such that LD is fixed; Fig. 3a). Comparing two communities with identical abundances and spatial patterns of each species, the community with stronger spatial dependence between species will be less diverse in terms of the local species combinations. Clearly, if the spatial pattern of a species depends on that of another (leading to larger associatum), then some species combinations will never be realized in the sample (causing lower florula diversity). Inverting the same argument: if the individuals of all species are allocated to the habitat space independently of one another (no associatum), we can expect all potential species combinations to be realized in a large enough sample so that maximum florula diversity is closely approximated.

The primary successional process in the example above (Fig. 2c) produced a clear positive relationship between FD_{max} and AS_{max} with time, obviously because the community was 'under development'. The appearance of new species by itself causes increases in LD , enabling a simultaneous increase in spatial complexity (FD) and spatial dependence (AS) (Fig. 3bc). Note, however, that this relation is not as clear-cut as it appears for the first sight, since the axes of the coenostate-plane are the *maxima* of FD and AS , which usually does not imply the same quadrat size j ! The strict $LD_j = FD_j + AS_j$ relationship applies only for the same j , which suggests that the FD_{max} - AS_{max} trajectories are best interpreted together with the complementary plot of the quadrat sizes at which the maxima appear (CA_{FD} - CA_{AS}).

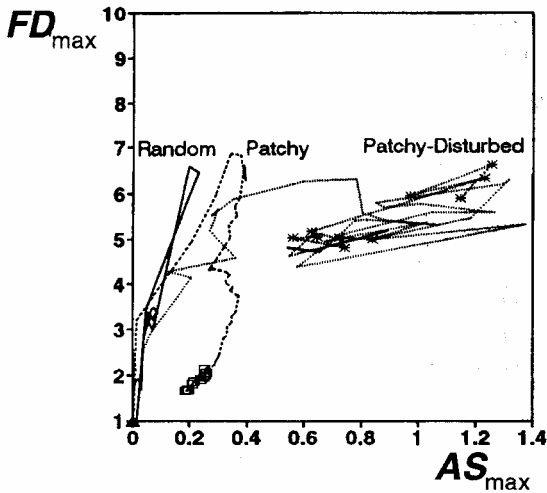


Figure 4. Examples for simulated trajectories in the FD_{max} - AS_{max} state-space (for explanation, see text).

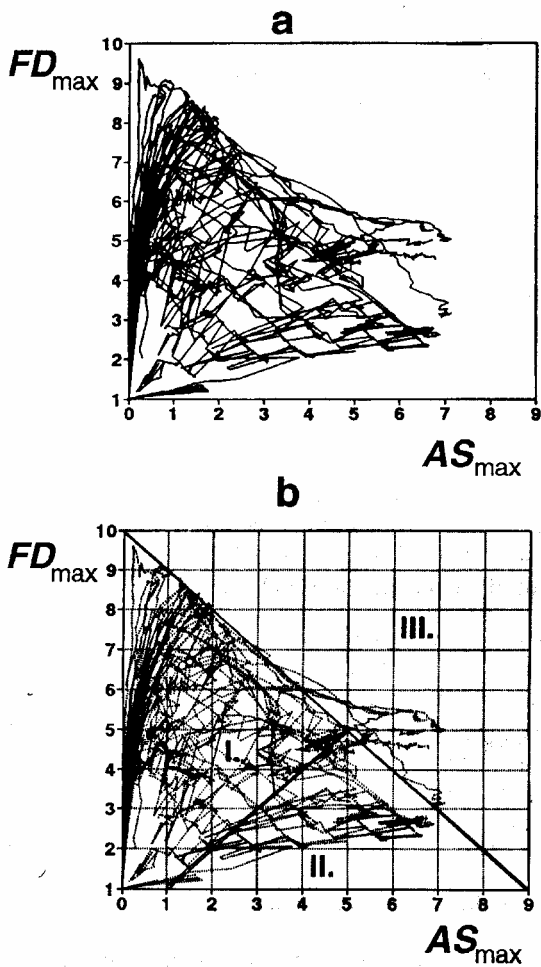


Figure 5a. Variation of the successional trajectories in the FD_{max} - AS_{max} state-space; b. The primary structure of the same state-space (for explanation, see text).

The rationale behind these considerations is that, although there is an obvious interdependence of the coenostate descriptors we use, no simple *a priori* positive or negative relationships between these two coenostate variables can be expected. Various mechanisms generate trajectories of various forms in the florula diversity - associatum state-space, representing different spatiotemporal dynamics of vegetation. Fig. 4 presents three typical examples of such trajectories, which were produced by a spatially explicit, individual-based Monte-Carlo competition model. Our simulated plant communities develop from a 10-species pool including both annuals and perennials, with trade-offs between their competitive abilities and reproductive capacities. There was a strong competitive hierarchy among the species, and competitive strength was also in trade-off with dispersal capacity. (For more details about the model and applications, see Czárán & Bartha 1989, and Czárán 1993, 1998). The model simulates typical primary successional processes with a gradual exchange of species, starting with well-spreading weak competitor annuals and ending with the dominance of a slow-growing, dispersal-limited, strong competitor perennial. Due to limited dispersal and the local nature of competitive interactions, markedly patchy patterns may develop and later vanish during succession. Such conspicuous spatial patterns slow competitive exclusions down, occasionally maintaining the diversity for quite a long time. Figure 4 shows that differences in the vital attributes (see Noble & Slatyer 1980, Pickett et al. 1987, Pickett & McDonnell 1989) of the simulated species (e.g., high or low degree of dispersal limitation producing 'patchy' or 'random' types of patterns) or the presence of spatially contagious disturbances might produce considerably different trajectories in the state-space.

IV. Exploring pattern and process within the FD_{max} - AS_{max} state space

The simulations were run with a wide range of input parameters, initializing various types of 'virtual succession', which yielded many markedly different trajectories within the florula diversity - associatum (FD_{max} - AS_{max}) state-space (Figure 5). Each simulation was allowed to run until the 50th generation, with a maximum of 10 species in all runs. The experiments differed with respect to the competitive relationships, the propagulum availability and the demographic structure of the populations, in addition to the disturbance regimes and the spatial heterogeneity of the environment. The range of parameters was chosen so as to fill the two-dimensional state-space by trajectories as much as possible.

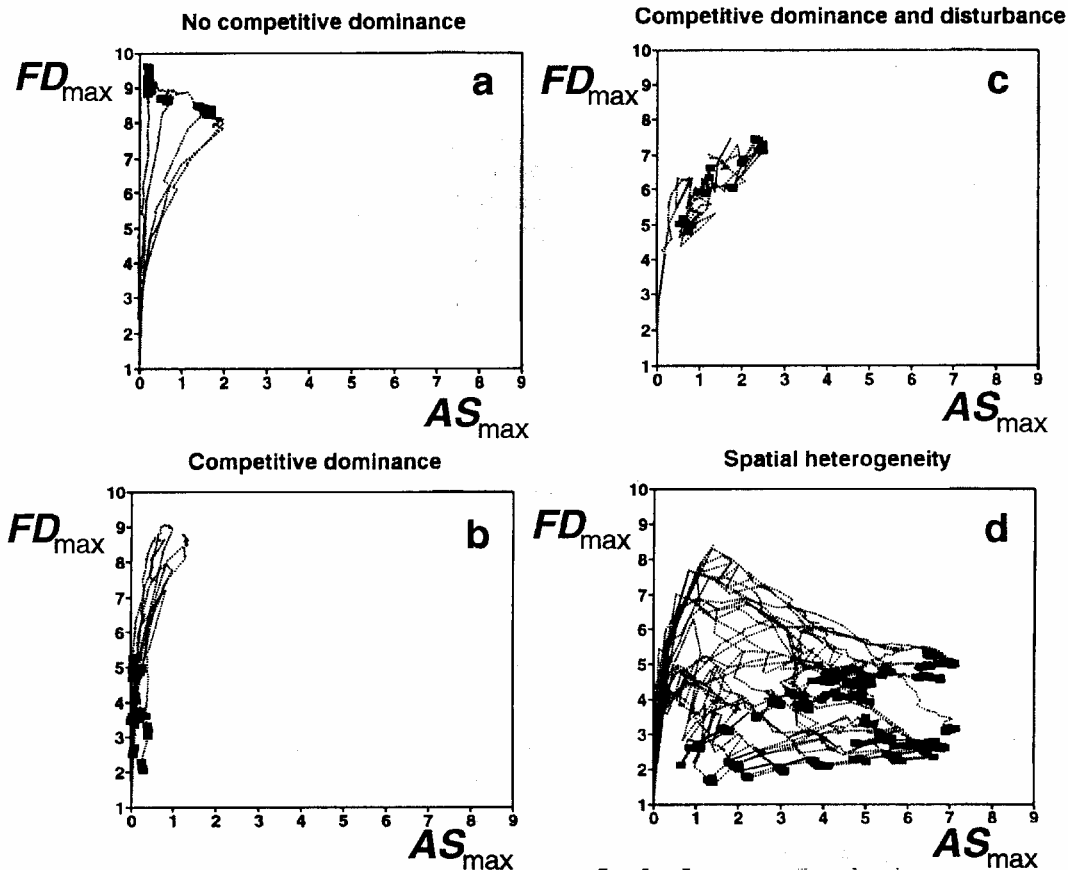


Figure 6. Relationships between patterns and mechanisms. Dotted lines: transitive stages of dynamics; ■: stationary stages of dynamics.

IV.1. The structure of the state space

It is clear from Figure 5 that the trajectories are unevenly dispersed over the FD_{max} - AS_{max} plane notwithstanding that many parameter combinations, some of them very unlikely under field conditions, were used. The domain of high florula diversity (FD) and high associatum (AS) remains empty. At first glance, this seems to be an obvious outcome of the canonical feature of local distinctiveness, florula diversity and associatum (Eq. [4]), but the boundaries around the simulated trajectories in the state space deserve more attention.

The number of species was 10 at most during the simulations, and thus $FD+AS=LD$ is constrained at a certain quadrat size: LD cannot be larger than $LD_{max} = 10 \times 2 \times (-1/2 \log_2 1/2) = 10$. LD_{max} is clearly the upper limit of the maximum of FD (in case of a random pattern with zero species associations). The lower limit of FD_{max} is 0 in theory, but this can be attained only if the pattern is empty (no individuals are present). Occurrence of a single species in the sampled region results

in there always being a quadrat size at which approximately half of the quadrats are empty, and each in the rest contains at least one individual of the species. At this quadrat size, compositional diversity is close to $FD = 2 \times (-1/2 \log_2 1/2) = 1$; i.e., the lower limit of the maximum of FD is unity.

For calculating the upper and the lower limits of maximum associatum (AS_{max}), recall that associatum is the difference between LD and FD (Equation 4). If LD is maximum (10 bits, in the example) and FD is minimum (1 bit), their difference is the highest (10-1=9 bits), i.e., 9 bits will be the upper limit for AS_{max} . If the spatial pattern of the simulated community is random, then $LD = FD$, that is, the lower bound for AS_{max} is 0.

The FD_{max} - AS_{max} state-plane is constrained within these limits for any 10-species community, and the canonical relation for FD and AS reduces the realistic part of the state-plane to an approximately triangular shape (Figure 5a). The diagonal line corresponds to the cases when the relationship $FD_{max} = 10 - AS_{max}$ holds. Because of this relationship high FD and high AS can-

not appear simultaneously, hence the almost empty area III. However, the diagonal line of this diagram is a rather fuzzy one: some trajectories do cross it and remain within the non-feasible triangle for a while, albeit always close to the diagonal. The reason for this was already mentioned in the previous section: for any given spatial pattern, the maxima of FD and AS may appear at different quadrat sizes, thus the sum of FD_{max} and AS_{max} may exceed the possible maximum of LD (10 bits in our example).

When we compare FD_{max} and AS_{max} , a logical expectation would be that the latter is always smaller than the former. In fact, if spatial dependence is extremely strong in the stand, it is possible that AS_{max} exceeds FD_{max} , and in such a case the two maxima may pertain to quite different spatial scales (quadrat sizes). Recall

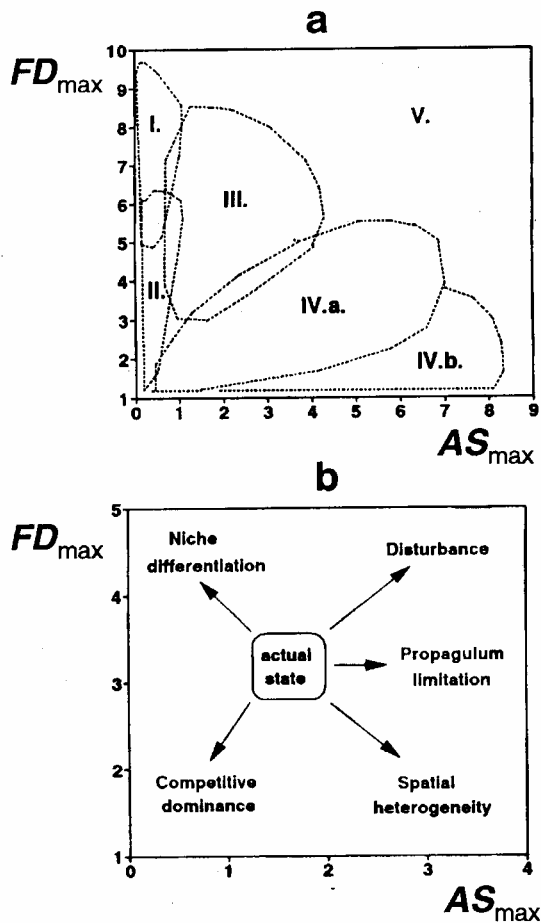


Figure 7a. Structure of the state-space based on the dominant pattern generation mechanisms. I. niche differentiation; II. competitive dominance; III. competitive dominance + disturbance; IV.a. environmental heterogeneity; IV.b. strong environmental heterogeneity; b, Direction of local transformations of an actual vegetation state due to the effect of a specific pattern generating mechanism.

that the value of florula diversity (FD) depends on the frequency distribution of realized species combinations in the sample. Let us imagine a step of succession whereby a new, colonizing species joins exclusively to an existing species or species combination very closely. Then, while the number of species is increased by the colonist, the frequency distribution of the realized species combinations (and thus FD) remains the same. As species richness increases, so does local distinctiveness (LD), and, due to the canonical relationship (Equation 4), AS increases as well. If such a successional process is repeated a sufficient number of times, AS may exceed FD , but the resulting pattern will show extreme spatial dependence that is obviously rare in nature. A situation approaching our hypothetical case might appear in extremely heterogeneous environments such as a salt-grassland with a two-phase mosaic of high and low concentrations of NaCl, where each phase has its own potential flora different from that of the other.

The region of the state space within the theoretical limits of FD_{max} and AS_{max} can be divided into three parts. In the first part $AS_{max} < FD_{max}$; this is where most of the simulated trajectories are found. It can be characterized by small or moderate spatial dependence of the simulated species patterns. Fewer trajectories stretched into part II in which $AS_{max} > FD_{max}$. This segment of the state space represents very marked spatial dependence relations, possibly – but not necessarily – connected to a patchy habitat or some kind of obligate (positive or negative) population interactions. Area III is the 'quasi-prohibited' part of the state plane, with a few trajectories occasionally crossing its margins through the diagonal boundary line, as explained above.

IV.2. Succession trajectories

A logical question is whether there are functional relationships between the pathways of trajectories and the starting parameter settings used in the simulations. A maximalist goal would be to characterize and to classify as many succession types as possible, based on the trajectories they produce within the FD_{max} - AS_{max} plane. Of course, we cannot complete such an ambitious program in this paper; only a few typical, markedly different types of vegetation dynamics can be analyzed here.

We classified the simulated trajectories according to the areas they run through in the state space, and found some correspondence between clusters of trajectories and the types of pattern generating mechanisms (Fig. 6). Since most of the simulated processes reached their quasi steady-state after ca. 40 generations, these may

be collectively termed the 'transitional phase' of the process, whereas the last 10 generations (41-50) are designated as the 'steady-state'.

Type I. Successional processes without competitive dominance in homogeneous environment

If intraspecific competition is stronger than the interspecific and the abiotic environment is homogeneous, then steady-state patterns with maximum florula diversity and minimum spatial dependence (associatum) develop (Fig. 6a). In this case, the species differ in their limiting environmental factors (we may refer to this by the metaphor 'niche differentiation'). Therefore, the spatial distributions of species are independent of one another. Contrary to the steady state condition, the early and mid-successional stages of this type have lower florula diversity (FD) and might show some spatial dependence. This is because the diversity of species combinations is limited by the low abundance and/or the limited dispersal of some component species. Associatum may be the highest in the transitional phase due to the autocorrelations in the patterns of individual species (patchiness) of transient patterns which disappear as the habitat becomes saturated. Demographic (fecundity, survival) and dispersal parameters are responsible for these transitional states.

Type II. Successional processes with competitive dominance in homogeneous environment

A marked competitive hierarchy between species or the existence of a single, absolutely dominant competitor results in patterns of low florula diversity and low spatial dependence (Fig. 6b). Succession ends with a single-species system of minimum florula diversity and associatum. The transitional states may have relatively high florula diversity in the early phase of succession, and some spatial dependence might also show up due to short-range propagule dispersal and spatiotemporal mass effects. Slow dispersal might isolate subordinated species from the dominants for a while, thus slowing down the process of extinctions, but the victory of the superior species is inevitable; it is just a matter of time. The steady state is hard-wired into the competitive relationships, but the demographic and the dispersal parameters can modify the speed of succession and the trace of the trajectories during the early transitional phases.

