25 June 2024, Preprint v1 · CC-BY 4.0 PREPRINT

Research Article

A Neg-Entropy Function Applied to Pairwise Co-Occurrence Matrices in the Study of Ecological Communities

Enrico Feoli[1]

1. Department of Life Sciences, University of Trieste, Italy

In this paper, I present the results of the application of an entropy function (H(C)), which I suggested in 1972, for matrices of pairwise co-occurrence of species in plant communities, to two sets of very simple "toy" matrices representing different situations of co-occurrence that species or "traits" may have in ecological communities. The first set consists of 7 binary matrices X(M, N) from which 7 pairwise co-occurrence matrices (C(M, M)) are generated by self-matrix multiplication. The second set consists of 6 co-occurrence matrices generated by 6 "toy" graphs that have the minimum number of connections (M-1) and that represent different situations as far as the centrality of nodes is concerned. The entropy (H(C)) is given by the difference between the entropy H(D) of the diagonal D of the co-occurrence matrix C(M, M) and the entropy H(T) of its upper or lower triangular part T, with M(M-1)/2 cells. H(C) is relativized in two ways to give origin to (K(C)) and (R(C)), which I call respectively "index of weighted connectance" and "redundancy"; these have been compared for both the sets of co-occurrence matrices with: a) the ratio (H(D)+H(T))/ ln ((M(M+1)/2), that is, the relative entropy or the evenness of the upper or lower triangular part of C(M, M) including the diagonal D (M, M) (i.e., D+T); b) the formula of Ricotta and Szeidel (2006), which in this case incorporates in the entropy H(D) the entropy of the matrix of co-occurrence itself with values relativized between 0 and 1; c) the evenness of the eigenvalues of C(M, M); d) the chi-square of C(M, M); e) the average nested free; and f) the nested-based similarity between the N vectors of the X(M, N) matrices; and g) the number of centres of the graphs corresponding to the co-occurrence matrices. I used the Pearson correlation coefficient to estimate the correlation between all the parameters used to calculate H(C) and the parameters for the comparisons from a) to g) of the two "toy" data sets. Finally, I used the two correlation matrices for ranking the parameters by the method suggested by Orlóci (1978), based on the sum of squares, in order to evaluate the capacity of each parameter in explaining the variation of the others. The results

doi.org/10.32388/T9WW4P

show that the correlation between the parameters is very high and that H(C) can be considered a good tool to describe the pattern of the graphs of the corresponding co-occurrence matrices.

Corresponding author: Enrico Feoli, feoli@units.it

Introduction

As a consequence of the fact that in 1970-71 I was speaking "everywhere almost fanatically" about information theory, I was invited by some of my botanist friends to write a "popular" article in Italian for the journal *Informatore Botanico Italiano* on its possible applications to phytosociology (Feoli 1972). Rather than taking care of mathematical formulas (in that paper, the minus sign in front of formula {7} was missing and I did not explain in detail the proposed formula {5} to be applied to pairwise co-occurrence matrices), I focused on transmitting the idea that vegetation should be interpreted as a system as a potential source of redundant messages. That is, redundant codes that we receive and organize in matrices X(M, N) in which M is the number of components of the system, species or other traits (Barkman 1979, Box 1981, Feoli 1984, Orlóci and Orlóci 1985, Pillar and Orlóci 1993, Natham et al. 2015, and references therein) and N is the number of sampled vegetation stands representing different vegetation states (see Orlóci 2020 for an innovative view on the vegetation stand).

In my mind, formula (5) of that paper (here numbered (2)) should be useful to measure the entropy of the co-occurrence matrices C(M, M) obtained by the self-multiplication of the matrices X(M, N):

$$\mathrm{C(M,M)} = \mathrm{X(M,N)}\, \mathrm{X}^{\mathrm{T}}\mathrm{(M,N)} \qquad (1)$$

or calculated by similarity functions between the M components (recall that the scalar product is also a measure of similarity in some sense).

The formula of entropy I proposed for co-occurrence matrices (Feoli 1972) is the following (in the original paper, it was number (5)):

$$\mathrm{H(C)} = -(\Sigma_{(\mathrm{ii})}\mathrm{p_{ii}}\mathrm{lnp_{ii}} - \Sigma_{(\mathrm{ih})}\mathrm{p_{ih}}\mathrm{lnp_{ih}}) \qquad (2)$$

with (ii)=1,..., M, and (ih) = 1,..., M(M-1)/2. The pii values are the ratio between the cross products and the total of the upper or lower part of the co-occurrence matrix that includes the diagonal values.

If we consider $-\Sigma_{(ii)}\, p_{ii} \ln p_{ii}$ = H(D), i.e., the entropy of the diagonal of C(M, M), and $-\Sigma_{(ih)}\, p_{ih} \ln p_{ih}$ = H(T), the entropy of the upper or lower triangular part of C(M, M), we can write the formula (2) as H(D)-H(T). In

doi.org/10.32388/T9WW4P

terms of graph theory (Diestel 2017), H(D) is the entropy of the diagonal matrix of the nodes (D), and H(T) is the entropy of the upper (or lower) triangular part (T) of the adjacency matrix.

The outcome of formula (2) should be useful if correlated with the environmental factors in order to explore their effects on the uncertainty (or heterogeneity) of vegetation systems.

However, as influenced by Orlóci (1968, 1978), I rather applied the formula of mutual information to C(M, M) to describe the spatial pattern in the grasslands of the Karst area (Feoli et al. 1980) and to measure the similarity between vegetation types corresponding to dynamical states of the coastal vegetation described by Pignatti (1960) (Feoli & Scoppola 1980).

In the present paper, I reconsider formula (2) and use some ad hoc “toy matrices” X(M, N) and matrices obtained from simple “toy graphs” to demonstrate its capacity in quantifying the pattern of graphs with respect to other formulas, e.g.:

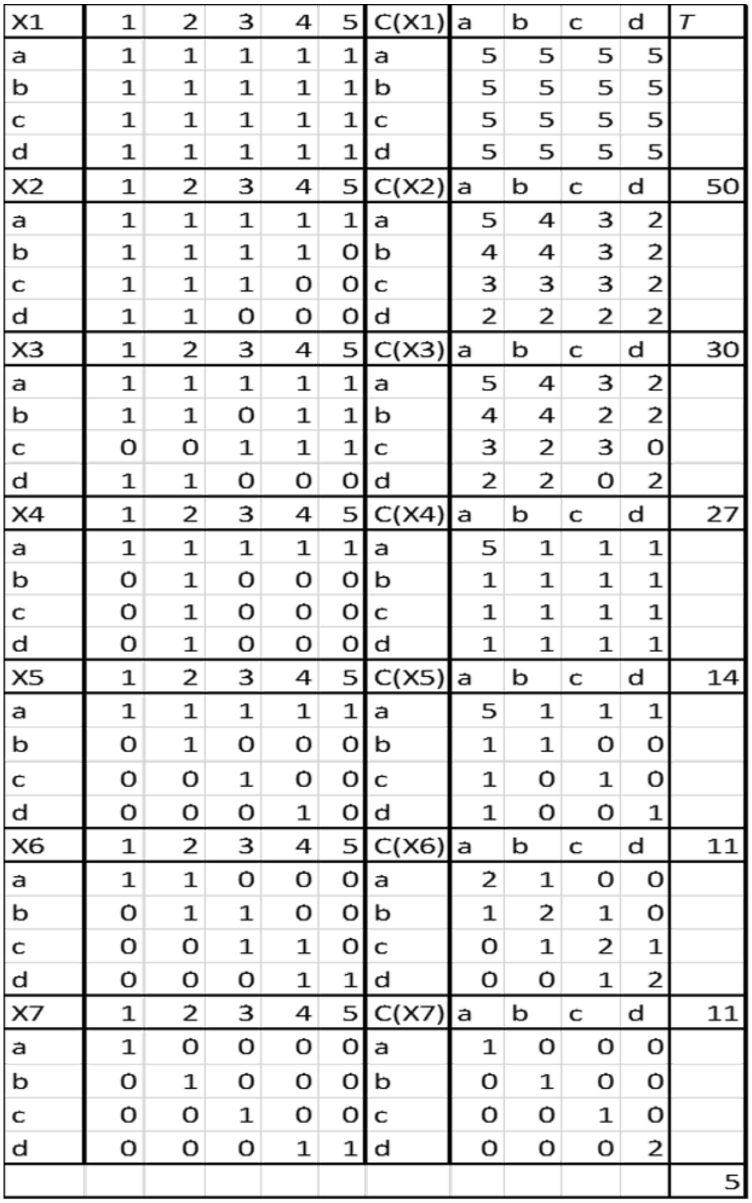
a. the ratio (H(D)+H(T))/ ln ((M(M+1)/2), that is, the relative entropy or evenness of the upper or lower triangular part of C(M, M) with diagonal D included (i.e., D+T);
b. the formula of Ricotta and Szeidl (2006), which incorporates in H(D) the entropy of the matrix of co-occurrences itself with values relativized between 0 and 1;
c. the evenness of the eigenvalues of C(M, M) that gives a measure of the independence of the M components (Feoli and Ganis 2021 and references therein); can components be independent??
d. the chi square of C(M, M), that according to Kulback (1959), approximates twice the mutual information (see Orloci 1978, Feoli et al. 1984 for ecological applications), and thus also expresses the independence of the M components;
e. the average nested free and
f. the nested based similarity (cf. Urlich et al. 2009; Feoli et al. 2019 and references therein) between the N vectors of X(M, N) that are indirect measures of the connections between the M components. It is clear that the similarity between the N sampled states of the system depends on the pattern of co-occurrence between the components (it is trivial to say that the higher the average similarity, the higher is the average co-occurrence); and
g. the number of centres of the graphs corresponding to the co-occurrence matrices; the maximum number of centres is M, i.e., when all the M components are connected with more than one component.

