Article

Fl. Medit. 31 (Special Issue): 23-41
https://doi.org/10.7320/FlMedit31SI.023

Enrico Feoli & Paola Ganis

Similarity, classification and diversity "an Eternal Golden Braid" in quantitative vegetation studies

Abstract

Feoli, E. & Ganis, P.: Similarity, classification and diversity "an Eternal Golden Braid" in quantitative vegetation studies. — Fl. Medit. 31 (Special Issue): 23-41. 2021. — ISSN: 1120-4052 printed, 2240-4538 online.

The role of similarity-dissimilarity matrices is discussed within a conceptual framework that shows the strict connections between similarity, classification and diversity in vegetation studies. Examples of application of the evenness of the eigenvalues of similarity matrices ($E(\lambda)$) to define classes, to measure correlation between biological communities and environmental factors and to measure diversity of vegetation systems as a parameter δ, by the formula $\kappa\delta$, where k is the number of classes and $\delta = E(\lambda)$, are given by considering two data sets regarding beech forests of the Italian peninsula.

Key words: beech forests, correlation, eigenanalysis, evenness, fuzzy sets, matrices, Occam's razor.

Introduction

Paraphrasing Douglas R. Hofstadter (1979, 1984) we can say that Similarity, Classification and Diversity are three concepts that in quantitative ecology constitute an "Eternal golden braid" as "Gödel, Escher and Bach" in the "world" of his book. Quantitative ecology is following Gödel when it is calculating similarity (dissimilarity) between ecosystem states (nothing is definable and explainable without comparisons), it follows Escher, when produces classes of ecosystem states (classification), sharp or fuzzy: from the indistinctiveness to distinct things (see the Escher's pictures "Liberation", "Methamorphosis II" and "Verbum" in the book of Hofstadter) and it follows Bach when it detects diversity, that in music is a property of a playable construct based on different notes and of combinations of different notes. It is not by chance that Pignatti (1982) introduces his flora of Italy by showing a piece of music of Bach (Concerto Italiano, Fig. 1). The definition of diversity of a system is the key to understand how the system is working or playing in the "immense game" of the nature in the sense of Margalef (1980) and Eigen & Winkler (1986).

After having quantified diversity of a given area on the basis of the species (alpha diversity) or on the basis of plant-animal communities (beta diversity), we are ready to calculate similarity between the classes of the hierarchical level for which we have calculated diversity and for defining new classes at higher hierarchical levels as new levels of diversity: we can walk up and down as in the picture of Escher at page 13 of the Italian edition of the book of Hofstadter (1984). The application of the three concepts: similarity, classification and diversity, in logical sequential loops, leads towards the knowledge generation about the ecosystems and ecosystem’s components. In this paper we reconsider in more detail the results we have obtained by Feoli & Ganis (2019) and we show how $E(\lambda)$ can be used to measure the diversity of the vegetation systems as suggested by Feoli (2018).

Similarity, classification, diversity and eigenvalues

Similarity

Similarity is a property of any object that emerges when it is compared with other objects. For the comparison we can use any set of characters that we are considering useful for a given purpose. Similarity between objects may be measured by mathematical functions called resemblance functions or similarity functions (Orlóci 1978) and it becomes a special case of data transformation: the description of a set of objects based on a set of characters is transformed in a description given by the set of similarities that each object has with the other objects in the set, i.e. the data matrix $m \times n$ is transformed in a symmetric similarity matrix $n \times n$ or $m \times m$. Similarity was introduced in sciences by Jaccard (1901, 1912) for comparing floras of different areas. It is a concept that is used in all branches of taxonomy and ecology and that had and has a prominent importance in phytosociology as defined by Bran-Blanquet (1964) (cf. Mueller-Dombois & Ellenberg 1974; Mucina & van der Maarel 1989; Podani 2000; Wildi 2017) thanks also to the contributions of Pignatti (e.g. Pignatti & Mengarda 1962; Maarel & al. 1976; Pignatti 1980). We do not enter in discussions about the similarity functions, we consider

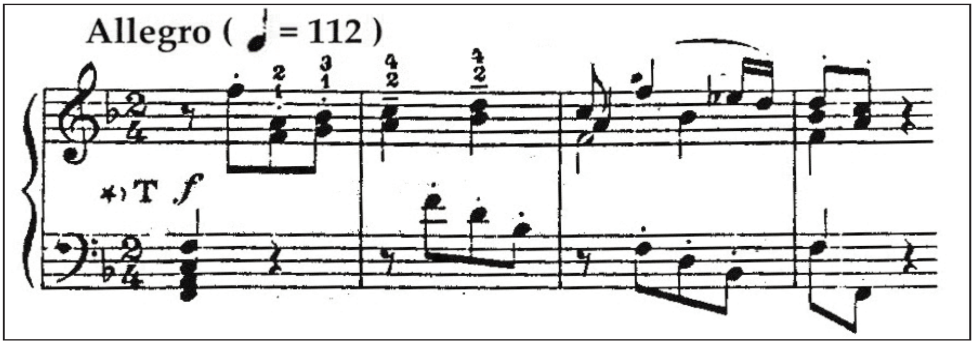


Fig. 1. A piece of Concerto Italiano of J. S. Bach from the book of Pignatti (1982).

that their choice is a matter of the circumstances and the knowledge of the researcher (Dale 1988a), but we want to stress that the similarity matrices are the basic tools for all the multivariate methods linear or non-linear, from multiple regression analysis to non-metric multidimensional scaling, to cluster analysis. In the first case the regression coefficients β are obtained by the cross product of the independent variables in a matrix $\mathbf{X}$, that gives rise to the similarity matrix $\mathbf{X'X}$ that is used in the equation $\mathbf{B}=(\mathbf{X'X})^{-1}\ \mathbf{X'Y}$ where $\mathbf{B}$ is the matrix of the regression coefficients β and $\mathbf{Y}$ is the vector of the dependent variable. In the second case the similarity matrix, can be obtained from different resemblance functions, and it is used, in an iterative process, to obtain a minimum number of axes that are able to fit, in terms of distance, as much as possible the similarity (dissimilarity) between the objects. In the third case, the similarity matrices are the basis to obtain clusters by different methods of cluster analysis (Anderberg 1973; Orloci 1978; Legendre & Legendre 2012). We want also to remember that all the methods known as factor analysis, principal component analysis, principal co-ordinate analysis, canonical correlation analysis, canonical correspondence analysis, discriminant analysis are based on similarity matrices of the cross - product type. The differences between the methods rely on the data transformations, and the algorithms by which the similarity matrices are treated (see Mardia & al. 1979 and Jongman & al. 1995, for a good mathematical presentation of the methods in terms of matrix algebra).

In vegetation science similarity can be calculated between single sample units (pieces of lands or volumes of water) described by species or other characters (descriptors) or between sets of sampling units (as done by Feoli & Scoppola 1980 to quantify the dynamical pattern of the plant communities defined by Pignatti 1959, for the dunes of Venice's Lagoon) or between the descriptors in terms of their behaviour in the sample units (presence-absence or quantities), that generally is called correlation or association, or on the basis of characters that can describe the descriptors (structural, chemical, evolutive etc., cf. Feoli 2018 and references therein). In Phytosociology (Braun-Blanquet 1964) such sampling units are called relevés. They can be done on the basis of different criteria (Mueller-Dombois & Ellenberg 1974) and are the basis for the discipline called syntaxonomy that is dedicated to the hierarchical classification of plant communities (Pignatti 1980, 1990). Besides to be the basis for numerical classifications and ordinations, similarity matrices can be very useful to correlate two or more sets of characters that describe the same set of objects, as was proposed by Mantel (1967), i.e. by calculating the correlation between similarity matrices, or as was proposed by Zerihun Woldu & al. (1989), by the correlation between fuzzy sets obtained by similarity matrices, or, as was proposed by Feoli & Ganis (1986), by the autocorrelation of single or composite variables in spaces defined by the similarity matrices. These ways of finding correlations between sets (representing classes of objects and or measures on single of groups of variables), overcomes all the problems related to different scales of measuring variables that we may find in traditional "parametric" methods treated in Mardia & al. (1979) and Jongman & al. (1995), because we can use different types of similarity functions including those suitable for mixed data such Goodall's (Goodall 1964, 1966), Burnaby's (1970) and Gower's (1970, 1971) ones (see Carranza & al. 1998).

