Hydrobiologia **300/301**: 425–431, 1995.
G. Balvay (ed.), *Space Partition within Aquatic Ecosystems.*

Using multivariate analyses for separating spatial and temporal effects within species–environment relationships

Evelyne Franquet[1], Sylvain Dolédec[2] & Daniel Chessel[2]
[1] *URA CNRS 1451, Equipe DESMID, Ecologie des Systèmes Fluviaux, 1, rue Parmentier, F-13200 Arles, France.*
[2] *URA CNRS 1451, Ecologie des Eaux Douces et des Grands Fleuves, Université Lyon 1, F-69 622 Villeurbanne cedex, France*

Key words: Multivariate methods, co-inertia analysis, between-site, species–environment relationships, aquatic macroinvertebrates

Abstract

We present a multivariate approach for the analysis of contingency tables involved in the study of species–environment relationships. The first table (species × sample) contains the abundance of p species collected in n samples. The second table (environmental variables × sample) contains values for q environmental variables measured in the n samples. The third table contains the indication of 'where' and 'when' samples were taken.

In this paper we demonstrate: (1) how to match an environmental table and a faunistic table using co-inertia analysis; (2) how to take into account a spatial effect using between-class analyses; and (3) how to combine point (1) and (2) to determine the spatial structure shared by fauna and habitat.

We illustrate such an approach by using a set of hydrobiological data concerning 13 *Ephemeroptera* species and ten physical and chemical variables which were collected in the same site at the same dates in a small river of the Prealps.

Introduction

Most studies in community ecology infer the relationships between species and their environment from community composition data and associated habitat measurements (ter Braak, 1986). Space and time information is often associated to the two resulting data tables. This information corresponds to 'where' (site) and 'when' (date) the samples were taken. Demonstrating the relationships between a set of faunistic data and a set of environmental data has several purposes: (1) to explain a spatial typology based on faunistic data using several environmental variables (e.g., Townsend *et al.*, 1984; Wright *et al.*, 1989; Johnson & Wiederholm, 1989; Storey *et al.*, 1990); (2) to estimate the values of environmental variables from species abundance (indicator species, e.g., Rutt *et al.*, 1990); and (3) to demonstrate the agreement between the typologies resulting from faunistic data and from environmental data (Chessel & Mercier, 1993; Mercier *et al.*, 1992; Dolédec & Chessel, 1994; Bornette *et al.*, 1994). Methods such as canonical correspondence analysis (CCA of ter Braak, 1986, 1987) and partial canonical correspondence analysis (ter Braak, 1988) have addressed points (1) and (2). The study of the co-structure between faunistic data and habitat measurements using co-inertia analysis has been shown as an alternative to CCA (Dolédec & Chessel, 1994).

As the co-structure between fauna and environment may vary according to space, we present in this paper an approach that makes it possible to incorporate spatial effect within species–environment relationships.

Material and methods

Data and site description

Data come from a study made on a stream in the Prealps (Pegaz-Maucet, 1980). Six sites (Fig. 1a) were sampled along the course of the river on four occasions (Fig. 1b). Sites 1 to 5 are situated on the Meaudret river; site 6 is situated on the Bourne tributary. Organic effluents cause local pollution at station 2. Ten physical and chemical variables were measured (Fig. 1c). At the same site and at the same date, 13 *Ephemeroptera* taxa (Fig. 1d) were collected.

Data processing

Each table appears as a multidimensional space with , e.g., p dimensions for the environmental table (noted as X in Fig. 2a), and q dimensions for the faunistic table (noted as Y in Fig. 2a). Each environmental variable and each taxa defines, respectively, a vector direction in each multidimensional space. Multivariate analyses such as principal component analysis (PCA) can be processed on each table separately. The analyses result in the finding of one or several axes, so that the projected inertia onto these axes is maximal. As a consequence, the reduction in dimensions of each data set leads to one structure within the environmental data set and another within the faunistic set. The question then arises as to whether there is an aggreement between these two independant structures. Such a question can be answered by using co-inertia analysis (Chessel & Mercier, 1993; Mercier *et al.*, 1992; Dolédec & Chessel, 1994). This method calculates axes maximizing the covariance between the factorial scores of samples. Tucker (1958) gave a solution called interbattery analysis, where each table X and Y are comprized of quantitative measurements. If each table X and Y is composed of categorical variables, then the co-structure between the two data sets results from a canonical analysis on categorical variables (Cazes, 1980). Finally, if X contains variables by category and Y contains the presence or absence of species, then the co-structure between the two data sets results from a correspondence analysis of ecological profiles (Romane, 1972).