Data

To show the performance of formula (2), I used two sets of data; the first is presented in Table 1. It consists of 7 small “toy” matrices (X (M, N)), each with 4 rows (a, b, c, d) that are the components of the systems and with 5 columns representing the states of the systems. These matrices, X1,..., X7, give different tipologies of pairwise co-occurrence matrices, abbreviated by C(X1), ..., C(X7) in Table 1. In this case, the nodes and edges of the corresponding graphs are weighted by the values obtained by the multiplication of each matrix by itself (formula 1)). For each matrix, the column total that is entering in formula 2) to obtain the p_{ii} and p_{ih} for calculating the entropies H(D) and H(T) is indicated.

If we consider pairwise co-occurrence as an expression of the connection between the nodes representing the components in a graph, we can say that matrix C(X1) represents a complete graph with the nodes and edges of the same weight; C(X2), corresponding to a matrix X(M, N) with a ”gradual” nested situation, represents a graph where the connection is still complete, but the nodes and edges have different weights; C(X3) represents a graph not completely connected, where the nodes of the graph have the same weight as those of C(X2), but the edges have different weights; C(X4) represents an extreme case of nestedness in which one component is linked to all others, those only with degree one; in this case, the connection between nodes is complete but weaker than those of the previous matrices; C(X5) represents a situation in which there is the minimal connection (M-1) between the nodes and where only node *a* is linked to all the others (strong centrality in the graph); C(X6) represents a situation in which there is the minimal connection between the nodes as in C(X5), but the nodes are connected only with 2 nodes at most. In this case, the graph is assuming the shape of a chain (these two last matrices represent two situations of the minimal connected graph as those of the graphs of Fig. 1 but with different centrality); C(X7) represents a completely disconnected situation, i.e., a “graph with only nodes.” In terms of beta diversity of Whittaker (1972), i.e., the ratio between M and m, where m is the average number of components in the matrices X(M, N), it is the following: X1=1, X2=1,43, X3=1,54, X4=2,5, X5=2,5, X6=2,5 and X7=4. Three matrices have the same beta diversity (B=2.5), but they present quite different patterns.

The second data set is derived from the blackboard in Fig. 1, whose graphs were drawn by Will Hunting (a young mathematical genius interpreted by Matt Damon) in the movie “Good Will Hunting.” These graphs were chosen because I realized that they all have the same number of connections (M-1) (the edges of the graph), which represents the minimal connection, but they are all different as far as the degree of the nodes is concerned. They give rise easily (for this reason, I do not show them) to X(10,9) and C(10,10) matrices; those last ones are with the same trace (Dt=18) and the same total of the upper or lower triangular part of the co-occurrence matrix (Tt =9). Beside the co-occurrence matrices of these graphs (G), I considered also the co-occurrence matrix (the number 6 in Table 2) obtained by the graph corresponding to the one in which each node has at most only two edges, i.e., it is connected with only two nodes (G6); topologically, it corresponds to the graph of matrix X6 in Table 1. Whittaker's beta diversity of the X(10,9) matrices is equal to 5 for all of them, the average number of components all being equal to 2.



X1	1	2	3	4	5	C(X1)	a	b	c	d	*T*
a	1	1	1	1	1	a	5	5	5	5	
b	1	1	1	1	1	b	5	5	5	5	
c	1	1	1	1	1	c	5	5	5	5	
d	1	1	1	1	1	d	5	5	5	5	
X2	1	2	3	4	5	C(X2)	a	b	c	d	50
a	1	1	1	1	1	a	5	4	3	2	
b	1	1	1	1	0	b	4	4	3	2	
c	1	1	1	0	0	c	3	3	3	2	
d	1	1	0	0	0	d	2	2	2	2	
X3	1	2	3	4	5	C(X3)	a	b	c	d	30
a	1	1	1	1	1	a	5	4	3	2	
b	1	1	0	1	1	b	4	4	2	2	
c	0	0	1	1	1	c	3	2	3	0	
d	1	1	0	0	0	d	2	2	0	2	
X4	1	2	3	4	5	C(X4)	a	b	c	d	27
a	1	1	1	1	1	a	5	1	1	1	
b	0	1	0	0	0	b	1	1	1	1	
c	0	1	0	0	0	c	1	1	1	1	
d	0	1	0	0	0	d	1	1	1	1	
X5	1	2	3	4	5	C(X5)	a	b	c	d	14
a	1	1	1	1	1	a	5	1	1	1	
b	0	1	0	0	0	b	1	1	0	0	
c	0	0	1	0	0	c	1	0	1	0	
d	0	0	0	1	0	d	1	0	0	1	
X6	1	2	3	4	5	C(X6)	a	b	c	d	11
a	1	1	0	0	0	a	2	1	0	0	
b	0	1	1	0	0	b	1	2	1	0	
c	0	0	1	1	0	c	0	1	2	1	
d	0	0	0	1	1	d	0	0	1	2	
X7	1	2	3	4	5	C(X7)	a	b	c	d	11
a	1	0	0	0	0	a	1	0	0	0	
b	0	1	0	0	0	b	0	1	0	0	
c	0	0	1	0	0	c	0	0	1	0	
d	0	0	0	1	1	d	0	0	0	2	
											5

Table 1. Seven “toy” matrices X(M, N): X1,.., X7, giving rise to seven different co-occurrence matrices C (M, M): C(X1),…, C(X7) of the components a, b, c, d (see text).

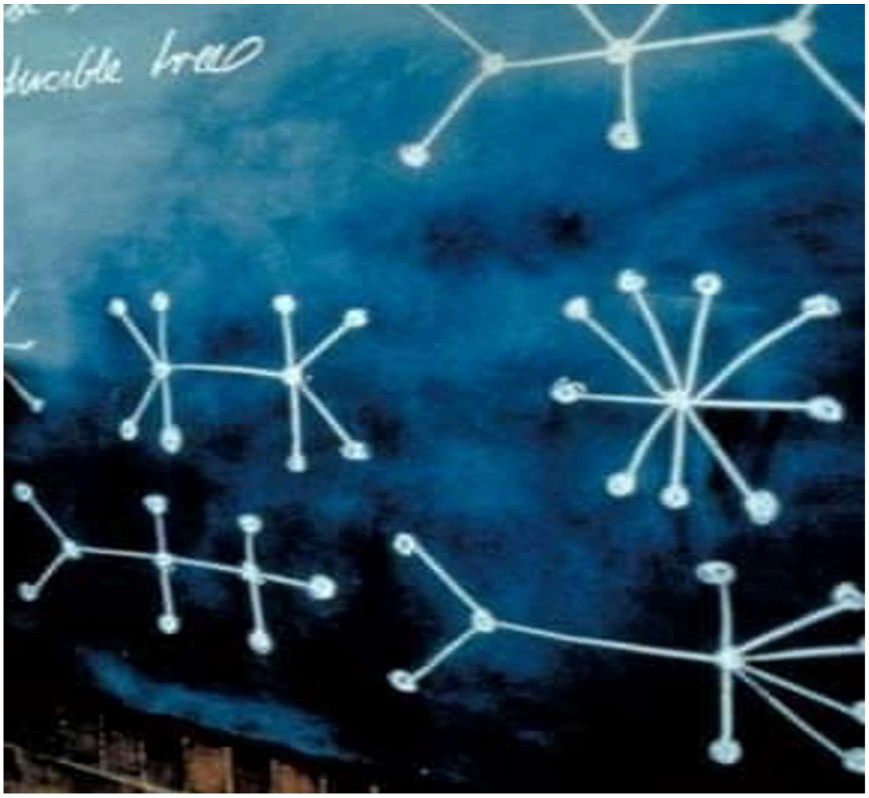


Figure 1. This figure shows the 5 graphs (G) drawn by Will Hunting (a young mathematical genius interpreted by Matt Damon) in the movie “Good Will Hunting”, from which I obtained 5 co-occurrence matrices. They are expressed as 10 by 9 data matrices and as 10 by 10 co-occurrence matrices that can be easily written. All of them have 10 nodes and 9 edges. To these graphs, I have added the one (G 6 in Table 2) with 8 centres, i.e., the one with less centrality. G1, with 3 centres, is on the top right, although partially visible. In the middle row are G2 with 2 centres on the left and G 3 with only one centre on the right. On the bottom, G4 has 3 centres and G5 has 2 centres.

Methods

All the toy matrices of co-occurrence obtained from the X matrices in Table 1 and from the graphs of Fig. 1 represent different patterns of connection between the nodes of the corresponding graphs (as said before, only the patterns of the graph of matrix X6 and that of G6 of Table 2 are topologically similar, with each node in connection with at maximum 2 nodes). What we could expect from a good formula of connectance is that it would give different values for each different “graph.” Formula 2) quantifies in a weighted way the connections within a system: the lower the entropy, the higher the negentropy, the higher the connectance should be. This can be obtained by dividing the observed negentropy by the value of negentropy corresponding to the situation in which there is the highest number of connections. This corresponds to the graph with all connected nodes and with the same weight (or equivalence) of nodes and edges (Table 1, matrix X1). It is easy to show that the maximum value of negentropy is given by the following formula:

$$\mathrm{H(C)\ ref\ max} = -((2/(\mathrm{M}+1))\ \ln(2/\mathrm{M}(\mathrm{M}+1)) - ((\mathrm{M}-1)/(\mathrm{M}+1))\ \ln(2/(\mathrm{M}(\mathrm{M}+1))) \qquad (3)$$

The ratio between 2) and 3) could be considered a relative weighted measure of connectance (based on co-occurrence):

$$\mathrm{K(C)} = \mathrm{H(C)/H(C)\ ref\ max} \qquad (4)$$

K(C) can be positive or negative; it is negative when H(C) is positive, i.e., when the entropy of the diagonal of the co-occurrence matrix is higher than the entropy of the upper or lower triangular part of the matrix (i.e., H(D)>H(T)).

In the paper of 1972, I proposed to calculate the maximum entropy of a community data matrix **X**(M, N) by the formula

$$\mathrm{Hmax} = -\ \mathrm{m\ ln\ m/M} \qquad (5)$$

where m is the average number of components for the N communities in **X**(M, N). I suggested this because it is unrealistic to think that a community would have only one component. In other words, each vegetation system or/and subsystem would have an entropy that depends on its average number of components in the stands by which it is described. The closer the average number of components is to the total number, the more homogeneous the vegetation system described by X(M, N), and the lower the entropy, or uncertainty, relative to the system's definition. The uncertainty of a system of completely connected components would be zero. This formula is in fact equal to 0 when m=M, and it is ln M when m=1. If it is used to calculate the redundancy of a matrix (R(**X**(M, N)) as I suggested in 1972:

$$\mathrm{R(X(M,\ N)} = 1\text{-}\ \mathrm{H(C)/(Hmax)} \qquad (6)$$

The values of redundancy R(**X**(M, N)) would become higher than 1 for H(C) negative (this happens when H(T) is higher than H(D)), and it will approach 2 as the absolute value of H(C) approaches Hmax. If Hmax is 0, i.e., when m=M, then H(C) is equal to H(C) ref max (Formula 3), and R(X(M, N)) has an indeterminate value. In this case, the redundancy is maximum. This can be viewed as a drawback of the formula; however, I think it useful to be compared with values obtained by K(C). It should be clear that when we are dealing with the triangular part of matrices of co-occurrence given by D+T, their maximum entropy values are equal to:

$$\mathrm{H\ max\ (D+T) = ln\ M(M+1)/2} \qquad (7)$$

So I compared K(C) also with the ratio:

$$\mathrm{H(D+T)\ rel = H(D+T)/Hmax(D+T)} \qquad (8)$$

H(C) was compared also with the formula of Ricotta and Szeidl (2006) that introduces in Shannon's formula the matrix of co-occurrence ρ between the components where the values of co-occurrence are transformed to the range [0,1]:

$$\mathrm{D(O)} = -\Sigma_{\mathrm{i}}\mathrm{p_i}\ \ln(\Sigma\ j \neq i\ \rho \mathrm{ijp_j}) \qquad (9)$$

In this equation, $\mathrm{p_i}$ is one element of the vector P (1, M) that shows the proportions between the weights of the nodes, ρ_{ij} is the ij-th element in the co-occurrence matrix C(M, M), which is obtained by the Jaccard coefficient (Podani 2022) applied between the M components of the matrix X(M, N), $\mathrm{p_j}$ is the component j-th of the vector of the proportions of the weight of the M nodes P $(1, \mathrm{M})^{\mathrm{T}}$ transposed, i.e., P(M,1).

In summary, I have considered for each of the co-occurrence matrices, those obtained by the matrices X (Table 1) and those obtained by graphs (Fig. 1), the following parameters as listed in Table 2 of the results:

1. the H(C) obtained by formula 2). It indicates in absolute terms how much the entropy of the edges H(T) is higher with respect to the entropy H(D) of the nodes: the higher the negative values in absolute terms, the higher the connection of the M components;
2. the H(C) ref max given by formula 3);
3. the K(C) obtained as the ratio between H (C) and H(C) ref max according to formula 4). It indicates a relative measure of connection (connectance weighted by nodes and edges) between the M components; the higher its value, the higher the connectance;
4. the redundancy (R) according to formula 6);
5. the H(D), i.e., the entropy of the diagonal D of the co-occurrence matrices;
6. the H(T), i.e., entropy of the upper or lower triangular part T of the co-occurrence matrices of Table 1 and those obtained by graphs of Fig. 1;

7. the H(D+T), i.e., the entropy of the triangular matrix obtained by H(D)+H(T);
8. H(D+T) max, i.e., the maximal entropy of the triangular matrices obtained by (D + T), obtained by formula 7);
9. the ratio H(D+T)/(H(D+T) max) (formula 8)), i.e., the relative entropy of matrix (D+T);
10. the entropy of Ricotta and Szeidel's (RISZ) by formula 9);
11. the evenness of the eigenvalues (Eλ) of the co-occurrence matrices (Feoli & Ganis 2021), it indicates how much the values of the matrix are concentrated on the diagonal. It is maximum when there are no connections (no edges) and the nodes have the same weight. In this case, the diagonal values are directly the eigenvalues of the matrices, and the neg-entropy of the eigenvalues is obviously equal to the entropy of the nodes based on their weight. The evenness becomes lower as the nodes are connected;
12. the chi-square of the co-occurrence matrices which, as shown by Kullback (1959), approximates twice the mutual information (cf. Orlóci 1978, Feoli et al. 1984 for applications in phytosociology). As in the case of Eλ, it is maximum when there are only values in the diagonal of the matrix and they are all equal; the chi-square is zero when the co-occurrence matrix has all equal values;
13. the number of components (M);
14. the average number of components (m);
15. Whittaker's Beta Diversity (Whittaker 1972), i.e., the ratio between M and m;
16. the nested free average similarity (NFS); and
17. the nested based similarity (NBS) according to Feoli et al. (2019). In both cases 16) and 17), the similarity has been calculated by considering the Jaccard similarity function (Podani 2022). It is to stress that for vectors given by only two components, such as those obtained by the graphs in Fig. 1, the two values (NBS and NFS) are identical;
18. and finally, the number of centres of each graph obtained by all the co-occurrence matrices.