Type III. Successional processes with competitive dominance and disturbances

The process of competitive exclusion might be halted or modified by local disturbances if there is a trade-off between the competitive abilities of the species and their capacity to tolerate disturbances or to regenerate

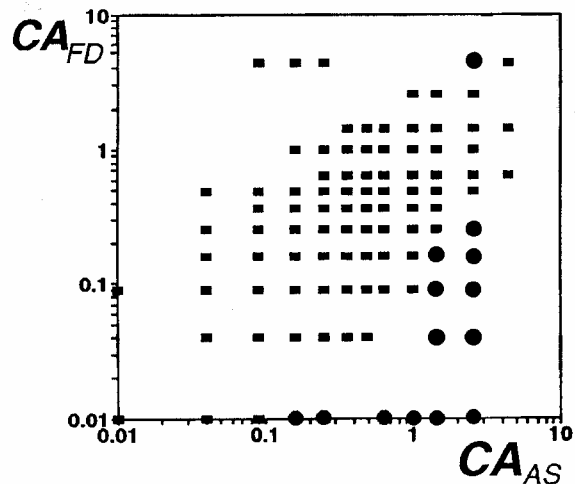


Figure 8. Variation of the simulated characteristic maximum scales in the state-space (● $AS > FD$, ■ $AS \leq FD$).

after disturbance. It can produce steady-states of high or medium florula diversity and low or medium spatial dependence (Fig. 6c), the actual outcome strongly depending on the spatiotemporal pattern of disturbance and that of propagule dispersion. Considering that stand-level diversity in a disturbed community is maintained by an externally driven, spatially heterogeneous process, demographic and dispersal parameters are even more important here in defining the position of trajectories than they are in other types.

Type IV. Successional processes with spatial heterogeneity

Suppose there is some patchy environmental spatial variation that is invariant in time. On each of the patches, there is an on-going selection in favour of the species combination best adapted to the local environment - each patch supports a certain species combination. In such a spatially heterogeneous selective environment, the number and the diversity of species combinations (FD) will be medium or low, combined with spatial dependence (AS) of varying extent. There is a wide range of steady-states that belong to this type (Fig. 6d). For example, if the local diversity and the environmental heterogeneity (selectivity) of the patches are low, the steady-state pattern will show small florula diversity and low spatial dependence. Many species but high heterogeneity (selectivity) of the environment might result in vegetation states of low florula diversity (i.e., a two-phase mosaic) but very high spatial dependence (i.e., a complete association of many species to the same components of the environmental mosaic). A multi-phase environmental

mosaic with many species and complex coalitional structure might result in steady-states of medium florula diversity and very high spatial dependence. Trajectories for early phases of successional processes can run through various areas in the state-space. These transitional dynamics are determined mainly by the demographic and dispersal parameters of the species, together with the actual environmental pattern of the habitat.

Our hypotheses about the structure of the florula diversity - associatum ($FS_{max} - AS_{max}$) state space based on theoretical considerations and simulation experiments are summarized in Fig. 7a. Field evidence from various plant communities would be necessary to validate these conclusions. However, in most field situations neither the potential species pool nor the dominant pattern-generating factors are known and they are far from being constant over time. In case of the usual complex pattern of mechanisms (Walker & Chapin 1987), we might reduce the bias of complexity

by following and testing only relatively uniform spatiotemporal domains for a start. To perform such smaller extent records, we need an alternative hypothesis about the direction of local transformations of an actual vegetation state due to the effect of specified pattern generating forces, as shown in Fig. 7b.

V. The state-space of characteristic areas ($CA_{FD} - CA_{AS}$)

The alternative state-space representation depicts coenological structure in terms of the characteristic areas of patterns (Fig. 2d). The state space that is complementary to the $FS_{max} - AS_{max}$ space is the $CA_{FD} - CA_{AS}$ plane, in which the coordinates of a stand are the quadrat sizes at which the maxima of florula diversity (FD) and associatum (AS) were detected. The structure of this plane is explored in a way analogous to the investigations described in the previous section. Figure 8 summarizes the characteristic maxima of vegetation

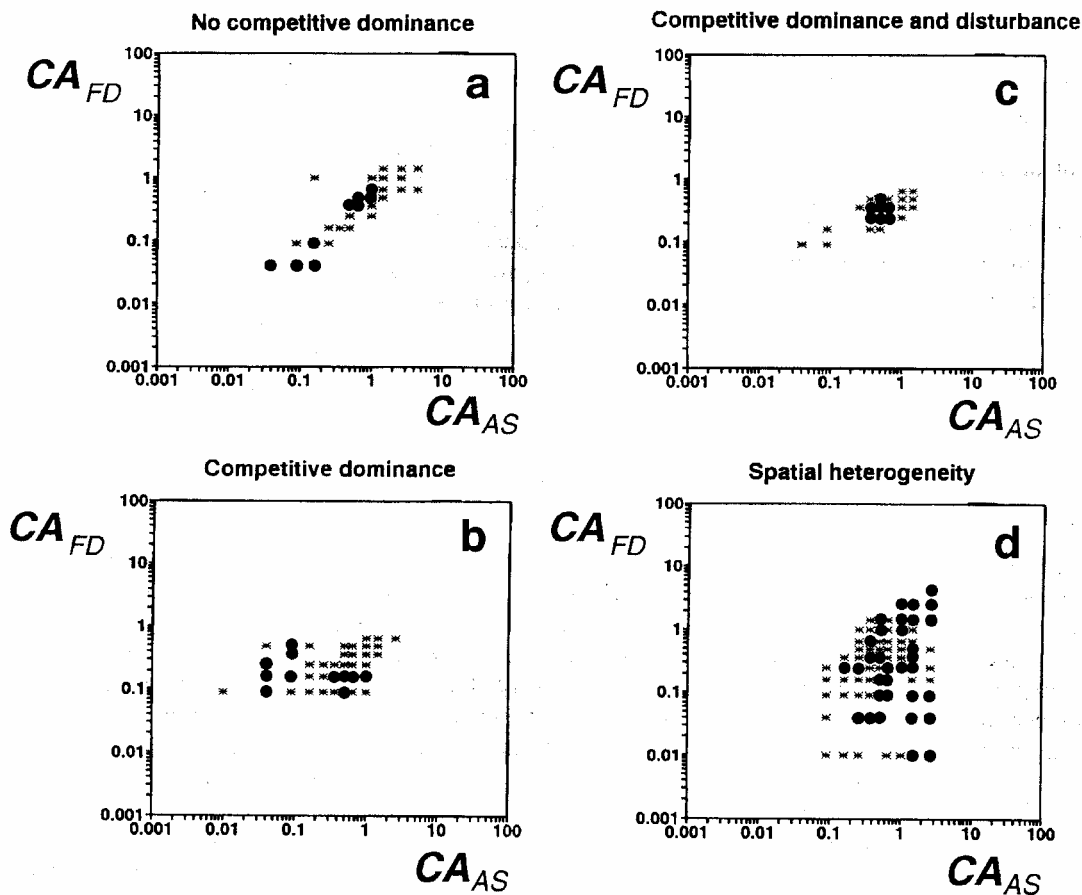


Figure 9. Relationships between pattern generating mechanisms and the resulting characteristic maximum scales (* early successional stages, i.e., the 1-39. generations of simulations, ● late successional stages, i.e., the 40-50. generations of simulations).

states simulated with the same wide range of simulation parameters.

There are no theoretical bounds to exclude any combinations of maximum areas and, indeed, the points representing the simulated successional series are rather evenly distributed within the CA_{FS} - CA_{AS} space. However, if the patterns of extremely large spatial dependence (where the maximum value of associatum is larger than that of florula diversity) are treated separately, the remaining cloud of points shows a slight trend of positive relationship, suggesting that the maximum areas of the two coenological characteristics are not independent. If the points are grouped according to the four types of pattern generating mechanisms (Fig. 9), we obtain a series of somewhat more complex patterns. The general positive relationship is the strongest for the patterns of Type I (succession with no competitive dominance), and the weakest (or partly disappeared) for Type IV (succession with spatial heterogeneity). If we focus only on the upper bound of the point cloud, Type IV shows a positive relationship as well. However, there were additional states with small-scale compositional diversity (FD) maxima but larger-scale maxima of associatum (AS), which might be a result of the strong spatial differentiation of the simulated species.

The trend of decreasing characteristic areas during succession was observed in several field studies (Juhász-Nagy 1980, Juhász-Nagy & Podani 1983, Bartha 1993) but the simulation experiments produced much less consistent results. There was a slight decreasing trend found in Type I and II (Figure 9ab). Late successional (steady-state) characteristic areas in Type III and IV (Fig. 9cd) reflected the actual spatial pattern of disturbance and environmental heterogeneity rather than a dynamical trend.

Concluding remarks

We emphasize throughout the paper that possibilities to depict spatiotemporal patterns of vegetation in abstract state spaces are by no means exhausted. In particular, we suggest that information theory statistics designed to explore multispecies patterns based on presence/absence data offer a sophisticated tool for such analysis. The advantage of this approach lies in its ability to integrate scale and basic coenological phenomena such as diversity, preference and independence at a multivariate level – a synthesis for which no other methods are available. Trajectories in the FD_{max} - AS_{max} coenostate space are especially useful as phenomenological traces of plant community dynamical processes. We have shown (Section 4.2) that different dynamical effects yield different successional trajec-

tories, which means that their careful analysis may provide useful information on the corresponding overall driving forces behind the observed phenomenological transformations, even in remote and floristically very different communities. The picture can be further tinged by observing the CA_{fl} - CA_{as} trajectories of the successional processes. It is important to see, however, that the phenomenological analysis does not replace detailed dynamical studies – the actual 'microscopic' mechanisms of plant-to-plant interactions and propagule dispersion are implicit in the coenostate transformations, thus the study of the former cannot be spared by using the latter.

In some sense, the purpose of the information statistical functions we propose for tracing coenostate transformations in plant communities resembles the role of intensive state variables – like temperature and pressure – in the statistical physics of gases. Temperature and pressure are phenomenological variables expressing the macroscopic state of a gas or a gas mixture of known volume, derived from the microscopic properties of the constituent molecules, the rules of classical mechanics, and statistical principles. Similarly, florula diversity and associatum are phenomenological variables expressing overall spatial relationships among the constituent populations of plant individuals, without any direct reference to the actual details of their spatial positions and forms of dynamical interaction. To obtain information on the isolated pattern of any particular population, or the combined pattern of any particular multiplet of populations, FD and AS can be decomposed to additive entropy terms for any set of species combinations (florulas), in a way somewhat similar to how it can be done for gas mixtures using the appropriate partial terms. We did not tackle decompositions into 'partial communities' in this paper – the reader is referred to the original papers of Juhász-Nagy (1984, 1993) for details in this respect.

Exploring coenostate transformations in abstract coenostate spaces is a laborious task, given that it requires an intensive sampling of the community stand in a spatial series (Podani 1993). The most efficient way of doing this is mapping all plant individuals within the sample area, and performing sampling and the entire analysis by computer in an automated fashion. With classical mapping procedures, even this requires a lot of time and energy, part of which could be saved by using remote sensing techniques, or by combining these with the classical methods. Remote sensing overcomes much of the limitations of field data collection and provides huge spatio-temporal data sets at a relatively low cost and without destroying the vegetation. Based on remotely sensed data, a state-

space approach interestingly similar to ours was developed by Henebry & Su (1993) for the study of vegetation cover transformations. They used spatial variability and spatial dependence state-variables for univariate data in a framework similar to that of Juhász-Nagy (1993). We believe that the future technical development of sensors used for small-scale and satellite imageries will promote the widespread use of scaled coenostate space methods in studies of community organization.

Acknowledgements: The authors thank T. Hochstrasser, L. Mucina and M. Kertész for their suggestions and comments. This work was partially funded by Hungarian National Research Grant (OTKA) No. T16225 (to S. B.), No. T19524 (to T. C.) and No. T19364 (to J. P.).