A further step, which is addressed in the paper, may be incorporated into the above approaches by considering the spatio-temporal information (the 'where' and 'when') associated with the samples. Data concerning space and time may be considered as categories. Spatial (or temporal) structures of a faunistic or an environmental table may be discovered using between-site analysis (Fig. 2b). Such an operation consists of centering the analysis on the spatial (or temporal) effect by dispersing the centers of classes (Dolédec & Chessel, 1987, 1989, 1991; Yoccoz & Chessel, 1988; Lebreton *et al.*, 1991). As a result, between-site analysis and co-inertia analysis may be combined to study the spatial co-structure between species and their habitats. The principle for matching two between-site analyses consists in discovering combinations of variables in each tables of average values (*i.e.*, mean abundance of species and mean values for environmental variables, at each site) (Fig. 2c). As a consequence, co-inertia axes are an expression of the spatial co-structure between species and environmental variables.

Results

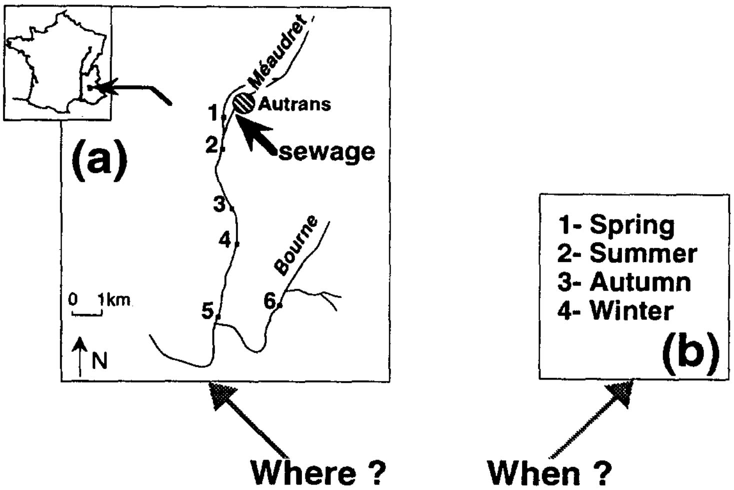
Co-inertia analysis

Co-inertia analysis was processed using the above data sets. The environmental variables were normalized by variables, whereas abundance of species were centered. There is a significant co-structure (permutation test significant, $p<0.001$) between species and physical and chemical variables. Furthermore, the arrows linking the co-inertia scores resulting from the environmental data set and that resulting from the faunistic data set are rather short (Fig. 3a). Co-inertia axis F1 separates polluted sites from non-polluted ones. A significant reduction in *Ephemeroptera* abundance (F1 >0 in Fig. 3b) is associated with high values of 5-day BOD, oxidation potential, conductivity, and phosphate and ammonia concentrations (F1>0 in Fig. 3c). Less polluted sites are characterized by higher pH, and higher oxygen concentrations (F1<0 in Fig. 3c), and are associated with a greater diversity of *Ephemeroptera* species relative to polluted sites (all species except *Caenis* sp. are situated on the negative side of co-inertia axis F1 in Fig. 3b). Co-inertia axis F2 distinguishes winter samples, (F2<0 in Fig. 3a) with low water temperature and nitrate concentrations (F2<0 in Fig. 3c), from summer samples with high water temperature and nitrate concentrations (F2>0 in Fig. 3c). *Ephemeroptera* taxa are positioned along co-inertia axis F2 according to season.

This analysis demonstrates an important overlap between the spatial effect characterized by the organic pollution and the temporal effect characterized mainly by water temperature variations (Fig. 3d).