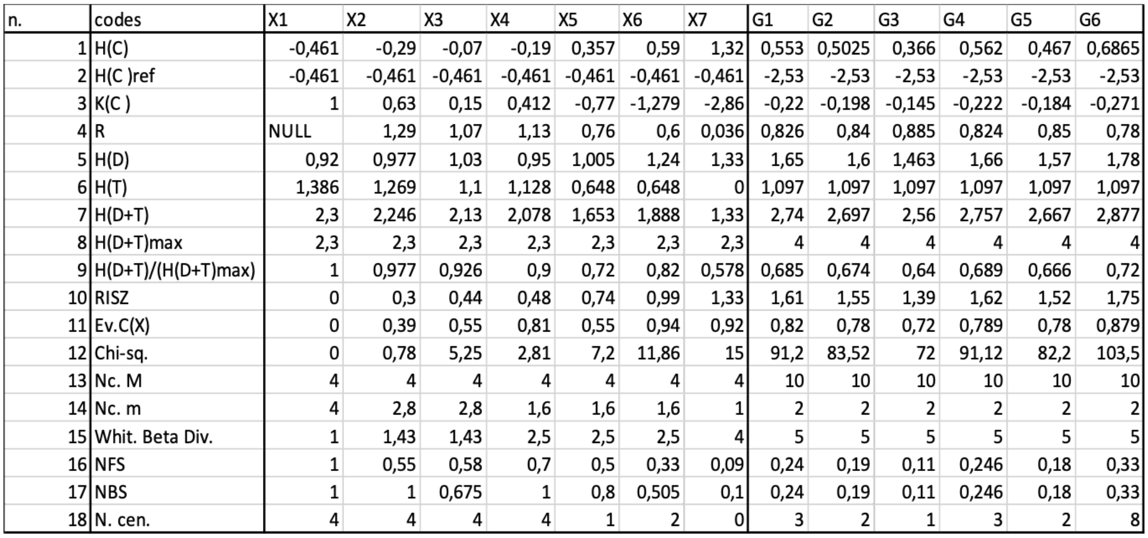
Between these parameters, I have calculated the Pearson correlation coefficient, and to the correlation matrix, I have applied the ranking procedure suggested by Orloci (1978) for weighting variables. I did this procedure following the thought that a good parameter, among those chosen to describe the connection pattern of graphs, should be the one that would explain most of the variability in the other parameters. Finally, I calculated the similarity between the graphs of the toy matrices and between the toy graphs of Fig. 1, both on the basis of all the parameters, excluding those that are constant, and considering only the single parameter which explains the maximum cumulative variance with respect to all the others, by the

doi.org/10.32388/T9WW4P

index of Gower (see Podani 2000), and I classified separately the graphs of the 7 matrices and those of Fig. 1. All the computations were done by the program MATEDIT (Burba et al., 2008).

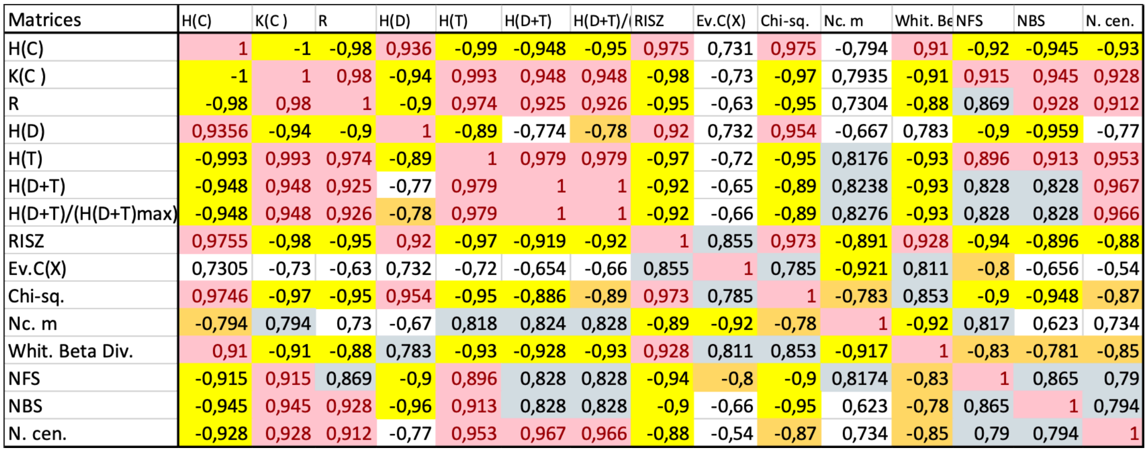
Results

Table 2 shows the values of the 18 parameters calculated for each of the 7 matrices (X) in Table 1 and for the co-occurrence matrices obtained for the graphs (G). Table 3 and Table 4 show the Pearson correlation coefficients between the considered parameters (the constants are excluded). Table 5 shows the results of the ranking procedure applied to the matrices in Tables 3 and 4. It is clear from these tables that the correlation between H(C) and K(C) with all the other parameters is significant at a very high level of probability, i.e., 0.01 (pink colour for positive correlation, yellow colour for negative ones). In proportion, the significative correlations are more frequent for the co-occurrence matrices obtained by graphs; this can be explained by the fact that graphs give a more homogeneous set concerning the connections, which are in all cases 9.



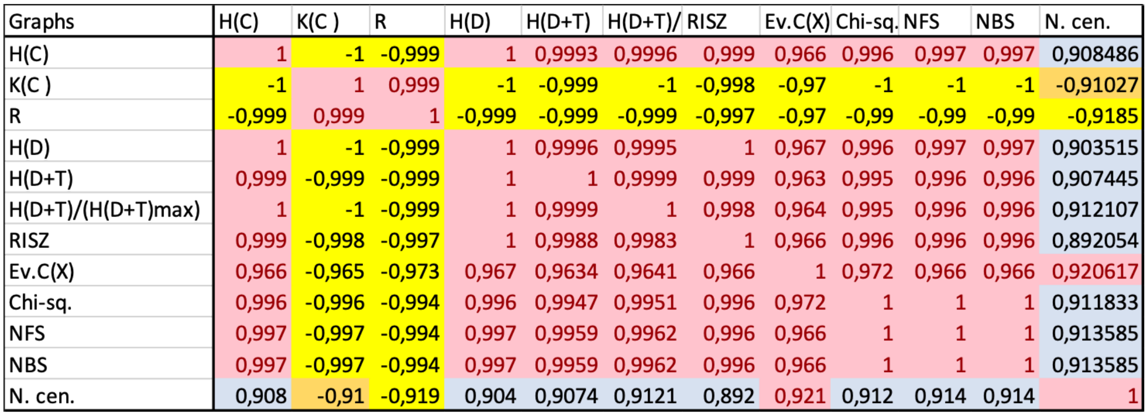
n.	codes	X1	X2	X3	X4	X5	X6	X7	G1	G2	G3	G4	G5	G6
1	H(C)	-0,461	-0,29	-0,07	-0,19	0,357	0,59	1,32	0,553	0,5025	0,366	0,562	0,467	0,6865
2	H(C)ref	-0,461	-0,461	-0,461	-0,461	-0,461	-0,461	-0,461	-2,53	-2,53	-2,53	-2,53	-2,53	-2,53
3	K(C)	1	0,63	0,15	0,412	-0,77	-1,279	-2,86	-0,22	-0,198	-0,145	-0,222	-0,184	-0,271
4	R	NULL	1,29	1,07	1,13	0,76	0,6	0,036	0,826	0,84	0,885	0,824	0,85	0,78
5	H(D)	0,92	0,977	1,03	0,95	1,005	1,24	1,33	1,65	1,6	1,463	1,66	1,57	1,78
6	H(T)	1,386	1,269	1,1	1,128	0,648	0,648	0	1,097	1,097	1,097	1,097	1,097	1,097
7	H(D+T)	2,3	2,246	2,13	2,078	1,653	1,888	1,33	2,74	2,697	2,56	2,757	2,667	2,877
8	H(D+T)max	2,3	2,3	2,3	2,3	2,3	2,3	2,3	4	4	4	4	4	4
9	H(D+T)/(H(D+T)max)	1	0,977	0,926	0,9	0,72	0,82	0,578	0,685	0,674	0,64	0,689	0,666	0,72
10	RISZ	0	0,3	0,44	0,48	0,74	0,99	1,33	1,61	1,55	1,39	1,62	1,52	1,75
11	Ev.C(X)	0	0,39	0,55	0,81	0,55	0,94	0,92	0,82	0,78	0,72	0,789	0,78	0,879
12	Chi-sq.	0	0,78	5,25	2,81	7,2	11,86	15	91,2	83,52	72	91,12	82,2	103,5
13	Nc. M	4	4	4	4	4	4	4	10	10	10	10	10	10
14	Nc. m	4	2,8	2,8	1,6	1,6	1,6	1	2	2	2	2	2	2
15	Whit. Beta Div.	1	1,43	1,43	2,5	2,5	2,5	4	5	5	5	5	5	5
16	NFS	1	0,55	0,58	0,7	0,5	0,33	0,09	0,24	0,19	0,11	0,246	0,18	0,33
17	NBS	1	1	0,675	1	0,8	0,505	0,1	0,24	0,19	0,11	0,246	0,18	0,33
18	N. cen.	4	4	4	4	1	2	0	3	2	1	3	2	8

Table 2. This table shows the values of the 18 parameters for each of the matrices Xi in Table 1 (matrices) and for the matrices obtained from the graphs of Fig. 1 (graphs G). Symbols or abbreviations: 1) H (C)= entropy of the co-occurrence matrices given by formula 2); 2) H(C) ref max given by formula 3); 3) K(C)= ratio between H (C) and H(C) ref max according to formula 5); 4) R= redundancy according to formula 7); 5) H(D) = entropy of the diagonal D of the co-occurrence matrices; 6) H(T) = entropy of the upper or lower triangular part T of the co-occurrence matrices; 7) H(D+T)= entropy of the triangular matrix obtained by D+T; 8) H(D+T) max= the maximal entropy of the triangular matrices obtained by (D + T); 9) H(D+T)/(H(D+T) max)= relative entropy of matrix (D+T); 10) RISZ= Ricotta and Szeidel's entropy formula (2006); 11) Ev.C(X)=evenness of the eigenvalues of the co-occurrence matrices (Feoli and Ganis 2021); 12) Chi-sq.= Chi-square of the co-occurrence matrices; 13) Nc.M= number of components M; 14) Nc. m= average number of components; 15) Whit.Beta Div = Whittaker's Beta Diversity =.; 16) NFS= nested free average similarity; 17) NBS= nested based similarity; 18) N. cen.= number of centres.



Matrices	H(C)	K(C)	R	H(D)	H(T)	H(D+T)	H(D+T)/(	RISZ	Ev.C(X)	Chi-sq.	Nc. m	Whit. Be	NFS	NBS	N. cen.
H(C)	1	-1	-0,98	0,936	-0,99	-0,948	-0,95	0,975	0,731	0,975	-0,794	0,91	-0,92	-0,945	-0,93
K(C)	-1	1	0,98	-0,94	0,993	0,948	0,948	-0,98	-0,73	-0,97	0,7935	-0,91	0,915	0,945	0,928
R	-0,98	0,98	1	-0,9	0,974	0,925	0,926	-0,95	-0,63	-0,95	0,7304	-0,88	0,869	0,928	0,912
H(D)	0,9356	-0,94	-0,9	1	-0,89	-0,774	-0,78	0,92	0,732	0,954	-0,667	0,783	-0,9	-0,959	-0,77
H(T)	-0,993	0,993	0,974	-0,89	1	0,979	0,979	-0,97	-0,72	-0,95	0,8176	-0,93	0,896	0,913	0,953
H(D+T)	-0,948	0,948	0,925	-0,77	0,979	1	1	-0,92	-0,65	-0,89	0,8238	-0,93	0,828	0,828	0,967
H(D+T)/(H(D+T)max)	-0,948	0,948	0,926	-0,78	0,979	1	1	-0,92	-0,66	-0,89	0,8276	-0,93	0,828	0,828	0,966
RISZ	0,9755	-0,98	-0,95	0,92	-0,97	-0,919	-0,92	1	0,855	0,973	-0,891	0,928	-0,94	-0,896	-0,88
Ev.C(X)	0,7305	-0,73	-0,63	0,732	-0,72	-0,654	-0,66	0,855	1	0,785	-0,921	0,811	-0,8	-0,656	-0,54
Chi-sq.	0,9746	-0,97	-0,95	0,954	-0,95	-0,886	-0,89	0,973	0,785	1	-0,783	0,853	-0,9	-0,948	-0,87
Nc. m	-0,794	0,794	0,73	-0,67	0,818	0,824	0,828	-0,89	-0,92	-0,78	1	-0,92	0,817	0,623	0,734
Whit. Beta Div.	0,91	-0,91	-0,88	0,783	-0,93	-0,928	-0,93	0,928	0,811	0,853	-0,917	1	-0,83	-0,781	-0,85
NFS	-0,915	0,915	0,869	-0,9	0,896	0,828	0,828	-0,94	-0,8	-0,9	0,8174	-0,83	1	0,865	0,79
NBS	-0,945	0,945	0,928	-0,96	0,913	0,828	0,828	-0,9	-0,66	-0,95	0,623	-0,78	0,865	1	0,794
N. cen.	-0,928	0,928	0,912	-0,77	0,953	0,967	0,966	-0,88	-0,54	-0,87	0,734	-0,85	0,79	0,794	1

Table 3. Correlations between the parameters of the matrices (X1,...X7) in Table 2 (p=0.01, r= 0.875; p=0.05, r=0.775). The 3 constants, H(C) ref, H(D+T) max, and Nc.M, are excluded from the table. Negative and positive correlations for probability lower or equal to 0.01 are respectively in yellow and pink; correlations for p lower or equal 0.05 and greater than 0.01 are in light blue.



Graphs	H(C)	K(C)	R	H(D)	H(D+T)	H(D+T)/	RISZ	Ev.C(X)	Chi-sq.	NFS	NBS	N. cen.
H(C)	1	-1	-0,999	1	0,9993	0,9996	0,999	0,966	0,996	0,997	0,997	0,908486
K(C)	-1	1	0,999	-1	-0,999	-1	-0,998	-0,97	-1	-1	-1	-0,91027
R	-0,999	0,999	1	-0,999	-0,999	-0,999	-0,997	-0,97	-0,99	-0,99	-0,99	-0,9185
H(D)	1	-1	-0,999	1	0,9996	0,9995	1	0,967	0,996	0,997	0,997	0,903515
H(D+T)	0,999	-0,999	-0,999	1	1	0,9999	0,999	0,963	0,995	0,996	0,996	0,907445
H(D+T)/(H(D+T)max)	1	-1	-0,999	1	0,9999	1	0,998	0,964	0,995	0,996	0,996	0,912107
RISZ	0,999	-0,998	-0,997	1	0,9988	0,9983	1	0,966	0,996	0,996	0,996	0,892054
Ev.C(X)	0,966	-0,965	-0,973	0,967	0,9634	0,9641	0,966	1	0,972	0,966	0,966	0,920617
Chi-sq.	0,996	-0,996	-0,994	0,996	0,9947	0,9951	0,996	0,972	1	1	1	0,911833
NFS	0,997	-0,997	-0,994	0,997	0,9959	0,9962	0,996	0,966	1	1	1	0,913585
NBS	0,997	-0,997	-0,994	0,997	0,9959	0,9962	0,996	0,966	1	1	1	0,913585
N. cen.	0,908	-0,91	-0,919	0,904	0,9074	0,9121	0,892	0,921	0,912	0,914	0,914	1

Table 4. Correlations between the parameters of the graphs in Table 2 (p=0.01, r=0.917; p=0.05, r=0.811). The constants, H(C) ref, H(T), H(D+T) max, Nc.M, Nc.m, Whit. Beta Div., are excluded from the table. Negative and positive correlations for probability lower or equal to 0.01 are in yellow and pink, respectively; correlations for probability lower or equal 0.05 and greater than 0.01 are in light blue.

By comparing the two matrices, we can see that the graphs of Fig. 1 show that the correlations between the parameters are all significant, at least at the level of probability of 0.05, while the correlations between the parameters calculated for Table 1 are not all significant. In both cases, 5 parameters are enough to explain 100% of the total variance.

Rank Matrices	Sum of sq.	% spec.	% cum.	Rank Graphs	Sum of sq.	% spec.	% cum.
H(C)	13,1051	87,3674	87,3674	R	11,7386	97,822	97,8219
Nc. m	0,9429	6,2865	93,654	N. cen.	0,1675	1,3962	99,2181
K(C)	0,7886	4,2573	97,911	Ev.C(X)	0,05221	0,4351	99,65
Ev.C(X)	0,0622	1,1526	99,06	H(D+T)/(H(D+T)max)	0,0402	0,3353	99,9886
RISZ	0,0231	0,9395	100	NFS	0,0014	0,0114	100

Table 5. Results of the ranking procedure suggested by Orlóci (1978), applied respectively to the matrix of Table 3 (Rank Matrices) and to the matrix of Table 4 (Rank Graphs). Sum of sq. = Sum of squares; %spec. = sum of squares specific to the single parameter with respect to the others; % cum. = cumulative sum of squares.

From Table 5, it is clear that for the matrices of Table 1, H(C) is the most redundant parameter, i.e., the one that explains better than others the pattern of the graphs, while for the graphs of Fig. 1 and Table 2, the most redundant parameter is the redundancy R (X(M, N)) of formula 6), which explains better the variability of the others in terms of sum of squares.

The classification of the 7 matrices X1, ..., X7 on the basis of all the parameters and on the basis of H(C), that is, the parameter with the highest specific sum of squares, is presented by the dendrograms of Fig. 2 a) and b). These are obtained by complete linkage clustering and show almost equal topology both for a) and b); only the position of matrices X3 and X4 is changed within the same main cluster (X2, X3, X4). This proves that H(C) is a good parameter that can explain all the variability of the other parameters. The classification of the graphs of Fig. 1 based on all the parameters and only on the redundancy (R) of formula 6) shows dendrograms c) and d) that are topologically identical. In any case, thanks to the very high correlation between R(X(M, N)) and H(C) (see Tables 3 and 4), the dendrogram e), obtained by using H(C), is topologically identical to dendrograms c) and d).

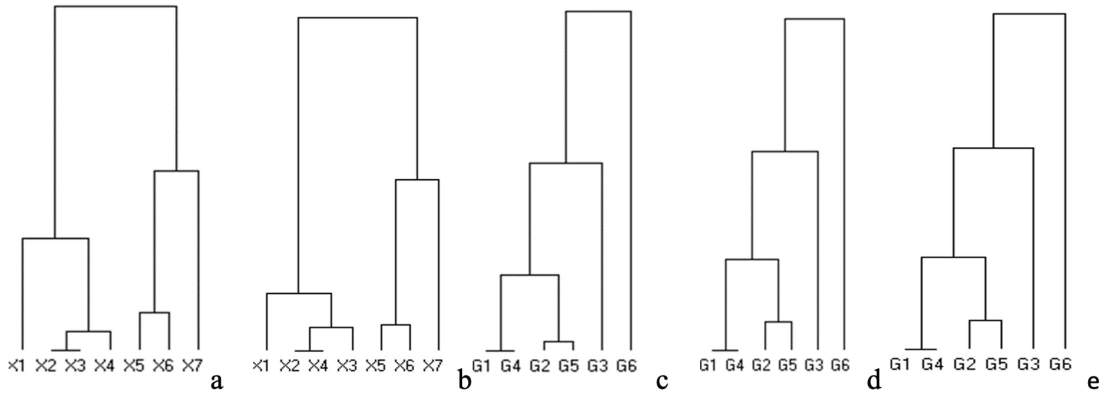


Figure 2. Classification of the seven graphs originated by the matrices X of Table 1 by using all the parameters in Table 2 (a) and by using only H(C) (b), and classification of the graphs of Fig. 1 and Table 2 by using all the parameters (c) and by using only the redundancy of formula 6) (d). The dendrogram (e) shows the classification of the six graphs (G) using only H(C).

Discussion and Conclusion

- *Co-occurrence matrices, graphs, connections, connectivity, connectance, and negentropy as a conceptual tool to define community patterns*

I think that, today, and among ecologists, it is trivial to mention that: "What makes a system a system, and not simply a collection of elements, is the connections and interactions between its components, as well as the effect that these linkages have on its behaviour," but I consider it useful to recall the concept anyway (https://www.britannica.com/science/complexity-scientific-theory/Connectivity).

In general terms, every "entity" can be considered a system when it is seen as a set of connected components by one or all of the following relationships: contact, exchange of materials and energy, and exchange of information. See Allen and Starr (1982) for a discussion of connectedness, connectivity, and connectance in the context of complex systems. I want to stress here that systems may be open or closed, static or dynamic, and the connections between their components may be direct or indirect, i.e., via interposed components.

When we analyse ecological communities, we have to take into consideration that the concept of coexistence is strictly related to the spatial and temporal scales. It is obvious that if I considered sampling units of a few cm^2, the co-occurrence would be realized only for the microbes or small species that are found in these sampling units; if I considered sampling units of 1 km^2, I would find in them several species of microbes, plants, and animals of different sizes. It is instructive to visit the web with the keywords: scale, species-area curves, biodiversity, to have a lot of information on the area–species richness relationship. The time interval is also very important, especially in the study of animal communities. If I stay in a site just for a few hours, I can see very few animals, but if I observe the area for days, then I can record several animals that are visiting the area for grazing or hunting. Even for plants, time could be important; annual species are "visible" in some periods of the year and not in others.