References

- Agnew, A. D. Q., Collins, S. L. & van der Maarel, E. (eds.) 1993. Mechanisms and processes in vegetation dynamics. *J. Veg. Sci.* vol. 4.
- Austin, M. P. 1977. Use of ordination and other multivariate descriptive methods to study succession. *Vegetatio* 35: 165-175.
- Bartha, S. 1990. Spatial processes in developing plant communities: pattern formation detected using information theory. In: Krahulec, F., Agnew, A. D. Q., Agnew, S. & Willems, J. H. (eds.), *Spatial processes in plant communities*. pp. 31-47. Academia, Prague and SPB Academic Publishing, The Hague.
- Bartha, S. 1992. Preliminary scaling for multi-species coalitions in primary succession. *Abstracta Botanica* 16: 31-41.
- Bartha, S. 1993. Gyomnövényközösségek szünmorfogenezise külszíni szénbányák meddőhányóin. (Vegetation development on dumps from strip-coal mining) (In Hung.) Kandidátusi Disszertáció, (Ph. D. thesis) Vácrátót
- Bartha, S., Collins, S. L., Glenn, S. M. & Kertész, M. 1995b. Fine-scale spatial organization of tallgrass prairie vegetation along a topographic gradient. *Folia Geobotanica and Phytotaxonomica* 30: 169-184.
- Bartha, S., Czárán, T. & Scheuring, I. 1997. Spatio-temporal scales of non-equilibrium community dynamics: a methodological challenge. *New Zealand Journal of Ecology* 21: 199-206.
- Bartha, S., Czárán, T., Oborny, B., Podani, J. & Kertész, M. 1994. JNP-models 1.0: Program package for detecting coexistential patterns in multi-species assemblages. Abstracts of the 3rd Hungarian Congr. of Ecology, Szeged, July 4-6, 1994, p. 17.
- Bartha, S. & Kertész, M. (in press) The importance of null-models in detecting interspecific spatial associations from 'transect' data. *Tiscia*
- Braun-Blanquet, J. 1964. *Pflanzensoziologie*. 3. Aufl. Springer, Wien-New York.
- Christensen, N. L. & Peet, R. K. 1984. Convergence during secondary forest succession. *J. Ecol.* 72: 25-36.
- Collins, S. L., Adams, D. E. 1983. Succession in grasslands: Thirty-two years of change in a central Oklahoma tallgrass prairie. *Vegetatio* 51: 181-190.
- 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. & Bartha, S. 1989. The effect of spatial pattern on community dynamics: a comparison of simulated and field data. *Vegetatio* 83: 229-239.
- Diggle, P. J. 1983. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.
- Feoli, E. & Orlóci, L. (eds.) 1991. *Computer Assisted Vegetation Analysis*. Kluwer, Dordrecht.
- Gauch, H. G. 1982. *Multivariate Analysis in Community Ecology*. Cambridge Univ. Press, Cambridge.
- Green, R. H. 1979. *Sampling Design and Statistical Methods for environmental Biologists*. Wiley, New York.
- Greig-Smith, P. 1983. *Quantitative Plant Ecology*. 3rd ed. Blackwell, Oxford.
- Henebry, G. M. & Su, H. 1993. Using landscape trajectories to assess the effects of radiometric rectification. *Int. J. Remote Sensing* 14: 2417-2423.
- Hogeweg, P., Hesper, B., van Schaik, C. P. & Beeftink, W. G. 1985. Patterns in vegetation succession, an ecomorphological study. In: White, J. (ed.): *The population structure of vegetation*. Dr. W. Junk Publ., Dordrecht, pp. 637-666.
- Hochstrasser, T. 1995. The structure of different loess grasslands types in Hungary. M. Sc. thesis, Vácrátót
- Juhász-Nagy, P. 1967. On some 'characteristic areas' of plant community stands. *Proc. Colloq. Inf. Theory, Bolyai Math. Soc., Debrecen* pp. 269-282.
- Juhász-Nagy, P. 1980. Modeling of coexistential structures in coenology. Doctoral Thesis (Manuscript, in Hungarian), 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. 1984. Spatial dependence of plant populations. Part 2. A family of new models. *Acta Bot. Acad. Sci. Hung.* 30: 363-402.
- Juhász-Nagy, P. 1993a. Notes on compositional diversity. *Hydrobiologia* 249: 173-182.
- Juhász-Nagy, P. 1993b. The fading diversity. *Scientia Publishing, Budapest*. (in Hungarian)
- Juhász-Nagy, P. & Podani, J. 1983. Information theory methods for the study of spatial processes and succession. *Vegetatio* 51: 129-140.
- Legendre, L. & Legendre, P. 1983. *Numerical Ecology*. Elsevier, Amsterdam.
- Magurran, A. E. 1988. *Ecological Diversity and its Measurement*. Princeton Univ. Press, Princeton.
- Manly, B. F. J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London.