Classification and fuzzy sets

The higher is the similarity between two objects the higher is the possibility that two objects would belong to the same class of objects (Zadeh 1978). The similarity matrices of objects can be interpreted mathematically in terms of fuzzy sets theory: the similarity between two objects is, indeed, the degree of belonging of one object to the set defined by the other object (Zhao 1986; Roberts 1986; Feoli & Zuccarello 1986; Feoli & Orlóci 2011). The definition of sets (classes) is the result of a classification or of an identification process and the *k* sets (classes), that are established, represent always a classification. The degrees of belonging of objects to each of the *k* classes may be calculated by averaging their similarity in the classes or by keeping the minimum or the maximum values of their similarity within the classes, according to the three main strategies of cluster analysis, i.e. average, complete and single linkage (Anderberg 1973; Orlóci 1978; Podani 2000). In this way one object can belong to more than one set with the same or different degrees of belonging. Once we have decided the number of *k* classes on the basis of a given set of characters, we can describe the classes on the basis of other characters (external characters) suitable to infer on the causes that could have been responsible for that classification, e.g. the importance of the environmental factors in determining the *k* classes, or in general to find correlations between environmental factors and other characters that can be used to describe the classes (Feoli & Ganis 2019). The utility of defining fuzzy sets from similarity matrices after classification processes is the fact that the fuzzy sets so obtained, may be used as new variables that have the property to be correlated with the eigenvectors of the similarity matrix from which they are calculated (Feoli & Zuccarello 2013) and that allow to correlate directly classes of objects with variables measured on the objects (in the case of the example of this paper the vegetation types with the environmental indicator values describing the vegetation types).

Diversity

The higher is the number of classes (*k*) of elements we can define within one system (living or not), by a classification or an identification process, the higher is its diversity. Feoli (2018) discussed the concept of fuzzy diversity and proposed to measure it by the formula:

$$D= k\delta \qquad 1)$$

where k is the number of classes and δ a parameter ranging between 0 and 1, representing the dissimilarity within-between (or only between the classes, depending if we consider or not that the elements of one class are completely alike). The parameter δ should not to be seen as a compulsory parameter, it should be used only if we think that the similarity within-between the classes is important to be taken into consideration when measuring the diversity of a system: "the higher the similarity between its components the lower is its diversity". If this assumption is not considered relevant, δ is placed equal to 1 and the diversity is just the richness of the system. Another alternative formula should be $D = k^{\delta}$, however, this is less sensitive to the number of classes. Feoli (2018) distinguished between crisp and fuzzy diversity according to the way by which δ is calculated. If we consider the complete homogeneity within the class (i.e. the similarity between the elements of the class is 1 and dissimilarity is 0) and we consider the similarity between the classes equal to 0, then the parameter δ is considered crisp and it will depend only by the proportion of the classes. The diversity in this case is a crisp diversity and is the one that is traditionally calculated with the Gini-Simpson index or with the Shannon's index. To clarify the matter: if we consider a vector, in which each component represents a quantity of one class of elements, the vector could be interpreted as a crisp similarity matrix with disjoint submatrices with elements $s_{ij} = 1$, representing the perfect similarity between the units that has been used to quantify the classes (components of the vector). For example, if we have a vector with three classes A, B, C with respective quantities (e.g. biomass in kg.): A=4.3, B=8.65 and C=10.68, the vector is representing a disjoint symmetric matrix of (430+865+1068) x (430+865+1068) cells (with values as decagrams) with 3 submatrices **A**, **B**, **C** full of 1s, with respectively 430 x 430, 865 x 865 and 1068 x 1068 cells. It follows that the most used diversity indices e.g. the one of Gini-Simpson and the one of Shannon, are respectively the average similarity-dissimilarity and the entropy of the eigenvalues of the crisp symmetric matrix and their evenness could be used as a crisp δ (cf. Feoli 2018).

If we are interested in considering the similarity between the classes, i.e. the fuzzy diversity, we should consider as a fuzzy δ, the evenness of the well-known index of Rao (cf. Rao 2010), when we are going to apply the Gini-Simpson index, or the evenness of the formula suggested by Ricotta & Szeidl (2006), if we are going to apply the Shannon's entropy. These indices are keeping into consideration both the proportions and the similarity between the classes represented by the components of the diversity vectors. Feoli (2018) stresses the fact that dissimilarity cannot be considered synonymous of "beta diversity". The difference between "alpha" and "beta" diversity are consisting only by the fact that alpha diversity is an individual-based diversity, while beta diversity is a community- based diversity, i.e. in the first case the objects grouped into classes are individuals of plants, animals, microbes, etc., in the second case, they are communities of plant, animals and microbes. Therefore, according to Feoli (2018) alpha diversity is the diversity of species, or other sets of characters, grouping the individuals living in a given ecosystem (these groups or classes can be defined at any hierarchical level we consider adequate for the research we are involved to, e.g. Pignatti 1960, 1980), while beta diversity is the diversity of the ecosystems (adequately defined) that are occupying a given area. In analogy with classifications and ordinations based on rows (R) and those based on columns (Q) of a matrix, when the rows represent classes of individuals defined by a certain set of characters and the columns classes of objects (with one or more elements) described by the classes in rows, we can distinguish two type of diversity of a matrix, the diversity based on rows and the diversity based on columns, i.e. R and Q diversity. Both may be fuzzy or crisp. While Q diversity could correspond to the Whittaker's beta diversity, R could include alpha diversity, when we consider the richness of the single column vector or the average richness of all the column vectors of a matrix, and the gamma diversity of Whittaker when we consider the richness of the complete matrix i.e. the number of its rows.

Feoli (2018) suggests to use the diversity measures of formula 1) to calculate the redundancy of the system by the following formula:

$$\rho = \frac{k - k\delta}{k} = 1 - \delta \qquad 2)$$

that is the complement to one of δ i.e. a measure of a relative "global" similarity within a system.

Eigenvalues

The definition of classes is the aim of any classification process. To represent a useful classification the classes should be significantly separated in order to be in line with the Occam's razor principle "Entities are not to be multiplied beyond necessity" also called principle of ontological economy, principle of parsimony, or principle of simplicity Goodall (1986). There are many methods to measure the separation between the classes that are used in ecological context (Dale 1988b; Pillar 1999a, b; Pillar & Orlóci 1996; Tichy & al. 2010). We do not want to compare all these methods here, we want just to show the use of the evenness of the eigenvalues of similarity matrices ($E(\lambda)$) to measure the separation between classes, as already done by Feoli & Ganis (2019) and to calculate fuzzy diversity of vegetation systems at different hierarchical levels. The reasoning on which the proposal is based, that underlies also nodal ordination (Noy-Meier 1971, 1973; Orlóci 1978), is the following. If a similarity matrix is filled by 1s, it corresponds to only one set of objects, while if it has 1s and 0s, it represents different sets of objects. When there is only one homogeneous set, the matrix has only one eigenvalue that is equal to the number of objects or classes k, i.e. the square root of k x k (Wilkinson 1965). It is enough that 1 element is different from all the others that the number of eigenvalues of the similarity matrix becomes 2. If all the k elements are completely different from the others, then the similarity matrix is the identity matrix (with 1s only in the diagonal and 0s outside the diagonal). In this case the number of eigenvalues is k and all are equal to 1. If instead of single objects we compare classes of objects, the diagonal elements of a similarity matrix may represent (if we think it useful, otherwise we consider them equal to 1) the within classes similarity, while the elements outside the diagonal represent the between classes similarity. As the elements outside the diagonal approaches to 1, the lower becomes the entropy of the eigenvalues of the similarity matrix, because the number of eigenvalues decreases and the first one becomes more and more big. It follows that the evenness of the Shannon's entropy (Shannon & Weaver 1949) of the positive (p) eigenvalues (λ) of the similarity matrices (**S**)

$$E(\lambda) = \frac{-\sum_i \frac{\lambda_i}{\sum_i \lambda_i} \ln \frac{\lambda_i}{\sum_i \lambda_i}}{\ln k} \qquad i= 1,..., p \qquad 3)$$

is a relative measure of the global dissimilarity between objects compared in the similarity matrices. These matrices can be defined at different hierarchical levels by grouping the classes according to some hierarchical criterion. The values of $E(\lambda)$ calculated at different hierarchical levels can be used to find the level for which the classes are significantly separated on the basis of the set of characters used to calculate the given similarity matrix **S**. The k classes corresponding to the highest separation i.e. to the highest $E(\lambda)$, should constitute the optimal classification. The statistical significance of the separation between the classes expressed by $E(\lambda)$ is tested by permutation techniques (cf. Biondini & al. 1991; Pesarin 2001; Pillar 1996; Pillar & Orlóci 1996). Once the classes are defined based on the similarity matrices, they can be used to obtain fuzzy sets for ordinations (Feoli & Zuccarello 1986) and for measuring correlations between them and other variables both environmental or biological, as shown in the example. The $E(\lambda)$ of a similarity matrix obtained by considering its partition in *k* groups (or classes) can be used also as one alternative of the non-parametric analysis of variance between-within the *k* groups (see the example). The definition of the classes can be based on different criteria; all those underlying the analysis of variance (Manova & Permanova e.g. Anderson 2017) can be used. Correspondence between the eigenvalues and eigenvectors with clusters have been shown by simple examples by Feoli & Zuccarello (2013).

It is clear that the higher is the evenness $E(\lambda)$ of a similarity matrix the higher is the difference between the objects or between the classes of objects that are compared in the matrix, therefore the evenness $E(\lambda)$ can be properly used as δ in formula 1) in order to correct the value of richness *k* (number of classes), by a measure of the overall dissimilarity within the system based on the characters we want to take into consideration (Feoli 2018).

Two types of beech forests as a case study

Data

The loop "similarity, classification and diversity" is applied in a simple case study aimed to put in evidence the structural and chorological differences between two vegetation types of Apennines that represent the two opposite states of the beech forests in the Mediterranean area along a North -South gradient. The first association is *Aquifolio - Fagetum* (AF) that includes three sub-associations, two of them have two variants; the second association is *Trochiscantho - Fagetum* (TF) that also includes three sub-associations of which two are subdivided in two variants. In total there are 10 variants: 5 in the first association and 5 in the second one. The variants have been defined by clustering methods using the species as characters by Feoli & Lagonegro (1982). AF is distributed in the Southern part of the Apennines and on the mountain of Sicily where the rainfall has the maximal peaks in winter and autumn. TF is distributed on the North part of the Apennines where the rainfall has the maximal peaks of rainfall in spring and autumn (Feoli & Lagonegro 1982). The description of the 10 variants is given in two data sets represented by two very simple data matrices. The data have been extracted from the paper of Feoli & Ganis (1985). The first (Table 1) consists in the description of the 10 variants by 8 environmental factors obtained by averaging each of the 8 ecological indicator values of Landolt (1977) of the species in each type (matrix **X**). The second data set (Table 2) shows the description of the same vegetation types by the sum of frequency classes (from 1 to 5) of species in cells of a three-way contingency table given by the chorological type (i.e. the type of geographical distribution), growth forms and life forms, i.e. structural characters (matrix **Y**).

Table 1. Description of 10 syntaxonomical variants (vegetation types) of beech forests of Central Italy by the average values of indicator values corresponding to environmental factors according to Landolt (1977). F = Humidity, R = Reaction, N = Nutrients, H = Humus, D = Dispersion, L = Light, T = Temperature, C = Continentality.

	Codes	F	R	N	H	D	L	T	C
Aquifolio-Fagetum cyclametosum	AF1	2.8	3.3	2.8	3.5	3.6	2.5	3.8	2.4
Aquifolio -Fagetum carpinetosum var. *Milium*	AF2.1	2.9	3.3	3.1	3.6	3.8	2.1	3.6	2.3
Aquifolio-Fagetum carpinetosum var. *Lamium*	AF2.2	3.1	3.2	2.9	3.6	3.8	2.2	3.7	2.3
Aquifolio -Fagetum brachypodietosum var. *Digitalis*	AF3.1	2.8	3.2	2.9	3.5	3.6	2.4	3.6	2.5
Aquifolio-Fagetum brachypodietosum var. *Quercus ilex*	AF3.2	2.8	3.1	2.9	3.5	3.6	2.4	3.6	2.6
Trochiscantho-Fagetum daphnetosum mezerei	TF1	2.9	3.2	3.0	3.6	3.8	2.2	3.0	2.6
Trochiscantho-Fagetum ranunculetosum lanuginosi	TF2.1	3.0	3.0	3.1	3.8	3.9	2.1	3.0	2.6
Trochiscantho-Fagetum ranunculetosum var. *Acer pseudoplatanus*	TF2.2	3.1	3.2	3.2	3.5	3.9	2.1	3.0	2.5
Trochiscantho-Fagetum luzuletosum var. *Sesleria autumnalis*	TF3.1	2.8	3.3	2.8	3.6	3.8	2.3	3.2	2.6
Trochiscantho-Fagetum luzuletosum niveae	TF3.2	2.9	3.0	3.0	3.7	3.9	2.0	3.1	2.5

Table 2. Description of 10 variants of beech forests of Central Italy (listed in Table 1) by the combination of chorological types and the structural characters (Growth forms and Life forms). The numbers are the sum of Raunkier’s frequency classes of the species in the phytosociological tables given by Feoli & Lagonegro (1982) as obtained by Feoli & Ganis (1985).

	Codes	AF1	AF2.1	AF2.2	AF3.1	AF3.2	TF1	TF2.1	TF2.2	TF3.1	TF3.2
Circumboreal-Caespitose-Hemicryptophyte	C-C-H	0	0	0	0	0	3	3	0	5	7
Circumboreal-Scapose-Hemicriptophyte	C-Sc-H	0	0	0	0	3	2	1	6	3	2
Eurasiatic-Scapose-Hemicriptophyte	Eu-Sc-H	3	0	0	0	4	1	4	2	0	3
Euroasiatic-Rhizomatose-Geophyte	Eu-R-G	0	9	4	4	0	15	12	19	3	10
European-Caespitose-Hemicriptophyte	E-C-H	0	2	2	3	4	0	0	0	0	0
European-Scapose-Hemicriptophyte	E-Sc-H	11	6	4	9	4	12	18	18	10	11
European-Scapose-Phanerophyte	E-Sc-P	10	7	7	5	6	7	5	7	5	5
European-Suffruticose-Chamaephyte	E-Su-C	4	2	4	4	0	0	0	0	0	0
Mediterranean Atlantic-Caespitose-Phanerophyte	MA-C-P	8	7	9	7	3	0	0	0	3	2
Mediterranean Montane-Scapose-Hemicriptophyte	MM-Sc-H	9	5	8	2	6	10	2	9	8	6
South East European-Rhizomatose-Geophyte	SEE-R-G	3	3	9	2	0	0	0	0	0	0

Methods

With this case study we show how the use of similarity, classification, $E(\lambda)$ and fuzzy sets can answer easily the following questions:

a) Are the two plant associations, already defined by species by Feoli & Lagonegro (1982), significantly separated in the space defined by environmental factors and chorological-structural characters?

b) What is the effect of the environmental factors on the chorological and structural pattern of the two associations and of the single phytosociological variant?

c) What is the Q-diversity of the considered vegetation system at different hierarchical levels, on the basis of the similarity based on environmental factors and on that based on combination of chorological and structural characters?

To answer to the three questions, we have produced two similarity matrices **S**(X) and **S**(Y) for the 10 variants, one for matrix **X** and one for matrix **Y**. We have used the same similarity function for the two matrices. This is the complement to 1 of Euclidean distance after having transformed all the d_{ij} values of the distance matrix to range between 0 and 1 according to the following formula:

$$S_{ij} = 1 - \frac{d_{ij} - d_{min}}{d_{max} - d_{min}} \qquad 4)$$

where d_{min} and d_{max} are respectively the minimum and maximal Euclidean distance in the matrix. This transformation is suggested to relativize the distance to the maximal distance we can have between two objects in a dissimilarity matrix. As we have already said, there are many similarity functions that can be applied; we refer to Pavoine & al. (2005) for a comparison between several of them in measuring fuzzy diversity by the Rao's entropy (Rao 2010).