Having said this, I want to stress also that when we complete a list of species or "traits" of a given area (a sampling unit, stands, traps, etc.), we have originated a vector of co-occurrence. This is the basic unit in the study of variation of ecological communities with respect to environmental heterogeneity and/or gradients. It may be considered alone or together with other N-1 vectors that we could call diversity vectors of a matrix X(M, N). This matrix, with vectors j from 1 to N, can be considered the primary co-occurrence matrix that can be used to obtain a pairwise co-occurrence matrix by formula 1). The co-occurrence matrices C(M, M) may be considered as one of the quantitative expressions of the pattern of coexistence of the components of a system. It is obvious that just the co-occurrence of given M components of any type of system, as obtained from formula 1), would not explain the interactions between them; experiments and/or careful observations are necessary to understand the nature of co-occurrence (e.g., Bever 2003, Adler et al. 2018, and references therein). It is clear that a matrix of co-occurrence may have different meanings, but from a mathematical point of view, it is always a symmetric matrix that can be used to construct a graph according to graph theory (Diestel 2017). In this paper, I consider the pairwise co-occurrence matrix just as a mathematical expression, and I do not want to revise the vast literature on coexistence of species or other community components where the concept of co-occurrence is implicit. This literature spans from the biogeography of birds to different ecological aspects of food webs, soil microbiology, and landscape ecology (e.g., Diamond 1975, Pimm 1984; Jordan 2001, 2009; Jordan et al. 2007, Scotti and Jordan 2010; Heleno et al. 2012, Veech 2013, Poisot and Gravel 2014, van Altena et al. 2016, Turnbull et al. 2018, Liccari et al. 2022, and references therein). I do not want to review all the formulas and models related to co-occurrence, but I want just to say that the study of the connection between the system components led to the concept of connectivity (see Turnbull et al. 2018 for a deep discussion about its meanings in different disciplines), and to that of connectance (cf. Allen and Starr 1982). The first is an absolute measure of the "consistency" of a graph, the second is a relative measure of connections with respect to what I call the maximum "diversity of connections". With diversity of connections, I mean the combination between the richness of the nodes and the edges and the proportion of their importance. The "maximum diversity of connections" is realized when all the nodes of a graph are connected with all the others and nodes and edges have the same weight. The diagonal component (D) and the triangular upper or lower part of a C(M, M) matrix (T) are representing the triangular matrix of the "diversity of connections" with M(M+1)/2 maximum number of cells to be occupied by non-zero values. When we use a co-occurrence matrix to obtain a graph of the "connections" between the M components of a system, the nodes and edges are always weighted as a consequence of C(M, M) being the self-product of the matrix X(M, N). It is obvious that if X(M, N) is a binary matrix (Avena et al. 1981, Wilson 2012), C(M, M) is a symmetric matrix in which the diagonal is showing the frequency of each of the M components and the elements outside the diagonal are the frequency with which each component occurs with each other component (pairwise co-occurrence); in this context, I do not find the room to discuss about "null models" on which there is a plethora of papers (e.g., Gotelli 2000, 2001).

The novelty of my proposal in 1972, with respect to the applications of information theory in ecology, by the seminal works of Margalef (1958, 1968), Orlóci (1968), and Lausi (1970), is the fact that I suggested the formula of neg-entropy specifically for the upper or lower triangular part—diagonal included—of the pairwise co-occurrence matrices C(M, M), rather than for the usual single vectors X(M,1), as is done by the application of several indices of diversity. In this way, the formula 2) can be interpreted as a diversity measure based on Shannon's index that includes a quantification of the connection between the components of the communities and, if we consider the co-occurrence a way to express connections, the formula 4), which is a relativization of formula 2), is a measure of weighted "connectance." It is true that Margalef and Gutierrez (1983) suggested a method to incorporate "connectance" in the diversity measures; however, they do not consider X(M, N) matrices, but they applied a formula similar to that proposed by Rao (1982) to the matrix of the cross product of the M elements of a single diversity vector. In this way, the graph is supposed to be completely connected, and different weights are given to the nodes as a consequence of the scalar product between the values of the vector X(M,1). However, I do not think that we should accept the idea of complete connections between all the species of a community on the basis of a single diversity vector, being the ecological communities subjected to great variability along ecoclines and ecotones (Whittaker 1975, Odum and Barrett 2005).

The formula (2) is an expression of "negentropy" because it depends on H(T), which is a measure of the links or connections between the M components in an X(M, N) matrix, a direct measure of their interdependence and/or cohesion: the higher H(T) is, the higher the homogeneity of the system represented by X(M, N), and the lower its entropy, i.e., its uncertainty. It is important to stress that the interpretation of the entropy of a co-occurrence matrix is different from the interpretation of the entropy of a diversity vector X(M,1). The formula has positive values only when H(D)>H(T), it is equal to zero if H(D)=H(T), and it has a negative value when H(T)> H(D). It is clear that if we consider the negative values of entropy, they are expressing a neg-entropy, i.e., the contrary of entropy and thus the contrary of uncertainty. The smaller H(D) is and the higher H(T) is, the higher the neg-entropy of H(C)) and the smaller the entropy. In summary, if we consider the co-occurrence values between the M components of a system as measures of their connection, the entropy gives a value of such a connection in positive or negative terms: the lower the entropy, the higher the neg-entropy of the system, i.e., its "cohesion" and predictability. In this paper, I considered two other relevant parameters to measure the uncertainty of a co-occurrence matrix, such as H(D)+H(T)/ H(D+T) max and the Ricotta and Szeidl (2006) entropy (formula 9)), and two parameters of the "primary" matrices of co-occurrence X(M, N) that are the average nested similarity and the average free nested similarity. The Pearson correlation between them has shown that all these 4 parameters are highly correlated (positively or negatively), so the following question arises automatically: "Why choose the H(C)-based ones, e.g., K(C) or R?". The answer is simple: they put in evidence directly the difference between the entropy of the nodes and the entropy of the edges of the graphs corresponding to the co-occurrence matrices, and because they are showing the highest capacity for explaining the variability of the other parameters.

I think and I suggest that the co-occurrence between the living components of ecosystems could be calculated at different hierarchical levels and could be interpreted always as an expression of connections between the components. In community ecology (the study of ecological communities), the idea to characterize an ecological system by parametrizing the "links" between the species (or other characters: "traits") within the combinations that could characterize community types at different hierarchical levels is not yet very well explored; for this reason, I would like to send the interested reader to the views of Aleksandrova (1973) and Dale (2001) and references therein, since they explicitly address the idea of hierarchical combinations. With this, I want to conclude that a matrix of co-occurrence can be obtained from community tables irrespective of the hierarchical level of the N vectors of the matrices X(M, N) and that formula 2) can help in quantifying the pattern of the coexistence of species or "traits" at different hierarchical levels of ecological communities. In this respect, I would like to remember the paper of Wilson (2011), in which he summarizes twelve theories of "co-existence" for plant communities. I think that it would be interesting to challenge these theories in other types of ecological communities where the hierarchical pattern is not yet explored in a syntaxonomical sense, such as in phytosociology (Pignatti 1980, 1990), notwithstanding the stimulating book of Allen and Starr (1982) on hierarchy as a context for mathematical modelling.

Acknowledgements

Many thanks are to be given to Prof. Dr. Janos Podani for his comments and corrections.

References

- Aleksandrova VD. 1973 Russian approaches to classification of vegetation. In R.H Whittaker ed. “Ordination and classification of communities”. Dr. W.Junk Publishers, The Hague
- Allen TFH, Starr TB (1982) Hierarchy. Perspectives for Ecological Complexity. The University of Chicago Press. Chicago and London.
- Avena G, Blasi C, Feoli E, Scoppola A (1981) Measurement of the predictive value of species lists for species cover in phytosociological samples. Vegetatio 45:77–84.
- Barkman JJ (1979) The investigation of vegetation texture and structure. In MJA. Werger (ed.) "The study of Vegetation". Junk, The Hague, Boston..
- Bever JD (2003) Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. New Phytol 157:465–473
- Burba N, Feoli E, Malaroda M (2008) MATEDIT: A software tool to integrate information in decision making processes. In: Neves R, Baretta JW, Mateus M (eds) Perspectives on Integrated Coastal Management in South America. IST Press, Lisbon, Portugal
- Adler PB, Smull D, Beard KH, Choi RT, Furniss T, Kulmatiski A, Meiners JM, Tredennick AT, Veblen KE (2018) Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. Ecol Lett 21:1319–1329
- Box EO (1981) Macroclimate and plant forms: an introduction to predictive modelling in phytogeography. Task for Vegetation Science. 1. Dr. W. Junk Publisher, The Hague.
- Dale, M. B. 2001. Functional synonyms and environmental homologues: an empirical approach to guild delimitation. Community Ecology 2:67-79
- Diestel R (2017) Graph Theory. 5th ed. Springer Verlag, Heidelberg

- Diamond J (1975) Assembly of species communities. In Ecology and evolution of communities, M. Cody and J. Diamond (eds.). Belknap Cambridge, Massachusetts. p. 342-444
- Evans TS, Chen B (2022) Linking the network centrality measures, closeness and degree. COMMUNICATIONS PHYSICS | (2022) 5:172 | https://doi.org/10.1038/s42005-022-00949-5 | www.nature.com/commsphys
- Feoli E (1972) Rudimenti della teoria dell'informazione in fitosociologia. Inform Bot Ital 4:202-208
- Feoli E (1984) Some aspects of classification and ordination of vegetation data in perspective. Studia Geobot 4:7–21.
- Feoli E, Ganis P (2021) Similarity, classification and diversity "an Eternal Golden Braid" in quantitative vegetation studies. Fl Medit 31 (Special Issue):23-41. https://doi.org/10.7320/FlMedit31SI.023
- Feoli E, Scoppola A (1980) Analisi informazionale degli schemi di dinamica della vegetazione. Un esempio sul popolamento vegetale delle dune di Venezia. Giorn Bot Ital 114:227-236
- Feoli E, Feoli-Chiapella L, Ganis P, Sorge A (1980) Spatial pattern analysis of abandoned grasslands of the Karst region by Trieste and Gorizia. Studia Geobot 1 (1), 213-221
- Feoli E, Ganis P, Ibáñez JJ, R Pérez-Gómez (2019) On the use of nestedness-based similarity functions (NBSF) to classify and/or order operational geographic units (OGUs). Community Ecology 20 (3): 223-229
- Feoli E, Lagonegro M, Orlóci L (1984) Information Analysis of Vegetation Data. Dr. W. Junk Publishers, The Hague
- Feoli E, Zuccarello V (1994). Naivete of fuzzy system space in vegetation dinamics? Coenoses 9: 25–32.
- Gotelli N J (2000) Null model analysis of species co-occurrence patterns. Ecology 81:2606–2621.
- Gotelli N J (2001). Research frontiers in null model analysis. Global Ecology and Biogeography Letters 10:337–343
- Goberna M, Verdú M (2022) Cautionary notes on the use of co-occurrence networks in soil ecology. Soil Biology and Biochemistry. 166 (2022) 108534
- Gower J C (1971). A general coefficient of similarity and some of its properties. Biometrics, 27, 857–871
- Griffith D M, Veech JA, Marsh CJ (2016) cooccur: Probabilistic Species Co-Occurrence Analysis in R. Journal of Statistical Software. Volume 69, Code Snippet 2. doi: 10.18637/jss.v069.c02
- Heleno R, Devoto M , Pocock M (2012) Connectance of species interaction networks and conservation value: Is it any good to be well connected? Ecological Indicators 14 (1): 7-10
- Jordán F (2001). Seasonal changes in the positional importance of components in the trophic flow network of the Chesapeake Bay. J. Marine Syst. 27: 289-300.

- Jordán, F. 2009. Keystone species and food webs. Phil. Trans. R. Soc. B 364: 1733-1741.
- Jordan 2009; Jordán F (2009) Keystone species and food webs. Philos Trans R Soc Lond B Biol Sci 364(1524):1733-1741
- Jordán F, Benedek Z, Podani J (2007) Quantifying positional importance in food webs: a comparison of centrality indices. Ecological Modelling 205(1):270-275
- Kullback, S (1959) Information Theory and Statistics. Wiley, New York.
- Lausi D (1970). Die Logik der Pflanzensoziologischen Vegetaionanalyse- Ein Deutungsversuch. Berict uber das “Internat. Symposium” in Rinteln, Den Haag.
- Liccari F, Boscutti F, Bacaro G, Sigura M (2022) Connectivity, landscape structure, and plant diversity across agricultural landscapes: novel insight into effective ecological network planning. Journal of Environmental Management, 26 May 2022, 317:115358 https://doi.org/10.1016/j.jenvman.2022.115358 PMID: 35636109
- Krafta N JB, Godoy O, Levined JM (2015). Plant functional traits and the multidimensional nature of species coexistence. PNAS 112: 797–802
- Margalef R (1958) Information Theory in ecology. In: L. van Bertalanffyand Rapoport (eds), General Systems, Yearbook of the Society for General System Research 3: 36-71.
- Margalef R (1968) Perspectives in Ecological Theory. Univ, Chicago Press, Chicago.
- Margalef R, Gutierrez E (1983) How to Introduce Connectance in the Frame of an Expression for Diversity. The American Naturalist 121(5):601-607
- Miklos I, Podani J (2004) Randomization of presence-absence matrices: Comments and new algorithms. Ecology 85: 86–92
- Mucina L, Maarel van der E (1989) Twenty years of numerical syntaxonomy. Vegetatio 81:1-15
- Odum EP, Barrett GW (2005) Fundamental of Ecology. Fith Edition, Cengage Learning India Private Limited, 2005.
- Orlóci L (1968) Information analysis in phytosociology: partition, classification and prediction. J. Theor. Biol. 20:271-284.
- Orlóci L (1978) Multivariate Analysis in Vegetation Research. 2nd ed. Dr. Junk, The Hague.
- Orlóci L (2020). Statistical quantum ecology. Essay on the resonator complex model of the vegetation stand. SCADA Publishing, Canada.
- Orlóci L, Orlóci M (1985). Comparison of communities without the use of species: model and examples. Ann. Bot. 43:275–285.

- Pignatti S (1960) Ricerche sull'ecologia e sul popolamento delle dune del litorale di Venezia. Il popolamento vegetale. Bull Mus Civ St Nat Venezia 12:61–142
- Pignatti S (1980) Reflections on the phytosociological approach and the epistemological basis of vegetation science. – Vegetatio 42: 181-185.
- Pignatti S (1990) Towards a prodrome of plant communities. – J. Veg. Sci. 1: 425-426
- Pillar V, Orlóci L (1993) Character-Based Community Analysis: The Theory and an Application Program. SPB Academic Publishing bv, The Hague, The Netherlands.
- Pimm SL (1984) The Complexity and Stability of Ecosystems. Nature 307(5949):321-326
- Podani J (2000) Introduction to the Exploration of Multivariate Biological Data. Backhuys Publishers, Leiden
- Podani J (2022) The wonder of the Jaccard coefficient: from alpine floras to bipartite networks Fl. Medit. 31 (Special Issue): 105-123 https://doi.org/10.7320/FlMedit31SI.105
- Rao CR (1982) Diversity and dissimilarity measurements: a unified approach. Theor. Popul. Biol. 21:24–43.
- Ricotta C. and Szeidl L (2006). Towards a unifying approach to diversity measures: bridging the gap between Shannon entropy and Rao's quadratic index. Theor. Popul. Biol. 70:237–243
- Scotti M, Jordán F (2010) Relationships between centrality indices and trophic levels in food webs. Community Ecology 11:59-67
- Turnbull L, Hütt MT, Ioannides AA, Kininmonth S, Poeppl R, Tockner K, Bracken LJ, Keesstra S, Liu L, Rens Masselink R, Parsons AJ (2018) Connectivity and complex systems: learning from a multi-disciplinary perspective. Applied Network Science 3:11 https://doi.org/10.1007/s41109-018-0067-2
- Ulrich W, Almeida-Neto M and Gotelli NJ (2009) A consumer's guide to nestedness analysis. Oikos 118: 3- 17
- van Altena C, Hemerik L, de Ruiter PC (2016) Food web stability and weighted connectance: the complexity-stability debate revisited. Theoretical Ecology, 9:49-58. https://doi.org/10.1007/s12080-015-0291-7
- van der Maarel, E. 1975. The Braun-Blanquet approach in perspective. Vegetatio 30:213–219.
- VeechJ.A (2013) A probabilistic model for analysing species co-occurrence. Global Ecol. Biogeogr. 22: 252–260
- Whittaker RH (1972) Evolution and measurement of species diversity. Taxon 21:213-251
- Whittaker RH (1975) Community and Ecosystems. MacMillan Publishers, New York.

- Williams WT, Lambert JM (1959) Multivariate methods in plant ecology I. Association-analysis in plant communities. Journal of Ecology 47:83-101
- Wilson JB (2011) The twelve theories of co-existence in plant communities: the doubtful, the important and the unexplored. J. Veg. Sci. 22:184–195.
- Wilson, JB (2012) Species presence/absence sometimes represents a plant community as well as species abundances do, or better. J. Veg. Sci. 23:1013–1023.

Declarations

Funding: Logistic support from Department of Life Sciences, University of Trieste

Potential competing interests: No potential competing interests to declare.