- Matthews, J. A. 1979. A study of the variability of some successional and climax plant assemblage-types using multiple discriminant analysis. *J. Ecol.* 67: 255-271.
- Muller-Dombois, D. & Ellenberg 1974. *Aims and Methods of Vegetation Ecology*. Wiley, New York.
- Mucina, L., Grabherr, G. & Ellmauer, T. 1993. *Pflanzengesellschaften Österreichs. Wälder und Gebüsche*. New York.
- Noble, I. R. & Slatyer, R. O. 1980. The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbance. *Vegetatio* 43: 5-21.
- Oborny B. 1994. Growth rules in clonal plants and predictability of the environment: A simulation study. *Journal of Ecology* 82: 341-351.
- Orlói, L. 1978. *Multivariate Analysis in Vegetation Research*. 2nd ed. Junk, The Hague.
- Padisák, J. 1992. Spatial and temporal scales in phytoplankton ecology. *Abstr. Bot.* 16: 16-24.
- Padisák, J. 1994. Identification of relevant time-scale in non-equilibrium community dynamics: conclusions from phytoplankton surveys. *New Zealand Journal of Ecology* 18: 169-176.
- Palmer, M. W. & van der Maarel, E. 1995. Variance in species richness, species association, and niche limitation. *Oikos* 73:203-213.
- Pickett, S. T. A., Collins, S. L. & Armesto, J. J. 1987a. Models, Mechanisms and Pathways of Succession. *The Botanical Review* 53: 335-371.
- Pickett, S. T. A. & McDonnell, M. J. 1989. Changing perspectives in community dynamics: a theory of successional forces. *TREE* 4: 241-245.
- Pielou, E. C. 1975. *Ecological Diversity*. Wiley, New York.
- Podani, J. 1984a. Spatial processes in the analysis of vegetation. Theory and review. *Acta Bot. Acad. Sci. Hung.* 30:75-118.
- Podani, J. 1984b. Analysis of mapped and simulated vegetation patterns by means of computerized sampling techniques. *Acta Bot. Acad. Sci. Hung.* 30:403-425.
- Podani, J. 1987. Computerized sampling in vegetation studies. *Coenoses* 2: 9-18.
- Podani, J. 1992. Space series analysis of vegetation: processes reconsidered. *Abstracta Botanica* 16: 25-29.
- Podani, J. 1997. SYN-TAX 5.1: A new version for PC and Macintosh computers. *Coenoses* 12:149-152.
- Podani, J. & Czárán, T. 1997. Individual-centered analysis of mapped point patterns representing multi-species assemblages. *J. Veg. Sci.* 8: 259-270.
- Podani, J., Czárán, T. & Bartha, S. 1993. Pattern, area and diversity: the importance of spatial scale in species assemblages. *Abstracta Botanica* 17: 37-51.
- Reynolds, C. S. 1995. The intermediate disturbance hypothesis and its applicability to planktonic communities. Comments on the views expressed in Padisák -v- Wilson. *New Zealand Journal of Ecology* 19: 219-225.
- Schneider, D. C. 1994. *Quantitative Ecology. Spatial and Temporal Scaling*. Academic, San Diego.
- Szegi, J., Oláh, J., Fekete, G., Halász, T., Várallyay, Gy. & Bartha, S. 1988. Recultivation of the spoil banks created by open-cut mining activities in Hungary. *Ambio* 17: 137-143.
- Szollát, Gy. & Bartha, S. 1991. Pattern analyses of dolomite grassland communities using information theory models. *Abstracta Botanica* 15: 47-60.
- Tóthmérész, B. & Erdei, Zs. 1992. The effect of species dominance on information theory characteristics of plant communities. *Abstr. Bot.* 16: 43-47.
- Tóthmérész, B. 1994b. Statistical analysis of spatial pattern in plant communities. *Coenoses* 9: 33-41.
- van der Maarel, E. & Sykes, M. 1993. Small-scale spatial dynamics of plant species in a grassland community over six years. *J. Veg. Sci.* 4:179-188.
- Walker, L. R. & Chapin III, F. S. 1987. Interactions among processes controlling successional change. *Oikos* 50: 131-135.
- Wiens, J. A. 1989. Spatial scaling in ecology. *Funct. Ecol.* 3:385-397.
- Zobel, K. & Zobel, M. 1988. A new null hypothesis for measuring the degree of plant community organization. *Vegetatio* 75:17-25.

Appendix: Computer programs

Since the methodology outlined in this paper relies heavily on computerized implementations, we describe two sets of programs that are designed for calculating information statistics from field or simulated data.

SYN-PATT and MULPATT

Program **SYN-PATT** (formerly part of the **SYN-TAX** program package) has been developed for PC-DOS machines, whereas the module **MULPATT** does the same task on Mac OS systems (Podani 1997). The characteristic functions may be calculated from a series of data matrices obtained from field sampling, or from digitized point patterns indirectly. In the latter case, computerized sampling is performed first, with user specified sampling parameters (shape, size, arrangement and number of plots). The output includes a list of numerical results, and separate graphic displays of characteristic functions (as in Fig. 2ab, but for one community at a time). There is a utility to examine point patterns visually for all species, for a given pair of species, or for a single species. There is no direct option to display coenostate configurations; the appropriate information may be extracted from the output lists and then visualized by commercial graphics programs, such as **QUATTRO** or **Statistica**.

JNP-models, version 2.0

This package calculates Juhász-Nagy's information theory models without mixing it up with other spatial statistical indices. The main objective is to support routine applications for comparative analyses with large sets of data. In such analyses, we repeat the same applications with dozens of field and hundreds of simulated data sets. For practical reasons, the analyses based

on presence/absence data are collected in fine-resolution grids or along transects, because most of the available field data belong to this type. Maps representing point processes in continuous space, as in many simulated data sets, are first converted to grids, then analysed in the same way as the field data.

Compared to the previous version of this package (**JNP-models 1.0**, Bartha et al. 1994), emphasis is on the routine analyses with the same standard technique, in a automatized fashion. While the previous version performed each step of scaling, each related sampling and analysis with separate programs, the new version combines these steps into a single run to simplify and speed up the computations. The number of options for formulae calculated, as well as the number of scaling options applicable have been reduced to the most frequently used ones. The maximum dimensions of data were also reduced, but fitted to the usual dimensions of field data. Considering the scaling options and model selection, there is a hidden expert system built into this new version that is based on our 15 year routine of applying these information statistics. Because of the automatized nature, we recommend this version to experienced users only.

Common features

The programs were developed and tested by S. Bartha in Microsoft FORTRAN 77, version 5.1, for IBM PC or compatibles with built in coprocessor. The programs run under the DOS environment with at least 500 Kbytes of RAM and use ASCII text files in FORTRAN format for input and output.

Input format

The programs that calculate the informations statistics use standard input format. In case of grid data, input data are organized into separate blocks, where each block shows the presence/absence of an individual species in the grid. Blocks are started (separated) by a title line (I5,A6,A35) that is the id. code of species, followed by an alphabetic code, and the full name of the species. The latter two are optional. However, if the id. code is set to zero, the particular species will be omitted from the analyses. For transects, input data are organized into separate blocks as well, where each block shows the presence/absence of an individual species along the transect (I80). If a transect is longer than 80 units, it continues automatically in the next line. Blocks are started (separated) by the same type of title line (I5,A6,A35) as for grid data.

There are interactive programs available for creating and transforming the standard data format required by the other programs (see below **LINPREP**, **LINSET** etc.).

Input and sampling options:

Transect data can be of either linear or circular topology. Similarly, grid data can be of planar or toroid shape. (The latter is unrealistic in case of field data,

however, many simulation programs work with toroidal grids to avoid edge effects.)

Sampling can be performed by a fixed number of randomly positioned overlapping sampling units (random sampling). The alternative is the 'complete sampling' where overlapping sampling units are taken from all possible positions at a given sampling unit sizes. Spatial scaling that is an inherent feature of Juhász-Nagy's model are performed in spatial series (Podani 1992), i.e. the sampling procedure is repeated with gradually increasing sampling units.

There is an option to select frequent species from the total species pool based on their frequencies. Rare species might cause artefacts, therefore using frequency limit at least 1% is recommended.

Output format

Output lists are organized similarly to **SYN-PATT** and **MULPATT**, i.e. they can be exported to and displayed by commercial graphics programs (**QUATTRO**, **EXCEL**, etc).

1. Data preparation, transformations and testing

1.1. Transect data. For practical reasons, the simplest way to enter presence/absence data collected along a transect is to edit a file where each row is a micro-quadrat in the same sequence as along the transect. Each row has a list of id. codes (I3), i.e. the list of species that were present in the particular quadrat. This field notebook format is converted into binary transects of standard format by **LINPREP**. **LINTEX** calculates the frequency of species along the transect. The binary transect can be transformed by rearranging the sequence of species, for example, sorting them by decreasing frequencies or fusing them into functional groups. These operations are performed by **LINSET**. **LINSET** needs a parameter file where each row contains an integer, a species code, and the species name (I2,A6,A35). The integer shows the new position of the particular species in the new rearranged transect. If two (or more) species have the same position in the new transect, they will be fused.

1.2 Grids. **MGRID** transforms digitized coordinate data (species code, x coordinate, y coordinate) into binary grids of standard format. **GRIDTEX** calculates the frequency of species. The rearrangement or fusing of species can be done by **GRIDSET** based on the same type of parameter file used by **LINSET**.

2. Detailed analysis of a single community with randomization tests

All programs require two types of parameter files. **SCALE.PAR** is a simple list of scaling parameters, where each row shows the number of adjacent sampling units (I5) to be fused in the particular step of the spatial series. **SPNAME.PAR** contains the code for the functional group membership (if unknown, set it to 1), the species codes, and names. Input format is (I2,A6,A35).

2.1. **PRIMPRO1** (for transects) and **PRIMPRO2** (for grids) calculate the most important information statistics. There are separate output files for the community level (syncretic) and the population level (autocretic) models. Syncretic results include the number and diversity of species combinations, dissociatum, associatum, interassociatum, local distinctiveness, and two versions of relative associatum. Autocretic results are the local entropy and the complete associatum for each species. There is a third output file showing the difference between the field estimate and the average random reference of the autocretic models (Podani & Czárán 1997).

2.2. **NPASS1** (for transects) and **NPASS2** (for grids) calculate the pairwise spatial associations of species. We apply a new method here, where the signs of associations are also determined from randomization test based on the comparison of frequency distributions of observed and simulated random joint distributions of a species pair (Bartha & Kertész, in press). Two types of randomization can be chosen as neutral models. (1) Complete randomization (Diggle 1983) randomizes the positions of individuals (the presence in our case) at the finest resolution, but the number of species and the frequency of species remain the same as in the field. (2) Random shifts (Palmer & van der Maarel 1995) can be applied to circular transects or toroidal grids (in case of simulated data). Random shifts keep the spatial pattern of each species, but randomize the relative (interspecific) positions of the species by shifting (rotating) it randomly along the circular transect or toroidal grid. 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 is calculated by comparing it with the values obtained by the Monte Carlo randomizations, i.e. representations of the null models. The significance is expressed as the relative frequency of the randomizations resulted in higher or lower values than the observed one (Manly 1991).

The standard output shows the observed value, the average, maximum and the minimum of randomized values, and the probabilities that the observed values are higher or lower than the randomized ones. All these statistics are shown across a range of scales, i.e., calculated in spatial series (Podani 1992).

3. Comparative analyses of series of communities

SYNGEP1 (for transects) and **SYNGEP2** (for grids) were developed to support the comparative analyses. The programs perform the community level analyses with a series of transects or grids (up to 99). These non-interactive programs use standard parameter files with

information about the name of the input data files, dimensions and scaling options. (Instructions to prepare them are written into the parameter files.)

There are six separate standard output files, **NRC.PRN** (number of realized species combinations), **FD.PRN** (florula diversity), **ASS.PRN** (associatum), **INTER.PRN** (interassociatum and dissociatum), **RELASS.PRN** (relative associatum, i.e. associatum/florula diversity, **LD.PRN**) for most important syncretic models. **KVAD.PRN** is a summary of the dimensions and scaling options. Each row in the output shows the maximum value and the related maximum scale of the particular model followed by the series of observed values in the spatial series. The output includes stands x scales matrices of the observed values that are summarized to enhance the graphical comparisons at the characteristic maxima or at particular spatial scales.

SYNGEP1 and **SYNGEP2** can be used with the same input data but changing scaling options, for example, with increasing the number of species analysed. This method is particularly useful to optimize the scaling parameters of a new type of data.

No randomization tests are built in these programs.

4. Current maxima of dimensions

The maximum number of species is 120 in the programs for data preparation, transformations, and testing. The maximum length of transects is 3100 units, and the maximum sizes of grids are 64 x 64 units. The maximum number of species analysed by **PRIMPRO1** and **SYNGEP1** is 50. It is set to 35 for **PRIMPRO2** and **SYNGEP2**, and to 25 for **NPASS1** and **NPASS2**. In case of random sampling, the maximum sample size is 3100 for transects and 4100 for grids. The number of randomizations for the randomization tests is unlimited (but 2500 - 5000 randomizations are recommended). Considering the working effort needed for collecting field data, the extent of transects rarely exceeds 2000 units, and most of the grids sampled are 50 x 50 units or smaller. The maximum size of the sampled transects and grids delimits the sample size that can be taken in the spatial series, and consequently it delimits the detectable number of species combinations. At the usual sample sizes, information statistical analyses with no more than 15 species are recommended.

Availability

Demo files of all steps of data preparation, and analyses are available. This version of the program package is free and can be obtained from the first author upon request.