



**HAL**  
open science

## Les analyses multitableaux en ecologie factorielle.

J. Thioulouse, D. Chessel

► **To cite this version:**

J. Thioulouse, D. Chessel. Les analyses multitableaux en ecologie factorielle. : De la typologie d'état à la typologie de fonctionnement par l'analyse triadique. *Acta Oecologica, Oecologia Generalis*, 1987, 8 (4), pp.463-480. hal-00698250

**HAL Id: hal-00698250**

**<https://hal.science/hal-00698250>**

Submitted on 14 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Les analyses multitableaux en écologie factorielle

### I. — *De la typologie d'état à la typologie de fonctionnement par l'analyse triadique*

J. Thioulouse <sup>(1)</sup> et D. Chessel <sup>(2)</sup> (\*)

<sup>(1)</sup> *Laboratoire de Biométrie, U. A. C. N. R. S. n° 243, Université Lyon 1, 43, Boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex,*

<sup>(2)</sup> *Écologie des eaux douces, U. A. C. N. R. S. n° 367, Université Lyon 1, 43, Boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex*

#### RÉSUMÉ

L'analyse triadique permet d'aborder les problèmes soulevés par l'analyse statistique des tableaux de relevés de milieu multivariés et répétés au cours du temps. Elle constitue un cas particulier des analyses dites « multitableaux » et bénéficie donc de la puissance descriptive des outils théoriques sous-jacents à ces méthodes : l'interstructure propose une ordination des dates d'échantillonnage et une typologie globale des processus spatiaux; le compromis permet d'exprimer les structures communes aux diverses dates, du point de vue des variables et des stations; l'intrastructure offre une description fine des écarts à ce modèle commun, pour chacune des dates. La démarche empruntée ouvre la voie à de nombreuses généralisations de ce type d'analyses. Une procédure de calcul simple, ne faisant intervenir que l'analyse en composantes principales est proposée.

**MOTS-CLES :** *Variables physico-chimiques - Structures spatio-temporelles - Analyses multitableaux - Analyse triadique - Analyse en composantes principales.*

#### ABSTRACT

Triadic analysis is well adapted to the statistical study of multivariate surveys that are collected more than once. Triadic analysis is a particular case of the so-called "multi-table analyses", and therefore benefits from the descriptive power of the underlying theoretical tools of these methods. The interstructure proposes an ordination of the sampling dates and an overall typology of spatial processes. The compromise expresses the structures common to each date, from both the viewpoints of variables and sampling sites. The intrastructure offers a detailed description of deviations from the common model of the compromise for each date. The method proposed here leads to many generalizations concerning this kind of analysis. We present a simple computational procedure, using only principal components analysis.

**KEY-WORDS:** *Physico-chemical variables - Spatial-temporal structures - Multi-table analyses - Triadic analysis - Principal components analysis.*

---

(\*) Programme Interdisciplinaire de Recherche sur l'Environnement (P. I. R. E. N.) du C. N. R. S. « Recherches méthodologiques appliquées à la gestion écologique des systèmes fluviaux » dirigé par A. L. Roux.

## INTRODUCTION

L'analyse statistique des tableaux de relevés écologiques multivariés peut à l'heure actuelle être abordée à l'aide des méthodes classiques de l'écologie factorielle. Analyse en composantes principales (*ACP*), analyse factorielle des correspondances (*AFC*), analyse des correspondances multiples (*ACM*), analyse discriminante (*AD*), sont des méthodes issues de l'analyse des données, dont l'usage s'est largement répandu, en partie grâce aux progrès de la micro-informatique et à la diffusion de logiciels d'analyse multivariée.

Toutefois, il est fréquent que la série des relevés soit répétée (de façon plus ou moins rigoureuse) au cours du temps, conduisant ainsi à l'obtention de tableaux où chaque élément possède trois indices,  $x_{ijk}$  représentant la mesure dans la station  $i$  ( $i = 1, \dots, n$ ) de la variable  $j$  ( $J = 1, \dots, p$ ) à la date  $k$  ( $k = 1, \dots, t$ ). L'analyse de tels tableaux soulève alors des problèmes que les méthodes multivariées de base ne permettent pas d'aborder de façon satisfaisante. Il est de plus possible de généraliser ce point de vue, en introduisant la notion (issue de l'analyse de variance) de « plan d'observation » (ou encore de « plan d'expérience »). Dans le cas particulier d'une série de relevés multivariés répétée au cours du temps, le plan d'expérience est simplement constitué par la liste des valeurs prises par les indices  $i$  et  $k$  définis ci-dessus pour chaque variable. Lorsque ce plan est complet (*i. e.* lorsque toutes les stations ont été analysées à chaque date), il est possible de faire appel à, diverses méthodes statistiques adaptées au dépouillement de tels ensembles de données.

L'article de DOLEDEC et CHESSEL (1987) décrit un premier type d'approche basé sur l'*ACP* des tableaux obtenus par projection sur les sous-espaces vectoriels définis par le plan d'expérience, et l'*ACP* des résidus par rapport à ces projections. Cette stratégie introduit l'usage des méthodes de description par projection sur un sous-espace choisi par l'expérimentateur, dont l'intérêt dépasse le cadre de la simple prise en compte du protocole de récolte des données. Le principe intuitif de ces méthodes consiste à rechercher dans les données la part de variabilité qui dépend d'un facteur choisi *a priori* (par exemple l'espace ou le temps), ainsi que la part qui en est indépendante, ou qui ne peut s'expliquer que par une interaction avec d'autres facteurs.

L'approche que nous présentons ici est totalement différente de la précédente. Basée sur le principe des méthodes multitableaux, elle propose un cadre théorique entièrement adapté à l'étude des tableaux à trois indices; encore peu connues des écologues, ces méthodes demandent cependant une évaluation de leur impact. Contrairement à l'approche basée sur les projecteurs, on ne cherche pas à décomposer la variabilité des mesures sur divers facteurs (stratégie de l'analyse de variance) puis à décrire chacune de ces composantes, mais à décrire directement les structures induites dans les données par l'organisation du plan d'expérience. Chaque élément  $x_{ijk}$  est alors considéré comme appartenant à un tableau rectangulaire (par exemple un tableau [stations  $\times$  variables] à  $n$  lignes et  $p$  colonnes), la juxtaposition de ces tableaux formant le cube des données ( $t$  tableaux [stations  $\times$  variables] formant un cube [stations  $\times$  variables  $\times$  dates]).

L'analyse triadique (JAFFRENOU, 1978) est, parmi les méthodes multitableaux, l'une des plus aisées à mettre en œuvre mais, à notre connaissance, n'a jamais fait l'objet d'une utilisation effective. Pour exposer les résultats qu'elle permet d'obtenir, nous avons utilisé le même jeu de données que DOLEDEC et CHESSEL (1987),

*Acta Œcologica/Œcologia Generalis*

constitué par les mesures de dix variables physico-chimiques dans six stations réparties le long du Méaudret (ruisseau du Vercors septentrional, affluent de la Bourne), à quatre dates. Nous avons cependant retiré la station 6 (située en fait sur la Bourne) et la variable 5 (oxygène en % de saturation, qui prend la même valeur dans toutes les stations en février). Ces données figurent au tableau 1. La figure 1 montre une représentation graphique des données après centrage et réduction par variable à chaque date. Elle correspond au point de vue qui consiste à considérer les données comme une succession chronologique de  $t$  tableaux [stations  $\times$  variables].

TABLEAU 1. — Données traitées par l'analyse triadique. Les cinq groupes de quatre lignes correspondent aux cinq stations et aux quatre dates d'échantillonnage. Les 9 colonnes correspondent aux 9 variables mesurées (Doledec & Chessel, 1987) : 1 : Température de l'eau (°C). 2 : Débit (l/s). 3 : pH. 4 : Conductivité (mmho/cm). 5 : DBO<sub>5</sub> (mg/l d'oxygène). 6 : Oxydabilité (mg/l d'oxygène). 7 : Ammoniaque (mg/l de NH<sub>4</sub><sup>+</sup>). 8 : Nitrates (mg/l de NO<sub>3</sub><sup>-</sup>). 9 : Orthophosphates (mg/l de PO<sub>4</sub><sup>3-</sup>).

	1	2	3	4	5	6	7	8	9	
Juin	10.00	41.00	8.50	295.00	2.30	1.40	0.12	3.40	0.11	Station A
Août	13.00	62.00	8.30	325.00	2.30	1.80	0.11	3.00	0.13	
Nov.	1.00	25.00	8.40	315.00	1.60	0.50	0.07	6.40	0.03	
Fév.	3.00	118.00	8.00	325.00	1.60	1.20	0.17	1.80	0.19	
Juin	11.00	158.00	8.30	315.00	7.60	3.30	2.85	2.70	1.50	Station B
Août	13.00	80.00	7.60	380.00	21.00	5.70	9.80	0.80	3.65	
Nov.	3.00	63.00	8.00	425.00	36.00	8.00	12.50	2.20	6.50	
Fév.	3.00	252.00	8.30	360.00	9.50	2.90	2.52	4.60	1.60	
Juin	11.00	198.00	8.50	290.00	3.30	1.50	0.40	4.00	0.10	Station C
Août	15.00	100.00	7.80	385.00	15.00	2.50	7.90	7.70	4.50	
Nov.	2.00	79.00	8.10	350.00	7.10	1.90	2.70	13.20	3.70	
Fév.	3.00	315.00	8.30	370.00	8.70	2.80	2.80	4.80	2.85	
Juin	12.00	280.00	8.60	290.00	3.50	1.50	0.45	4.00	0.73	Station D
Août	16.00	140.00	8.00	360.00	12.00	2.60	4.90	8.40	3.45	
Nov.	3.00	85.00	8.30	330.00	2.00	1.40	0.42	12.00	1.60	
Fév.	3.00	498.00	8.30	330.00	4.80	1.60	1.04	4.40	0.82	
Juin	13.00	322.00	8.50	285.00	3.60	1.60	0.48	4.60	0.84	Station E
Août	15.00	160.00	8.40	345.00	1.70	1.90	0.22	10.00	1.74	
Nov.	2.00	72.00	8.60	305.00	1.60	0.90	0.10	9.50	1.25	
Fév.	2.00	390.00	8.20	330.00	1.70	1.20	0.56	5.00	0.60	

L'ANALYSE TRIADIQUE :  
CADRE D'UTILISATION ET OBJECTIFS

Si l'analyse triadique (AT) est effectivement simple à utiliser, le support théorique auquel elle fait appel est par contre plus complexe que celui des analyses à un seul tableau. