CLUSTER ANALYSIS AND VEGETATION MESO-XERIC ORDERING IN MORIȘCA VALLEY BASIN (BOTOȘANI COUNTY)

Corneliu TĂNASE1*

Abstract: The aim of this paper is to classify, describe, and interpret the floristic structure of meso-xeric plant communities in the Morișca Valley, by eco-cenotic data numerical analysis, interlinked with environmental gradients. A number of 92 relevés were analyzed by cluster analysis and grouped in 15 clusters. These clusters correspond to meso-xeric associations in Morișca Valley. These associations have been ordered directly by canonical correspondence analysis (CCA).

Keywords: Canonical Correspondence Analysis (CCA), Cluster analysis, ordering vegetation

Introduction

Morișca Valley is located in the Botoșani County in the N-E Plain of Moldavia and has been described in terms of natural context in a previous paper (Tănase and Ștefan, 2010).

The Morișca Valley vegetation is located in the northern region of Moldavia. Most steppe grassland communities, which occur fragmented in habitat area studies, fall within the Taraxaco serotinae - Festucetum valesiacae association (Burduja et al. 1956, Răvăruț et al. 1956, Sârbu, Coldea et Chifu, 1999). Among other associations derived and secondary to grasslands, we indicate Taraxaco serotinae - Botriochloetum ischaemi (Burduja et al. 1956, Sârbu, Coldea et Chifu, 1999).

The vegetation study had as main objectives the capture of vegetation discontinuities and classification of different types of vegetation in a hierarchical system where the basic unit is the plant association. The importance of the use of statistical and mathematical processing ensues from high demands for reliability and generalization of the obtained results in vegetation studies. These can be an important scientific foundation for highlighting ecological processes involved in formation, dynamics and functioning of plant communities.

Materials and methods

The numerical analysis of data

Gross phytosociological table was analyzed as a matrix of values with relative coverage. In this matrix the columns are of relevés (objects) and the lines correspond to species (variables). A second matrix contains data about abiotic environment that develops the phytocoenoses revealed. Lines represent environment variables and columns represent relevés.

1* “Al. I. Cuza” University of Iasi, Faculty of Biology, Carol I Bd., no. 20A, 700506, Iasi, Romania, email: tanase.corneliu@yahoo.com
For these matrices a geometric pattern is associated where objects (relevè) are represented as points in the multidimensional Euclidean space of variables (species). In Phytosociology, most relevant geometric patterns are those that represent relevè distribution in multidimensional space of species (indirect ordering) or of environmental variables (direct ordering) (Gauch, 1982; Cristea et al., 2004).

Data transformation

Because not all data could be used in the form in which it was collected, certain changes were necessary for the numerical analysis of data. Thus, a necessary first step was the transformation of information into a code of abundance-dominance in relative coverage values (percentage) after Braun – Blanquet scale completed of the Tüxen and Ellenberg (Cristea et al., 2004).

In order ecological ordering of vegetation types, for each relevè it was calculated the average weight for environmental value indicators in relation to the following factors: humidity (U), rich in nitrogen (N) and base (R) of the substrate and salinity (S). The relative coverage of species was used as weights, according to the formula:

\[ \sum_{i=1}^{S} AR_i \times v_i \]

\[ \sum_{i=1}^{S} AR_i \]

where:

- \( AR_i \) - Relative coverage (%) of the \( i \) specie;
- \( v_i \) - Ecological indicator value of the \( i \) specie;
- \( S \) - Total number of species found in the relevè

Cluster analysis and vegetation classification

In this study the classification of plant vegetation was made based on the similarity of floristic composition and hierarchical cluster analysis. The degree of floristic similarity was calculated for all possible pairs of relevès using the quantitative Bray – Curtis index. Further floristic dissimilar matrix was analyzed by UPGMA algorithm for grouping successive clusters of relevès increasingly higher (Legendre and Legendre, 1998). The result of the cluster analysis was plotted as a dendrogram. This was evaluated by distortion relation of similarity between the relevès, by cophenetic coefficient. This measure the correlation between initial similarity matrix and dendrogram derived exclusively. If the value of this coefficient is close to 1, the dendrogram better reflects the information contained in the input matrix. Due to the obvious dependence of the two matrices compared, a significance statistical test cannot be applied for the correlation coefficient cophenetetic (Cristea et al., 2004).