Between-site co-inertia analysis

To separate such an overlap, we decided to focus on the spatial effect, since seasonal variations were relat-



Temp	Flow	pH	Cond	O2	BOD	Oxyd	NH4	NO3	PO4
10	41	8.5	295	110	2	1	0	3	0.1
13	62	8.3	325	95	2	1	0	3	0.1
1	25	8.4	315	91	1	0	0	6	0.0
3	118	8.0	325	100	1	1	0	1	0.2
11	158	8.3	315	13	7	3	2	2	1.5
13	80	7.6	380	20	21	5	9	0	3.6
3	63	8.0	425	38	36	8	12	2	6.5
3	252	8.3	360	100	9	2	2	4	1.6
11	198	8.5	290	113	3	1	0	4	0.1
15	100	7.8	385	46	15	2	7	7	4.5
2	79	8.1	350	84	7	1	2	13	3.7
3	315	8.3	370	100	8	2	2	4	2.8
12	280	8.6	290	126	3	1	0	4	0.7
16	140	8.0	360	76	12	2	4	8	3.4
3	85	8.3	330	106	2	1	0	12	1.6
3	498	8.3	330	100	4	1	1	4	0.8
13	322	8.5	285	117	3	1	0	4	0.8
15	160	8.4	345	91	1	1	0	10	1.7
2	72	8.6	305	91	1	0	0	9	1.2
2	390	8.2	330	100	1	1	0	5	0.6
11	303	8.5	245	100	1	0	0	2	0.2
13	310	8.2	285	82	8	1	0	3	0.6
4	181	8.6	270	105	2	0	0	3	0.4
3	480	8.2	290	100	1	0	0	2	0.1

(c)

11
12
13
14
21
22
23
24
31
32
33
34
41
42
43
44
51
52
53
54
61
62
63
64

Eda	Bsp	Brh	Bnl	Bpu	Cen	Ecd	Rhi	Hla	Hab	Par	Cae	Eig
4	7	A	9	0	0	0	5	9	0	4	0	0
6	7	A	0	A	0	0	2	7	0	0	0	2
4	5	8	0	9	6	0	5	9	0	7	0	0
3	6	7	0	6	7	0	4	8	0	4	0	0
0	0	8	0	0	0	0	0	4	0	0	0	0
0	0	9	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0
0	3	6	0	0	5	0	4	3	0	1	0	0
0	5	5	0	0	0	0	2	5	0	0	0	0
0	6	8	0	0	2	0	0	0	0	0	0	0
0	9	A	0	0	0	0	0	4	0	3	0	0
0	0	3	0	0	1	0	1	0	0	0	0	0
0	3	6	0	0	0	0	3	6	0	0	0	0
0	7	B	0	0	2	0	0	2	0	0	5	5
0	A	D	0	0	3	0	5	5	1	4	2	4
0	6	A	0	0	5	1	3	5	0	2	0	0
0	5	6	0	0	0	5	0	4	0	0	0	4
0	6	9	2	3	0	4	0	0	0	0	2	7
2	A	C	0	4	0	8	4	4	2	5	1	6
1	9	B	0	3	6	8	3	5	2	5	0	0
2	7	9	4	0	0	0	5	6	0	2	0	A
0	8	9	3	4	0	2	4	3	0	0	0	8
0	C	D	0	8	3	6	9	7	0	0	0	5
0	9	C	0	7	4	5	8	6	0	1	0	0

(d)

Fig. 1. Data and site description. (a) Sampling sites. (b) Sampling dates. (c) Physical and chemical measurements (abbreviated as follows: Temp = Temperature; Flow = Flow; pH = pH; Cond = Conductivity; O2 = Oxygen; BOD = 5-day Biological Oxygen Demand; Oxyd = Oxidation potential; NH4 = Ammonia; NO3 = Nitrate; and PO4 = Phosphate). (d) Distribution of *Ephemeroptera* taxa (abbreviated as follows: Eda = *Ephemera danica*; Bsp = *Baetis sp.*; Brh = *Baetis rhodani*; Bni = *Baetis niger*; Bpu = *Baetis pumilus*; Cen = *Centroptilum sp.*; Ecd = *Ecdyonurus sp.*; Rhi = *Rhithrogena sp.*; Hla = *Habrophlebia lauta*; Hab = *Habroleptoides sp.*; Par = *Paraleptophlebia sp.*; Cae = *Caenis sp.*; and Eig = *Ephemerella ignita*. The values represent abundance classes (from 0 to D with A = 10, B = 11, C = 12, and D = 13).

ed to water tempearture as expected. By isolating this effect, we found an agreement between the between-site structures resulting from species and those resulting from physical and chemical variables. Since the distance between the centers of gravity of the environmental data set (as white circles in Fig. 4a) and the centers of gravity of the faunistic data set (as grey circles in Fig. 4a) are rather short, there is a significant co-structure (permutation test significant, $p<0.001$)

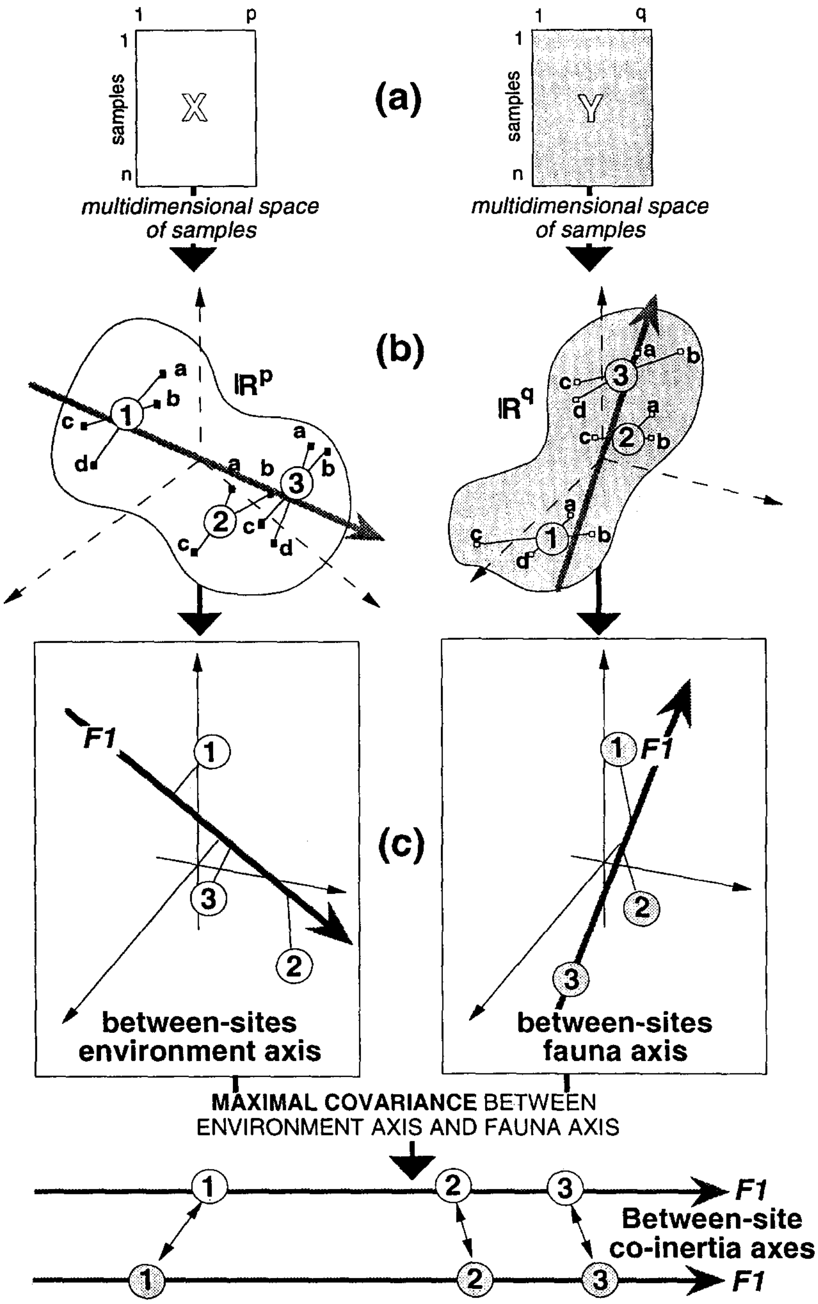


Fig. 2. Data processing. (a) The two tables analyzed (let X be the environmental table and Y be the faunistic table) appear in a '*p*' and a '*q*' multidimensional space respectively. (b) A between-site principal component analysis (PCA) can be processed on each data set separately. This results in finding one or several axes (grey arrows), so that the projected inertia of the centers of gravity of site-classes onto these axes is maximal (in white for the environmental data set, and in light grey for the faunistic data set). In this example, the measurements of environmental data and faunistic data were made in three sites on four occasions. The centers of gravity of site-classes are identified by a circle number (from 1 to 3), and the sampling dates are identified by a letter (from a to d). Each sampling date is linked to the corresponding centre of gravity (site) by a line. As often encountered in aquatic ecology, and for various reasons (flood, vandalism, etc.), a sample may be removed (2d, *i.e.*, site 2 was not sampled at date d). However, the method works even if the sampling design is not complete. (c) Between-site co-inertia analysis consists in matching two between-site analyses. For example, the analysis consists in finding out a combination in each table of average values maximizing the covariance among the between-site environmental axes and the between-site faunistic axes. In this example, only the first axes (F1) are presented. The standardization of the two resulting sets of between-site co-inertia scores makes it possible to compare the ordination of sites (centers of gravity) at the same scale.

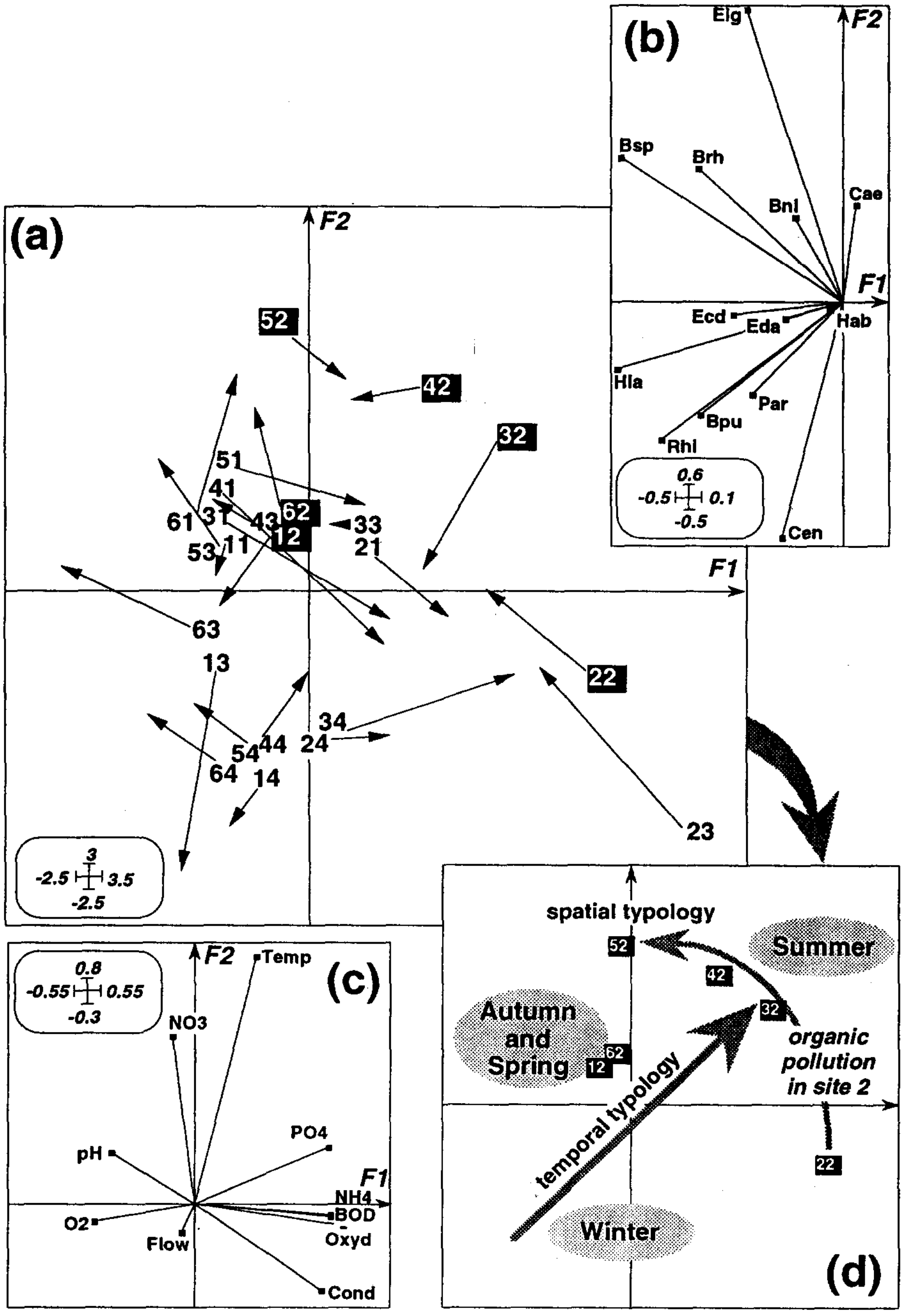


Fig. 3. Results of the co-inertia analysis processed on the tables of Fig. 1c and d. (a) Standardized co-inertia scores of the environmental and faunistic data sets projected onto the F1 × F2 factorial map. Arrows link environmental scores to faunistic ones. Numbers are situated at the environmental end of the arrows and they identify the positions in space (from 1 to 6 sampling sites) and in time (from 1 to 4 sampling dates) of the samples. (b) Co-inertia scores of the *Ephemeroptera* taxa on the F1 × F2 factorial map (see legend for taxa in Fig. 1d). (c) Co-inertia scores of the physical and chemical variables on the F1 × F2 factorial map (see legend for physical and chemical variables in Fig. 1c). (d) Interpretation of the F1 × F2 factorial map of samples shown in (a) demonstrating the overlap between temporal and spatial typologies.

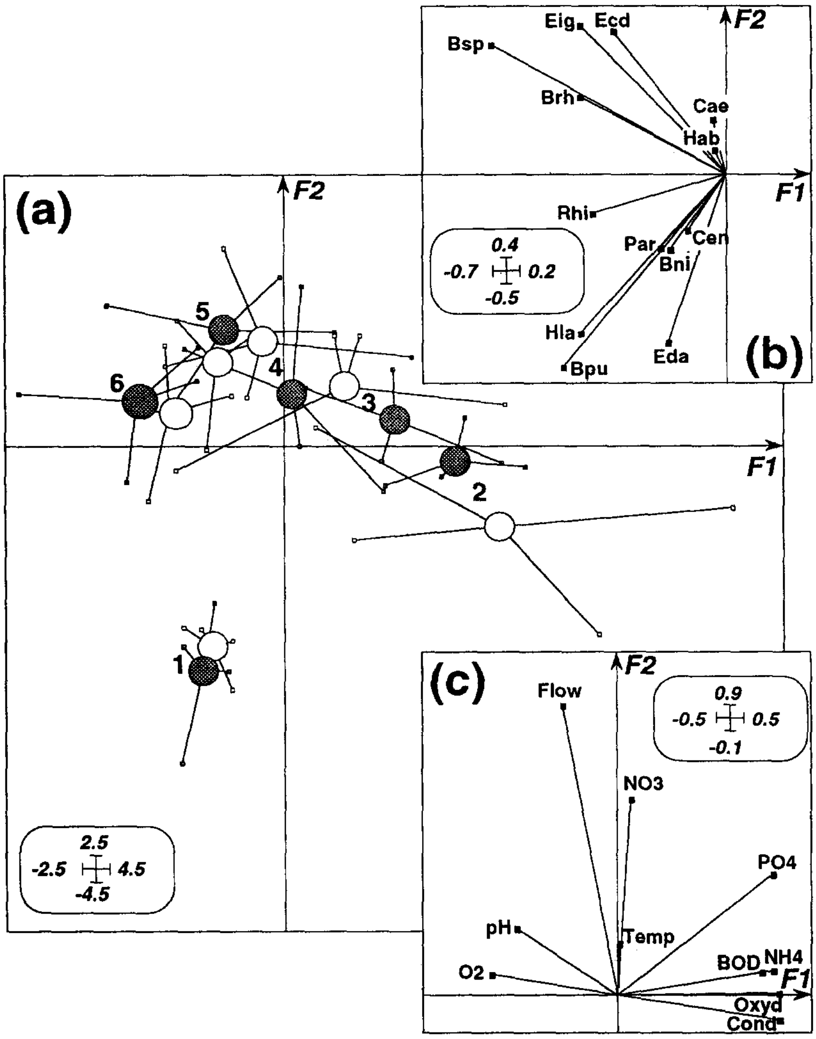


Fig. 4. Between-site co-inertia analysis. (a) Standardized co-inertia of the environmental and faunistic data sets are projected onto the F1 × F2 factorial map. Circles identify the positions of the centers of classes (from site 1 to 6) of the between-site environmental (white circles) and the between-site faunistic (grey circles) ordinations. The small black squares indicate the position of the sampling dates for the environmental (white) and faunistic (grey) data sets. (b) Co-inertia scores of the *Ephemeroptera* taxa on the F1 × F2 factorial map (see legend for taxa in Fig. 1d). (c) Co-inertia scores of the physical and chemical variables on the F1 × F2 factorial map (see legend for physical and chemical variables in Fig. 1c).

between *Ephemeroptera* and physical and chemical variables.

As a result, we were able to separate non-polluted stations (1 and 6) characterized by high values of pH and oxygen (Fig. 4a and Fig. 4c) from polluted ones (2 and 3) showing high values of conductivity, Oxidation potential, BOD, and ammonia concentration. Restoration of the river, demonstrated by the variation of nitrate concentrations, occurs from station 3 to station 5. Flow discharge increases from upstream (station 1) to downstream as expected. A low abundance of all the taxa is found in sites 2 and 3 (Fig. 4b). High abundances of *Habrophlebia lauta* (Hla), *Baetis pumilus* (Bpu), and *Ephemera danica* (Eda) characterize the upstream, whereas high abundances of *Baetis sp.*, *Baetis rhodani* (Brh), *Ephemerella ignita* (Eig), and *Ecdyonurus sp.* (Ecd) are found commonly downstream (Fig. 4b). Family, one should note the distance between the two typologies (Fig. 4a), indicating that the link between the environmental typology and the faunistic typology may be more or less pronounced according to site. In particular, in site 2, the co-structure between fauna and environmental variables is significantly affected.

Conclusion

The approach presented in this paper demonstrates that the physical and chemical variables and the faunistic composition are significantly related considering the spatial co-structure. As a result, between-site co-inertia analysis make it possible to ignore the temporal effect, which interferes with the spatial effect in the simple co-inertia analysis. The incorporation of space (or time) in the analysis enables the study of a hidden spatial (or temporal) structure, as it has already been demonstrated in simple multivariate methods (Beffy & Dolédec, 1991). Moreover, the study of the spatial variation of the co-structure demonstrates that the impact of a particular event such as a pollution discharge can reduce the intensity of such a co-structure.

Co-inertia analysis shows a great plasticity since it is now possible to make between-class co-inertia analyses. Within-class co-inertia analyses are also available, and even a combination of between- and within class co-inertia analyses may be processed. The use of such an alternative depends on the objectives of the researcher.

Software to perform co-inertia analysis and between-classes co-inertia analysis is incorporated in ADE version 3.6 (Chessel & Dolédec, 1993). ADE is available free for research and teaching on request to the last author of this paper.

References

Beffy, J. L. & S. Dolédec, 1991. Mise en évidence d'une typologie spatiale dans le cas d'un fort effet temporel: un exemple en hydrobiologie. Bull. Ecol. 22: 169–178.

Bornette, G., C. Amoros & D. Chessel, 1994. Rejuvenation in former braided channels of the Rhone River: successional patterns and allogenic processes. J. Veg. Sci. 5: 237–246.

Cazes, P., 1980. L'analyse de certains tableaux rectangulaires décomposés en blocs: généralisation des propriétés rencontrées dans l'étude des correspondances multiples. I. Définitions et applications à l'analyse canonique des variables qualitatives. Cah. de l'An. des Données 5: 145–161.

Chessel, D. & S. Dolédec, 1993. A.D.E. version 3.6: Hypercard Stacks & QuickBasic Microsoft Programme Library for the Analysis of Environmental data. User's manual. URA CNRS 1451, Université Lyon 1 (France), 725 pp.

Chessel, D. & P. Mercier, 1993. Couplage de triplets statistiques et liaisons espèces-environnement in J. D. Lebreton and B. Asselain (eds), Biométrie et environnement. Masson, in press.

Dolédec, S. & D. Chessel, 1987. Rythmes saisonniers et composantes stationnelles en milieu aquatique. I. Description d'un plan d'observation complet par projection de variables. Acta Oecol., Oecol. Gen. 8: 403–426.

Dolédec, S. & D. Chessel, 1989. Rythmes saisonniers et composantes stationnelles en milieu aquatique. II. Prise en compte et élimination d'effets dans un tableau faunistique. Acta Oecol., Oecol. Gen. 10: 207–232.

Dolédec, S. & D. Chessel, 1991. Recent developments in linear ordination methods for environmental sciences. Adv. Ecol. 1: 133–155.

Dolédec, S. & D. Chessel, 1994. Co-inertia analysis: an alternative method for studying species-environment relationships. Freshwat. Biol. 31: 277–294.

Johnson, R. K. & T. Wiederholm, 1989. Classification and ordination of profundal macroinvertebrate communities in nutrient poor, oligo-mesohumic lakes in relation to environmental data. Freshwat. Biol. 21: 375–386.

Lebreton, J. D., R. Sabatier, G. Banco & A. M. Bacou, 1991. Principal component and correspondence analyses with respect to instrumental variables: an overview of their role in studies of structure-activity and species-environment relationships. In Applied Multivariate Analysis in SAR and Environmental Studies, J. Devillers & W. Karcher (eds). Kluwer Academic Publishers: 85–114.

Mercier, P., D. Chessel & S. Dolédec, 1992. Complete correspondence analysis of an ecological profile data table: a central ordination method. Acta Oecol. 13: 25–43.

Pegaz-Mauret, D., 1980. Impact d'une perturbation d'origine organique sur la dérive des macroinvertébrés d'un cours d'eau. Comparaison avec le benthos. Thèse de 3ème cycle, Université Lyon 1 (France), 130 pp.

Romane, F., 1972. Un exemple d'utilisation de l'analyse factorielle des correspondances en écologie végétale. In E. van der Maarel & R. Tüxen (eds), Basic problems and methods in phytosociology, Dr W. Junk Publishers, The Hague: 151–167.

Rutt, G. P., N. S. Weatherly & S. J. Ormerod, 1990. Relationships between the physicochemistry and macroinvertebrates of british upland streams: the development of modelling and indicators systems for predicting fauna and detecting acidity. Freshwat. Biol. 24: 463–480.

Storey, A. W., S. E. Bunn, P. M. Davies & D. H. Edward, 1990. Classification of the macroinvertebrate fauna of two rivers systems in southwestern australia in relation to physical and chemical parameters. Regulated Rivers 5: 217–232.

ter Braak, C. J. F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

ter Braak, C. J. F., 1987. The analysis of vegetation–environment relationships by canonical correspondence analysis. Vegetatio 69: 69–77.

ter Braak, C. J. F., 1988. Partial Canonical Analysis. In: Classification and related methods of data analysis, H. H. Bock, (ed.). Elsevier Science Publishers: 551–558.

Townsend, C. R., A. G. Hildrew & J. Francis, 1984. Community structure in some southern English streams: the influence of physicochemical factors. Freshwat. Biol. 14: 297–310.

Tucker, L. R., 1958. An inter-battery method of factor analysis. Psychometrika 23: 111–134.

Wright, J. F., P. D. Armitage, M. T. Furse & D. Moss, 1989. Prediction of invertebrtaes communities using stream measurements. Regulated Rivers 4: 147–155.

Yoccoz, N. & D. Chessel, 1988. Ordination sous contraintes de relevés d'avifaune: élimination d'effets dans un plan d'observations à deux facteurs. C.R. acad. Sci. Paris 307: 189–194.