To answer question a) we have calculated $E(\lambda)$ with the similarity matrices **S**(X) and **S**(Y) by grouping the 10 vegetation types (variants) according to the 2 associations. We have measured the separation between the two associations also in terms of the single environmental factors and in terms of the single chorological-structural characters, in this way $E(\lambda)$ is used as an alternative of the test of Kruskal-Wallis (i.e. a univariate non-parametric analysis of variance). To do this we have obtained 8 similarity matrices by comparing the 10 vegetation types on the basis of each of the 8 environmental factors in **X** and 11 similarity matrices on the basis of the 11 characters in **Y**. The characters may be combined in several ways to test the capacity of their combinations to separate the classes, but we do not enter in such an exercise that does not add new meanings to the aims of the paper as we have said in Feoli & Ganis (2019).

To answer to question b) i.e. to define what are the environmental factors defining the chorological and structural differences between the two associations, we have calculated the corresponding fuzzy partitions. These have been obtained by averaging the similarity values of the variants within and between the two associations, based on matrix **X**, and based of matrix **Y** (Feoli & Zuccarello 1986) and then by squaring the normalized average similarity values. In this way the degrees of belonging become complementary i.e. their sum is equal to 1. With the 4 fuzzy sets, 2 for matrix **X** and 2 for matrix **Y** each corresponding to one association, we have calculated the following correlations:

- between the 4 fuzzy sets in order to have a general correlation between the environmental factors and the chorological and structural variables (in analogy of what is done in Canonical Correlation Analysis (Legendre & Legendre 2012), but conditioned by the classification of the vegetation types in two associations);
- between the 4 fuzzy sets and the 8 environmental factors in order to put in evidence the importance of the single environmental factors in defining the two fuzzy sets, i.e. the two plant associations;
- between the 4 fuzzy sets and the 11 chorological-structural variables in order to put in evidence what are, among these, the characters that are characterizing the two associations.

We have also analysed the correlation between the single variant (the vegetation type at lower hierarchical level) and the variables in **X** and in **Y** by the Pearson's and the

Spearman’s rank correlation coefficients, considering the single vectors of similarity in matrices **S**(X) and **S**(Y) and the variables in **X** and in **Y**. The single vectors of the two similarity matrices are fuzzy sets representing the degrees of belonging of each variant to the set represented by another variant (Zhao 1986; Feoli & Zuccarello 1986).

It is clear that all the correlations based on fuzzy sets are contextual correlations, i.e. they depend on the set of objects we are considering. It should be also clear that the correlation is just a measure of similarity ranging between -1 and 1 that can be easily transformed in one ranging between 0 and 1 with the formula $(r_{ij}+1)/2$, or $(r_{ij} - r_{min})/(r_{max} - r_{min})$, in analogy with formula 4).

To answer question c) we have applied the formula 1) and 2) by considering the $E(\lambda)$ of the similarity matrix 10 x 10, corresponding to the 10 vegetation variants, to the matrix 6 x 6 corresponding to the six sub-associations, to the matrix 2 x 2 corresponding to the two associations. Formula 2) has been used to measure redundancy of the two sets of characters in **X** and **Y**. We have considered the data set as representing a vegetation system that can be subdivided significantly in different vegetation types at three hierarchical levels. The redundancy (formula 2) can show how the system is going to be sharp as we pass from one level to the upper one.

Results

The results related to question a) confirm that the separation between the two associations in the space defined by environmental factors and in the space defined by the combination of chorological and structural characters is significant (Table 3).

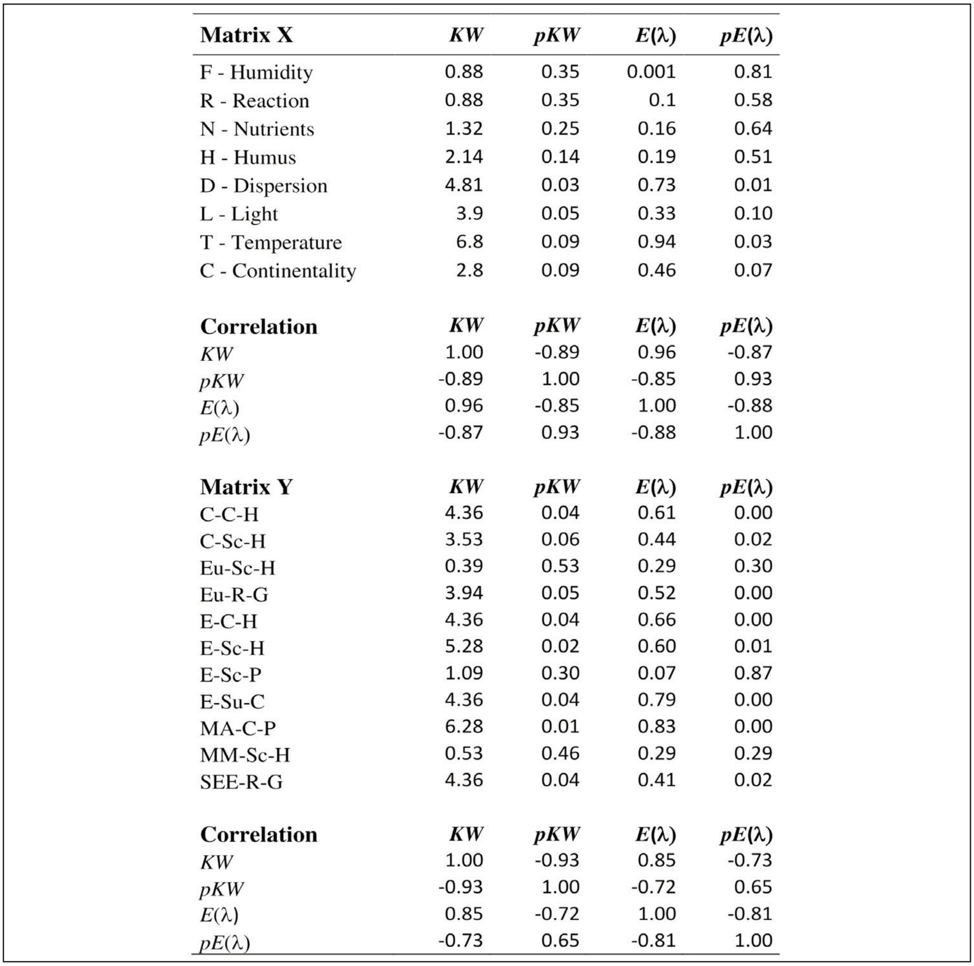
As we have stressed in Feoli & Ganis (2019), the $E(\lambda)$ is higher for the two associations when described by the environmental factors ($E(\lambda) = 0.824$), than when described by the combination chorological-structural characters ($E(\lambda) = 0.785$). However, the separation is maximal when we consider only the temperature ($E(\lambda) = 0.94$). If we consider separately the chorological characters and the structural characters, the separation is higher for the structural characters ($E(\lambda) = 0.797$ versus 0.72) (Table 3).

Table 3. Evenness ($E(\lambda)$) and its probability ($pE(\lambda)$) of the similarity matrix 2 x 2 obtained by grouping the 10 variants in two associations by averaging the similarity values of the matrix 10 x 10. X indicates the evenness ($E(\lambda)$) for the similarity matrix based on matrix **X**, Y indicates the evenness based on matrix **Y**, c indicates the evenness of the similarity matrix obtained by only the chorological data, s indicates the matrix obtained only by the structural data.

	$E(\lambda)$	$pE(\lambda)$
$E(\lambda)\ X$	0.824	0.0001
$E(\lambda)\ Y$	0.785	0.0044
$E(\lambda)\ Yc$	0.724	0.007
$E(\lambda)\ Ys$	0.797	0.00021

This means that the two associations defined on the basis of floristic data, occupy two different community niches well separated from the environmental point of view, with the temperature that is playing the most important role in differentiating the structure and the chorology of the two associations. The fact that they are more separated from the environmental point of view rather than from structural one would suggest that there is a certain convergence towards a common functionality of the beech forest. The results of the separation between the associations based on single environmental factors and on the single chorological and structural characters is given in Table 4.

Table 4. Results of application of Kruskal Wallis test (*KW*) to the original data and $E(\lambda)$ to matrices 2 x 2 obtained by the similarity matrices 10 x 10 corresponding respectively to matrices **X** and **Y,** by considering the classification of variants in two associations, *pKW* is the probability of test *KW*, $pE(\lambda)$ is the probability of the evenness test. The correlations between the two tests and their probability are also given. Symbols as in Tables 1 and 2.



Matrix X	***KW***	***pKW***	$E(\lambda)$	$pE(\lambda)$
F - Humidity	0.88	0.35	0.001	0.81
R - Reaction	0.88	0.35	0.1	0.58
N - Nutrients	1.32	0.25	0.16	0.64
H - Humus	2.14	0.14	0.19	0.51
D - Dispersion	4.81	0.03	0.73	0.01
L - Light	3.9	0.05	0.33	0.10
T - Temperature	6.8	0.09	0.94	0.03
C - Continentality	2.8	0.09	0.46	0.07

Correlation	***KW***	***pKW***	$E(\lambda)$	$pE(\lambda)$
KW	1.00	-0.89	0.96	-0.87
pKW	-0.89	1.00	-0.85	0.93
$E(\lambda)$	0.96	-0.85	1.00	-0.88
$pE(\lambda)$	-0.87	0.93	-0.88	1.00

Matrix Y	***KW***	***pKW***	$E(\lambda)$	$pE(\lambda)$
C-C-H	4.36	0.04	0.61	0.00
C-Sc-H	3.53	0.06	0.44	0.02
Eu-Sc-H	0.39	0.53	0.29	0.30
Eu-R-G	3.94	0.05	0.52	0.00
E-C-H	4.36	0.04	0.66	0.00
E-Sc-H	5.28	0.02	0.60	0.01
E-Sc-P	1.09	0.30	0.07	0.87
E-Su-C	4.36	0.04	0.79	0.00
MA-C-P	6.28	0.01	0.83	0.00
MM-Sc-H	0.53	0.46	0.29	0.29
SEE-R-G	4.36	0.04	0.41	0.02

Correlation	***KW***	***pKW***	$E(\lambda)$	$pE(\lambda)$
KW	1.00	-0.93	0.85	-0.73
pKW	-0.93	1.00	-0.72	0.65
$E(\lambda)$	0.85	-0.72	1.00	-0.81
$pE(\lambda)$	-0.73	0.65	-0.81	1.00

The values of the Kruskal Wallis (*KW*) test (non-parametric analysis of variance) are also reported for a comparison. The correlation between $E(\lambda)$ and the *KW* test is very high and significant (Table 4). The two tests fully agree in showing the characters that significantly separate the two associations. The temperature and the dispersion (i.e. the dimension of the soil particles), are the only two environmental factors that are strongly significant, while almost all the chorological-structural characters are significant except 3 of them (see Table 4). This means that these two environmental factors, temperature and dispersion, are influencing very strongly the structure of vegetation.

Concerning the question b) we have to say that the application of Pearson’s and Spearman’s correlation coefficients has given the same results, so we present only those of the first coefficient. Table 5 shows the fuzzy partitions corresponding to the two association.

Table 6 shows the correlations between the fuzzy sets corresponding to the two associations (*Aquifolio-Fagetum* and *Trochiscantho-Fagetum*, the environmental factors in Table 1 (matrix **X**) and the combination of chorological and structural variables in Table 2 (matrix **Y**). What is relevant from Table 4 and Table 6 is the fact that the most significant environmental factors influencing the definition of the two associations are also significantly correlated with the structure of vegetation (i.e temperature is correlated with European - Caespitose - Hemicriptophyte, European - Suffutricose - Chamaephyte, Mediterranean Atlantic - Caespitose - Phanerophyte and South East European - Rhizomatose - Geophyte).

From Table 6 we can see in detail how the two associations are different in structure, since the corresponding fuzzy partition allows to see what are the environmental and chorological-structural characters that are contrasting (i.e. correlated in opposite way) among the two associations in a significant way, for example N, L, D, T and Euroasiatic - Rhizomatose - Geophyte, European - Caespitose - Hemicriptophyte, European - Scapose - Hemicriptophyte, European -Suffruticose - Cahmaephyte and Mediterranean Atlantic - Caespitose - Phanerophyte.

Table 5. Fuzzy partitions of the fuzzy sets originated by similarity matrices based on matrix **X** (G1X(FP)and G2X(FP)) and matrix **Y** (G1Y(FP)and G2Y(FP)).

	codes	**G1X(FP)**	**G2X(FP)**	**G1Y(FP)**	**G2Y(FP)**
Aquifolio-Fagetum cyclametosum	AF1	0.95	0.05	0.64	0.36
Aquifolio -Fagetum carpinetosum var. Milium	AF2.1	0.68	0.32	0.67	0.33
Aquifolio-Fagetum carpinetosum var. Lamium	AF2.2	0.78	0.22	0.67	0.33
Aquifolio -Fagetum brachypodietosum var. Digitalis	AF3.1	0.81	0.19	0.67	0.33
Aquifolio-Fagetum brachypodietosum var. Quercus ilex	AF3.2	0.78	0.22	0.46	0.54
Trochiscantho-Fagetum daphnetosum mezerei	TF1	0.17	0.83	0.28	0.72
Trochiscantho-Fagetum ranunculetosum lanuginosi	TF2.1	0.09	0.91	0.28	0.72
Trochiscantho-Fagetum ranunculetosum var. Acer pseudoplatanus	TF2.2	0.13	0.87	0.35	0.65
Trochiscantho-Fagetum luzuletosum var. Sesleria autumnalis	TF3.1	0.42	0.58	0.37	0.63
Trochiscantho-Fagetum luzuletosum niveae	TF3.2	0.17	0.83	0.32	0.68

Table 6. Matrix of correlations between the variables of **X** and **Y** and the two fuzzy sets (fuzzy partition) corresponding of the two associations based on similarity calculated with matrix **X** (G1X(FP)and G2X(FP)) and matrix **Y** (G1Y(FP)and G2Y(FP)). Significance of *r*: *p*=0.05, *r*=0.632; *p*=0.02, *r*=0.716; *p*=0.01, *r*=0.765; *p*=0.001, *r*= 0.872. Symbols as in Table 1 and Table 2.

Codes	F	R	N	H	D	L	T	C	G1X(FP)	G2X(FP)	G1Y(FP)	G2Y(FP)
C-C-H	-0.14	-0.43	-0.05	**0.64**	0.53	-0.46	-0.63	0.46	-0.63	**0.63**	-0.65	**0.65**
C-Sc-H	0.25	-0.13	0.39	-0.19	0.37	-0.25	-0.61	0.51	-0.56	0.56	-0.96	**0.96**
Eu-Sc-H	-0.04	-0.67	0.15	0.24	0.01	0.00	-0.20	0.42	-0.23	0.23	-0.64	**0.64**
Eu-R-G	0.62	-0.25	**0.87**	0.32	**0.77**	-0.76	-0.80	0.17	-0.84	**0.84**	-0.45	0.45
E-C-H	-0.25	0.01	-0.20	-0.43	-0.62	0.39	0.63	-0.18	**0.65**	-0.65	0.60	-0.60
E-Sc-H	0.33	-0.28	0.53	0.34	0.53	-0.35	-0.75	0.44	-0.74	**0.74**	0.41	-0.41
E-Sc-P	0.04	0.55	-0.15	-0.47	-0.33	0.38	0.46	-0.51	0.44	-0.44	**0.74**	-0.74
E-Su-C	-0.07	0.45	-0.43	-0.39	-0.55	0.49	**0.82**	-0.71	**0.79**	-0.79	-0.15	0.15
MA-C-P	-0.18	0.52	-0.52	-0.38	-0.56	0.47	**0.93**	-0.79	**0.89**	-0.89	**0.90**	-0.90
MM-Sc-H	0.15	0.45	-0.16	-0.37	0.08	0.08	-0.07	-0.08	0.04	-0.04	0.58	-0.58
SEE-R-G	0.37	0.32	-0.26	-0.13	-0.15	0.12	**0.66**	-0.82	0.59	-0.59	-0.67	**0.67**
G1X(FP) *Aquifolio Fagetum*	-0.42	0.51	-0.64	-0.61	-0.84	0.75	**0.99**	-0.53	**1.00**	-1.00	**0.90**	-0.90
G2X(FP) *Trochiscantho-*	0.42	-0.51	0.64	0.61	0.84	-0.75	-0.99	0.53	-1.00	**1.00**	-0.90	**0.90**
G1Y(FP) *Aquifolio Fagetum*	-0.16	0.54	-0.37	-0.50	-0.61	0.48	**0.93**	-0.78	**0.90**	-0.90	**1.00**	-1.00
G2Y(FP) *Trochiscantho-*	0.16	-0.54	0.37	0.50	0.61	-0.48	-0.93	0.78	-0.90	**0.90**	-1.00	**1.00**

In this way the two associations are quantitatively well characterized in a way that models could be easily produced about the changes in structure of the considered beech forests, due to climatic changes; e.g. we can say that by increasing temperature they will becomes richer in European - Caespitose – Hemicriptophyte, in European - Suffruticose - Chamaephyte and Mediterranean Atlantic - Caespitose - Phanerophyte, while the lost of fertility (N), and the reduction of soil dispersion (D) would reduce the frequency of Euroasiatic - Rhizomatose - Geophyte and European - Scapose - Hemicriptophyte. We can conclude that the temperature would have produced effects that have differentiated the beech forest in the two well distinguishable associations, both in terms of species and in terms of chorological structural characters. A detailed pattern of relationships between the variants, described by their similarity based on matrix **X** and on matrix **Y**, and the environmental factors and chorological - structural characters can be seen respectively in Table 7 and Table 8.

These results obtained with a very simple data set are in agreement with those obtained by a very large data set Bruelheide & al. (2018), i.e. "the trait combinations seem to be predominantly filtered by local-scale factors such as disturbance, fine-scale soil conditions, niche partitioning and biotic interactions", in fact if we consider Table 7 we can find that the correlations between the fuzzy sets of the single variants and the environmental factors are more numerous than those between the fuzzy sets of the associations and the environmental factors. For example, there are variants correlated with humidity (F) and with humus (H), factors that at hierarchical level of associations don't look to be significant. Table 8 shows that the majority of variants are highly correlated with some chorological structural characters, however the variant TF3.1 (*Trochiscantho - Fagetum luzuletosum* var. *Sesleria autumnalis*) does not show any high correlation both with environmental factors and with chorological structural characters, meaning that it has an intermediate structure in between the two associations as it is suggested also by the degrees of belonging of this variant to the sets of the two associations in Table 5.

The results related to question c) are presented in Table 9. The diversity is higher when we consider the matrix **X**, i.e. the environmental characters to calculate the similarity between the vegetation types. The redundancy ρ (formula 2) is decreasing more than twice from the lower hierarchical levels (respectively 2.3 and 2.27) to the highest ones in the case of environmental factors, while only 1.3 and 1.36 times in case of biological characters. This means that the classification at level of associations is the one that is respecting more the Occam’s razor principle when we consider the environmental rather than the chorological-structural characters. The fact that at the highest hierarchical level (two associations) the Q diversity is higher for the **X** data than for **Y** data, respectively D(X) = 1.724 and D(Y) = 1.58, would lead to the conclusion that the adaptation to the environment tends to uniformize the structure of the beech forests. If we consider just the R diversity of the two associations separately in terms of the richness of chorological-structural characters *Aquifolio - Fagetum* results slightly more diverse (k = 10) than *Trochiscantho - Fagetum* (k = 8). This confirms the general trend of biodiversity i.e. in spontaneous ecosystems the diversity is increasing as temperature is increasing provided that all the environmental factors (including temperature) remain within acceptable limits.

Table 7. Correlation between fuzzy sets defined by the single variant according to matrices **X** and **Y** and the environmental factors in matrix **X** (Significance of r: $p=0.05$, $r=0.632$; $p=0.02$, $r=0.716$; $p=0.01$, $r=0.765$; $p=0.001$, $r= 0.872$) (see Tables 1 for symbols).

Vegetation types	F	R	N	H	D	L	T	C
AF1 X	-0.57	0.52	-0.73	-0.65	-0.92	**0.87**	**0.93**	-0.38
AF2.1 X	0.00	0.48	0.02	-0.19	-0.17	-0.04	0.62	-0.79
AF2.2 X	0.11	0.39	-0.35	-0.29	-0.35	0.24	**0.82**	-0.80
AF3.1 X	-0.65	0.38	-0.67	-0.66	-0.95	**0.83**	**0.81**	-0.15
AF3.2 X	-0.67	0.24	-0.65	-0.63	-0.93	**0.82**	**0.74**	0.00
TF1 X	0.18	-0.32	0.42	0.43	**0.64**	-0.54	-0.94	**0.66**
TF2.1 X	0.41	-0.66	0.62	**0.77**	**0.82**	-0.76	-0.93	0.53
TF2.2 X	0.58	-0.33	**0.78**	0.34	**0.82**	-0.75	-0.91	0.37
TF3.1 X	-0.42	0.19	-0.34	0.06	0.08	0.08	-0.40	0.61
TF3.2 X	0.32	-0.66	0.54	**0.72**	**0.82**	-0.83	-0.88	0.42
AF1 Y	-0.59	0.49	-0.81	-0.37	-0.74	**0.74**	**0.84**	-0.41
AF2.1 Y	-0.33	0.40	-0.20	-0.09	-0.30	0.07	0.61	-0.61
AF2.2 Y	-0.11	0.40	-0.56	-0.24	-0.45	0.35	**0.84**	-0.70
AF3.1 Y	-0.59	0.27	-0.53	-0.25	-0.65	0.49	**0.70**	-0.31
AF3.2 Y	-0.69	0.11	-0.66	-0.34	-0.72	0.60	**0.65**	0.01
TF1 Y	0.18	-0.21	0.52	0.36	0.61	-0.57	-0.84	0.44
TF2.1 Y	0.21	-0.59	0.57	**0.75**	0.64	-0.57	-0.82	0.53
TF2.2 Y	0.50	-0.22	**0.78**	0.20	0.69	-0.60	-0.85	0.35
TF3.1 Y	-0.71	0.21	-0.67	0.06	-0.20	0.27	0.09	0.26
TF3.2 Y	-0.19	-0.47	0.15	0.61	0.51	-0.55	-0.62	0.39

Table 8. Correlation between fuzzy sets defined by the single variant according to matrices **X** and **Y** and the chorological and structural characters in matrix **Y** (Significance of *r*: *p*=0.05, *r*=0.632; *p*=0.02, *r*=0.716; *p*=0.01, *r*=0.765; *p*=0.001, *r*= 0.872). Symbols as in Tables 1 and 2.

	C-C-H	**C-Sc-H**	**Eu-Sc-H**	**Eu-R-G**	**E-C-H**	**E-Sc-H**	**E-Sc-P**	**E-Su-C**	**MA-C-P**	**MM-Sc-H**	**SEE-R-G**
AF1X	-0.56	-0.53	-0.14	-0.85	0.52	-0.62	0.49	**0.76**	**0.81**	0.01	0.43
AF2.1X	-0.44	-0.50	-0.57	-0.25	0.50	-0.68	0.18	0.49	**0.70**	-0.14	0.59
AF2.2X	-0.52	-0.58	-0.50	-0.54	0.57	-0.80	0.28	**0.75**	**0.89**	0.00	**0.92**
AF3.1X	-0.50	-0.42	-0.21	-0.81	**0.76**	-0.68	0.13	0.61	**0.67**	-0.21	0.26
AF3.2X	-0.45	-0.29	-0.05	-0.81	**0.81**	-0.70	0.06	0.42	0.52	-0.19	0.14
TF1X	**0.63**	0.52	0.02	0.72	-0.54	0.56	-0.44	-0.78	-0.87	0.18	-0.62
TF2.1X	**0.64**	0.39	0.31	0.75	-0.55	**0.69**	-0.55	-0.75	-0.85	-0.17	-0.54
TF2.2X	0.36	**0.71**	0.09	**0.93**	-0.53	**0.71**	-0.32	-0.71	-0.82	0.16	-0.48
TF3.1X	0.57	0.31	-0.32	-0.05	-0.21	0.02	-0.48	-0.47	-0.36	0.17	-0.46
TF3.2X	**0.79**	0.38	0.23	0.69	-0.52	0.52	-0.58	-0.73	-0.76	-0.10	-0.50
AF1Y	-0.26	-0.63	-0.10	-0.86	0.23	-0.54	0.50	**0.70**	**0.80**	0.06	0.43
AF2.1Y	-0.21	-0.67	-0.57	-0.40	0.50	-0.71	0.01	0.49	**0.69**	-0.30	0.44
AF2.2Y	-0.37	-0.69	-0.49	-0.69	0.53	-0.85	0.23	**0.76**	**0.92**	-0.01	**0.86**
AF3.1Y	-0.23	-0.69	-0.42	-0.68	0.62	-0.62	-0.15	**0.65**	**0.73**	-0.53	0.33
AF3.2Y	-0.16	-0.28	0.04	-0.86	**0.74**	-0.81	-0.08	0.18	0.44	-0.17	0.10
TF1Y	0.53	0.39	0.01	**0.81**	-0.60	0.56	-0.17	-0.65	-0.76	0.27	-0.56
TF2.1Y	0.54	0.21	0.39	**0.64**	-0.56	**0.81**	-0.46	-0.60	-0.75	-0.38	-0.61
TF2.2Y	0.23	**0.68**	0.15	**0.94**	-0.61	**0.83**	-0.10	-0.61	-0.79	0.21	-0.54
TF3.1Y	0.55	-0.12	-0.22	-0.55	-0.01	-0.34	-0.34	-0.12	0.12	-0.02	-0.21
TF3.2Y	**0.93**	0.11	0.14	0.32	-0.45	0.27	-0.54	-0.57	-0.51	-0.13	-0.55

Table 9. Hierarchical Q-diversities based on formula 1) and corresponding values of redundancy ρ. The hierarchical levels are given respectively by the 10 variants, the 6 subassociations and the 2 associations (see text). **X** means that the similarity between the vegetation types has been calculated on the basis of ecological indicators, **Y** that similarity was calculated on the basis of chorological and structural characters.

Hierarchic levels	**10X**	**6X**	**2X**	**10Y**	**6Y**	**2Y**
$E(\lambda)$	0.683	0.686	0.824	0.725	0.714	0.785
Q Diversity	6.83	4.11	1.724	7.25	4.284	1.58
Redundancy (ρ)	0.317	0.314	0.138	0.275	0.286	0.21

Discussion and conclusion

The case study is showing how, with similarity matrices, we can analyse in detail the relationships between different kind of environmental and biological variables and the states of ecological systems as defined by types of communities and how we can reach a level of information surprisingly high by very simple mathematical tools. We used data describing plant communities, but we could have been used data from any other type of communities and ecosystems and even from other systems described by two or more sets of variables. The choice of this very simple data set was done because the results could have been easily discussed and challenged with the results of studies already done with the same data set and with knowledge already consolidated on the type of vegetation system (i.e. beech forests) in the Mediterranean basin. We think that working with similarity matrices and classification methods, leads to obtain very clear results that could be a useful step towards modelling dynamics of vegetation or other ecosystems. As a matter of fact, the usefulness of the concept of similarity becomes more and more tangible in all the disciplines that are dealing with multivariate time series analysis (Norman & Streiner 2003; Liao 2005) and in general with complex systems and artificial intelligence (AI) (e.g. Rissland 2006; Hofmann & al. 2008). The application of the similarity concept leads to important conclusions in different fields of ecology for example about the behaviour of invasive species Divíšek & al. (2018) or on the role of evolution in the relationships between different ecological communities Podani & al. (2018). In this respect if we consider a set A of *k* classes of objects, δ can be calculated by considering different sets of characters, if for example one set represents the mutual phylogenetic position of the *k* classes (i.e. the similarity matrix **S**(1) is a topological distance matrix of the evolutionary tree of the species involved, e.g. Ricotta & al. (2018), and references therein) and the same *k* classes are described also by a set that may represent their similarity **S**(2) in terms of structure or functional characters, then if ρ_1, corresponding to **S**(1) is larger than ρ_2, corresponding to **S**(2), it would mean that there is a structural divergence, while $\rho_1<\rho_2$ would represent a structural convergence of the *k* classes (obviously for the characters we have chosen). The case study presented in this paper, is answering three questions that are very important in supporting similarity theory (Feoli & Orlóci 2011; Wildi 2018). The first question is concerned with the prediction of a classification based on certain characters with respect to other characters. The second is concerned with detailed correlations analysis between characters of different nature and fuzzy sets representing classes at different hierarchical levels (see Roberts 2008, 2009 for use of fuzzy sets incorporated in traditional methods of multivariate analysis). The third question is concerned with the analysis of fuzzy diversity (i.e. diversity of ecosystems, Q diversity) at different hierarchical levels. In all the three questions it emerges the importance of classification based on similarity, and therefore of numerical taxonomy and numerical syntaxonomy, as already stressed several years ago (Mucina & van der Maarel 1989). In this context the evenness of the eigenvalues of similarity matrices, that can be calculated by different similarity functions and characters at different hierarchical levels, playing a fundamental role in class definition according the Occam’s razor principle, can be considered a key mathematical tool in the “Eternal Golden Braid” of similarity, classification and diversity in vegetation science.

References

Anderberg, M. R. 1973: Cluster analysis for applications. – New York.

Anderson, M. J. 2017: Permutational Multivariate Analysis of Variance (PERMANOVA). – https://doi.org/10.1002/9781118445112.stat07841

Biondini, M. E., Mielke, P. W. Jr. & Redente, E. F. 1991: Permutation techniques based on Euclidean analysis spaces: A new and powerful statistic method for ecological research. – Pp. 221-240 in: Feoli, E. & Orlóci, L. (eds) Computer assisted vegetation analysis. – Boston.

Braun-Blanquet, J. 1964: Pflanzensoziologie. Gründzuge der Vegetationskunde. 3th ed. – Wien.

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F. & al. 2018: Global trait–environment relationships of plant communities. – Nat. Ecol. Evol. **2:** 1906-1917.

Burnaby, T. P. 1970: On a method for character weighting a similarity coefficient, employing the concept of information. – J. Int. Ass. Math. Geol. **2:** 25-38.

Carranza, L., Feoli, E. & Ganis, P. 1998: Analysis of vegetation structural diversity by Burnaby's similarity index. – Pl. Ecol. **138:** 77-87.

Dale, M. B. 1988a: Mutational and Non-mutational similarity Measures: A Preliminary Examination. – Coenoses **3:** 121-133.

— 1988b: Knowing when to stop: cluster concept–concept cluster. – Coenoses **1:** 11-31.

Divíšek, J., Chytrý, M., Beckage, B., Gotelli, N. J., Lososová, Z., Pyšek, P., Richardson, D. M. & Molofsky, J. 2018: Similarity of introduced plant species to native ones facilitates naturalization, but differences enhance invasion success. – Nature Communications **4631**. https://doi.org/10.1038/s41467-018-06995-4

Eigen, M. & Winkler, R. 1986: Il gioco. Le leggi naturali governano il caso. – Milano.

Feoli, E. 2018: Classification of plant communities and fuzzy diversity of vegetation systems. – Community Ecology **19(2):** 186-198.

— & Ganis, P. 1985: Comparison of floristic vegetation types by multiway contingency tables. – Abstr. Bot. **9:** 1-15.

— & — 1986: Autocorrelation for measuring Predictivity in Community Ecology: An example with structural and Chorological data from Mixed Forest Types of NE Italy. – Coenoses **1:** 53-56.

— & — 2019: The use of the evenness of eigenvalues of similarity matrices to test for predictivity of ecosystem classifications. – Mathematics **7:** 245. https://doi.org/10.3390/math7030245

— & Lagonegro, M. 1982: Syntaxonomical analysis of beech woods in the Apennines (Italy) using the program package IAHOPA. – Vegetatio **50:** 129-173.

— & Orlóci, L. 2011: Can similarity theory contribute to the development of a general theory of the plant community? – Community Ecology **12:** 135-141.

— & Scoppola, A. 1980: Analisi informazionale degli schemi di dinamica della vegetazione. Un esempio sul popolamento vegetale delle dune del litorale di Venezia – Giorn. Bot. Ital. **114:** 227-236.

— & Zuccarello, V. 1986: Ordination based on classification: yet another solution?! – Abstr. Bot. **10:** 203-219.

— & — 2013: Fuzzy sets and eigenanalysis in community study: classification and ordination are two faces of the same coin. – Community Ecology **14:** 164-171.

Goodall, D. W. 1964: A probabilistic similarity index. – Nature **203:** 1098.

— 1966: A new similarity index based on probability. – Biometrics **22:** 882-907.

— 1986: Classification and ordination: their nature and role in taxonomy and community studies. – Coenoses **1:** 3-9.

Gower, J. C. 1970. A note on Burnaby's character-weighted similarity coefficient. – J. Int. Ass. Math. Geol. **2:** 39–45.

— 1971. A general coefficient of similarity and some of its properties. – Biometrics **27:** 857-871.

Hofmann, T., Schölkopf, B. & Smola, A. J. 2008: Kernel methods in machine learning. – The Annals of Statistics **36(3):** 1171-1220.

Hofstadter, D. R. 1979: Gödel, Escher, Bach an Eternal Golden Braid. – New York. (1984 Italian edition, Milano).

Jaccard, P. 1901: Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. – Bull. Soc. Vaudoise Sci. Nat. **37:** 241-272.

— 1912: The distribution of the flora in the alpine zone. – New Phytologist **11(2):** 37-50.

Jongman, R. H. G., Ter Braak, C. J. F. & Van Tongheren, O. F. R. 1995: Data analysis in community and landscape ecology. – Cambridge.

Landolt, E. 1977: Okologische Zeigerwerte zur Schweizer Flora. – Ber. Geobot. Inst. ETH. **64:** 64-207.

Legendre, P. & Legendre, L. 2012. Numerical ecology; 3rd ed. – Amsterdam.

Liao, T. W. 2005: Clustering of time series data-a survey. – Pattern recognition **38:** 1857-1874.

Maarel van der, E., László Orlóci, L. & Pignatti, S. 1976: Data-processing in phytosociology, retrospect and anticipation. – Vegetatio **32:** 65-72.

Mantel, N. 1967: The detection of disease clustering and a generalized regression approach. – Cancer Res. **27:** 209-220.

Mardia, K. V., Kent, J. T. & Bibby, J. M. 1979: Multivariate analysis. – London.

Margalef, R. 1980: La biosfera entre la termodinamica y el juego. – Barcellona.

Mucina, L. & Maarel van der, E. 1989: Twenty years of numerical syntaxonomy. – Vegetatio **81:** 1-15.

Mueller-Dombois, D. & Ellenberg, H. 1974: Aims and methods of vegetation ecology. – New York.

Norman, G. R. & Streiner, D. L. 2003: Pretty Darned Quick Statistics. – Hamilton, London.

Noy-Meir, I. 1971: Multivariate analysis of the semi-arid vegetation in South-eastern Australia: nodal ordination by component analysis. – Pp. 159-193 in : Nix, N. A. (ed.) Quantifying Ecology. – Proc. Ecol. Soc. Aust. **6**.

— 1973: Data transformation in ecological ordination. I. Some advantages of non-centering. – J. Ecol. **61:** 329-341.

Orlóci, L. 1978: Multivariate analysis in vegetation research. 2nd ed. – The Hague, The Netherlands.

Pavoine, S., Ollier, S. & Pontier, D. 2005: Measuring diversity from dissimilarities with Rao's quadratic entropy: Are any dissimilarities suitable? – Theor. Popul. Biol. **67:** 231-239.

Pesarin, F. 2001: Multivariate Permutation Tests. With Applications in Biostatistics. – Chichester, Toronto.

Pignatti, S. 1959: Ricerche sull'ecologia e sul popolamento delle dune del litorale di Venezia. Il popolamento vegetale. – Boll. Mus. Civ. Venezia **12:** 61-141.

— 1960: Il significato delle specie poliploidi nelle associazioni vegetali. – Ist. Ven. Sci. Lett. Arti. **118:** 75-98.

— 1980: Reflections on the phytosociological approach and the epistemological basis of vegetation science. – Vegetatio **42:** 181-185.

— 1982: Flora d'Italia, **1-3.** – Bologna.

— 1990: Towards a prodrome of plant communities. – J. Veg. Sci. **1:** 425-426.

— & Mengarda, F. 1962: Un nuovo procedimento per l'elaborazione delle tabelle fitosociologiche. – Accad. Naz. Lincei Mem. Classe Sci. Fis. Mat. Nat. **8:** 215-222.

Pillar, V. D. 1996: A randomization-based solution for vegetation classification and homogeneity testing. – Coenoses **11:** 29-36.

— 1999a: How sharp are classifications? – Ecology **80:** 2508-2516.

— 1999b: On the identification of optimal plant functional types. – J. Veg. Sci. **10:** 631-640.

— & Orlóci, L. 1996: On randomization testing in vegetation science: Multifactor comparisons of relevé groups. – J. Veg. Sci. **7:** 585-592.

Podani, J. 2000: Introduction to the exploration of multivariate biological data. – Leiden.

—, Pavoine, S. & Ricotta, C. 2018: A Generalized Framework for Analyzing Taxonomic, Phylogenetic, and Functional Community Structure Based on Presence–Absence Data. – Mathematics **6**, 250. https://doi.org/10.3390/math6110250

Rao, C. R. 2010: Quadratic entropy and analysis of diversity. – Sankhyā: Ind. J. Stat. **72-A(1):** 70-80.

Ricotta, C. & 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.

—, Bacaro, G., Caccianiga, M., Cerabolini, B. E. L. & Pavoine, S. 2018: A new method for quantifying the phylogenetic redundancy of biological communities. – Oecologia **186:** 339-346.

Rissland, E. L. 2006: AI and Similarity. – IEEE Intelligent Systems **21(3):** 39-49.

Roberts, D. W. 1986: Ordination on the basis of fuzzy set theory. – Vegetatio **66:** 123-143.

— 2008: Statistical analysis of multidimensional fuzzy set ordinations. – Ecology **89:** 1246-1260.

— 2009: Comparison of multidimensional fuzzy set ordination with CCA and DB-RDA. – Ecology **90:** 2622-2634.

Shannon, C. E. & Weaver, W. 1949: The mathematical theory of communication. – Urbana.

Tichý, L., Chytrý, M., Hájek, M., Talbot, S. S. & Botta-Dukát, Z. 2010: OptimClass: using species-to-cluster fidelity to determine the optimal partition in classification of ecological communities. – J. Veg. Sci. **21:** 287-299.

Wildi, O. 2017: Data analysis in vegetation ecology; 3rd ed. – Oxfordshire.

— 2018: Evaluating the predictive power of ordination methods in ecological context. – Mathematics **6(12):** 295. https://doi.org/10.3390/math6120295

Wilkinson, J. H. 1965: The algebraic eigenvalue problem. – London.

Zadeh, L. A. 1978: Fuzzy sets as a basis for a theory of possibility. – Fuzzy Sets Syst. **1:** 3-28.

Woldu, Z., Feoli, E. & Nigatu, L. 1989: Partitioning an elevation gradient of vegetation from south-eastern Ethiopia by probabilistic methods. – Vegetatio **81:** 189-198.

Zhao, S. X. 1986: Discussion on fuzzy clustering. – Pp. 612-614 in: 8th Int. Conference on Pattern Recognition. – New York.

Addresses of the authors:
Enrico Feoli & Paola Ganis,
Department of Life Sciences, University of Trieste, via L. Giorgieri 10, 34127 Trieste, Italy.
Corresponding author: Enrico Feoli, Email: feoli@units.it