The optimum number of groups (clusters) in each dendrogram was assessed based on its node stability, which was expressed as a percentage obtained by 10000 permutations.
The more these values are closer to 100, the higher are the nodes’ stability, and hence the respective cluster’s.

The assignment of groups (clusters) by relevé for the various plant associations described was based on the proposed procedure of Central European School (Braun-Blanquet, 1964), by the presence of the abundance–dominance characteristic and differential species.

**Ordering vegetation**

Ordering data of vegetation is a collective term for multivariate techniques aimed at arranging samples (relevés) according to their specific structure (ter Braak and Tongeren, 1995), a process of representation as true relations between samples of vegetation, species and environmental factors in the reduced dimensional space of several axes (usually 2 to 3 axis). The final product is a graph (usually two-dimensional); the distance between points (species, associations) reflects the similarity between them in terms of their structure or environmental requirements (Gauch, 1982).

Statistically speaking, the ordering is a process of successive regression and calibration and stops when the relative position of species, relevés or associations can not be improved. The axes are considered latent variables or hypothetical environmental variables (ter Braak and Prentince, 1988; ter Braak and Smilauer, 2002; Cristea et al., 2004).

The relative position of a specie, relevé, or association, can be described as a function of environmental factors registered for direct analysis of gradients, or they can be positioned along the imaginary axis and can be interpreted as an ecological gradient, registering the analysis of indirect gradients (Cristea et al., 2004).

In this study we used one of the most common methods of ordering, namely The Canonical Correspondence Analysis (CCA) (ter Braak, 1986).

**Canonical Correspondence Analysis (CCA)**

This was used for direct ordering of analytical relevé along ecological gradients, measured or available. Input data were grouped into two matrices: matrix relevé/species and matrix relevé/environmental variables. The CCA produces optimal linear adjustment of the distribution of specie’s abundances depending on the environmental variables, axes extracted as linear combinations of environmental factors that best explain the variation in species distributions (Cristea et al., 2004). One of the basic CCA the size of matrix relevé/environmental variables, the number of relevé must be much larger than the number of environmental variables (Cristea et al., 2004), otherwise the analysis is reduced to a simple analysis of correspondence (Oksanen, et al., 2006).

Environmental gradients are plotted as vectors, whose length is proportional to the amplitude factor that they produce and can be interpreted in conjunction with species or relevés. Among other things, output CCA provides information about the extent of correlation between environmental variables and canonical axes, while the variance is explained by each axis and by the total variance of the data set. In addition, we conducted a multivariate test based on Pillai’s Trace index to check how the environmental variables
considered had a statistically significant effect on relative cover species. Statistically significant effect of Pillai’s Trace index was calculated based on 1000 permutations.

**Results and discussions**

For the classification of the association characteristic of meso-xeric habitats (grassland and grassy salting vegetation) a total 92 relevès were used. The cophenetic coefficient associated on dendrogram, has a value of 0.980. On the stability of nodes, and the principle of parsimony, for relevès that compose the dendrogram in Fig. 1 were chosen 15 clusters. These clusters circumscribe plant associations of the meso-xeric type that are present in the Morisca Valley.

**Cluster A** - *Puccinellietum limosae* Rapaics ex Soó 1933,
**Cluster B** - *Camphorosmetum annuae* (Rapaics 1916) Soó 1933,
**Cluster C** - *Staticeto - Artemisietum monogynae (santonicum)* Țopa 1939,
**Cluster D** - *Astero tripoli - Juncetum gerardii* Šmarda 1953.
**Cluster E** - *Taraxaco bessarabici - Caricetum distantis* Sanda et Popescu 1978,
**Cluster F** - *Rorippo austriacae - Agropyretum repentis* (Timár 1947) R. Tüxen 1950,
**Cluster G** - *Junco inflexi - Menthetum longifoliae* Lohmeyer 1953,
**Cluster H** - *Ranunculo repenti - Alopecuretum pratensis* Ellmauer et Mučina in Mučina et al. 1993,
**Cluster I** - *Agrostetum stoloniferae* (Ujvárosi 1941) Burduja et al. 1956,
**Cluster J** - *Epilobio - Juncetum effusi* Oberd. 1957,
**Cluster K** - *Scirpetum sylvatici* Ralski 1931,
**Cluster L** - *Potentilletum anserinae* Felföldy 1942,
**Cluster N** - *Taraxaco serotine - Festucetum valesiacae* (Burduja et al. 1956, Răvărăț et al. 1956) Sârbu, Coldea et Chifu, 1999,

**Canonical Correspondence Analysis (CCA)**

Of the total variance associated canonical axes, the first two CCA axes explained a cumulative 75.71% (Table 1) and a 3 axis have a lower ratio (24.23%).

<table>
<thead>
<tr>
<th>Axis</th>
<th>CCA 1</th>
<th>CCA 2</th>
<th>CCA 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eigenvalues</strong></td>
<td>0.92234</td>
<td>0.88056</td>
<td>0.5765</td>
</tr>
<tr>
<td><strong>Axis variance (%)</strong></td>
<td>38.76</td>
<td>37.01</td>
<td>24.23</td>
</tr>
<tr>
<td><strong>Cumulated variance (%)</strong></td>
<td>38.76</td>
<td>75.71</td>
<td>99.94</td>
</tr>
<tr>
<td><strong>Multivariate test</strong></td>
<td>Pillai’s trace = 2.379;</td>
<td>p = 0.001</td>
<td></td>
</tr>
</tbody>
</table>
Clusters number = 15  
Cophenetic coefficient = 0.9925

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Code</th>
<th>Legend</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Pl</td>
<td>Puccinellietum limosae Rapaics ex Soó 1933</td>
</tr>
<tr>
<td>B</td>
<td>Ca</td>
<td>Camphorosmetum annuae (Rapaics 1916) Soó 1933</td>
</tr>
<tr>
<td>C</td>
<td>SA</td>
<td>Staticeto - Artemisietum monogyne (santonicum) Topa 1939</td>
</tr>
<tr>
<td>D</td>
<td>AJ</td>
<td>Aster to turbi - Juncetum gerardii Smarda 1953</td>
</tr>
<tr>
<td>E</td>
<td>TC</td>
<td>Taraxaco bessarabici - Caricetum distinctis Sanda et Popescu 1978</td>
</tr>
<tr>
<td>F</td>
<td>RA</td>
<td>Rorippo austriaca - Agropyretum repentis (Timár 1947) R. Tüxen 1950</td>
</tr>
<tr>
<td>G</td>
<td>JM</td>
<td>Junco inflexi - Menthetum longifoliae Lohmeyer 1953</td>
</tr>
<tr>
<td>H</td>
<td>RrA</td>
<td>Ranunculo repentis - Alopecuretum pratensis Ellmayer et Mučina in Mučina et al. 1993</td>
</tr>
<tr>
<td>I</td>
<td>As</td>
<td>Agrostetum stoloniferae (Ujvárosi 1941) Burduja et al. 1956</td>
</tr>
<tr>
<td>J</td>
<td>EJ</td>
<td>Epilobio - Juncetum effusi Oberd. 1957</td>
</tr>
<tr>
<td>K</td>
<td>Ss</td>
<td>Scirpetum sylvatici Ralski 1931</td>
</tr>
<tr>
<td>L</td>
<td>Pa</td>
<td>Potentillietum anserinae Felldy 1942</td>
</tr>
<tr>
<td>M</td>
<td>TL</td>
<td>Trifolio repentis - Lolietum perennis Krippelova 1967</td>
</tr>
<tr>
<td>N</td>
<td>TF</td>
<td>Taraxaco serotine - Festucetum valesiacae (Burduja et al. 1956, Răvăruţ et al. 1956) Sârbu, Coldea et Chifu, 1999)</td>
</tr>
<tr>
<td>O</td>
<td>TB</td>
<td>Taraxaco serotine - Bothriochloetum ischaemi (Burduja et al. 1956, Răvăruţ et al. 1956) Sârbu, Coldea et Chifu, 1999)</td>
</tr>
</tbody>
</table>

Figure 1. The dendrogram of relevé of grasslands and halophytic habitats on Valley Morișca
Figure 2. Ordering of relevé of grassland and halophytic habitats in the space delimited by the first two CCA axes.

Legend
N_med – average for nitrogen content in soil indicator
U_med – average for humidity indicator
R_med – average for soil reaction indicator
S_med – average for soil salinity indicator
Allocation of relevè from different plant association was realized in the dendrogram in figure 2. Looking at Fig. 2, we see that all ecological gradients follow, are associated with CCA 1 axis. Thus CCA1 axis may be associated positively with the soil reaction and salinity gradients and negative with gradients for moisture and nitrogen content of soil.

From ordering analysis of relevè, species and environmental variables in space determined by the first two CCA axes (Fig. 2) is observed three clearly defined groups of associations.

The first group consists of three associations: Trifolio repenti – Lolietum perennis, Taraxaco serotinae – Festucetum valesiacae, Taraxaco serotinae – Bothriochloetum ischaemi. The Trifolio repenti – Lolietum perennis association is characteristic mesophile grasslands, located on the dendrogram near the intersection of the two CCA axes. Farther away from the point of intersection of the axes are the other two associations characterized by Taraxacum serotinum (Waldst et Kit) Poiret, Festuca valesiaca Schleicher ex Gaudin, and Bothriochloa ischaemum (L.) Keng. They are negatively correlated with humidity and soil nitrogen content and positive with soil reaction. This ordering of relevè and environmental variables in space determined by the first two CCA axes argues that these associations are characteristic of xero – mesophyll grassland.

Group 2 is represented by associations characteristic meso – hydrophilic meadows. This group includes seven associations: Scirpetum sylvatici, Epilobio – Juncetum effusi, Potentilletum anserinae, Rorippo austriacae – Agropyretum repentis, Agrostetum stoloniferae, Junco inflexi – Menthetum longifoliae, Ranunculo repenti – Alopecuretum pratensis. This group of associations is correlates well with CCA 1 axis that is associated with gradients for moisture and nitrogen content of soil. This supports the fact that these associations are characteristic of grassland with high moisture and rich in nutrients.

Group 3 containing associations Puccinellietum limosae, Staticeto – Artemisietum monogynae, Camphorosmetum annuae, Taraxaco bessarabici – Caricetum distantis and Astero tripoli – Juncetum gerardii is positive order along a gradient for soil salinity. This ordering supports the associations of group 3 are characteristic halophytic habitats. In addition to good correlation with the salinity gradient, associations of group 3 is positively correlated to soil reaction gradient.

Conclusions

Following cluster analysis were analyzed 92 relevè that were grouped in 15 clusters corresponding to 15 associations characteristic of grassland and halophytic habitats on the Valley Morisca.

From ordering analysis of relevè, species and environmental variables in space determined by the first two CCA axes, is observed associations grouping in the three areas of multidimensional space, depending on their relation with environmental indices introduced in the analysis. Affinity associations were observed in group 1 to drier soils and poor in nitrogen. In contrast, associations are in group 2 who prefer moist soils rich in nitrogen. Associations of group 3 prefers soils with higher salinity, is positively correlated to soil reaction gradient.
Acknowledgement

This paper was realised with the support of Phd. Nicolae Ştefan at „Al. I. Cuza” University of Iaşi, Faculty of Biology, and Phd. Dan Gafta at „Babeş – Bolyai” University, Department of Taxonomy and Ecology.

REFERENCES


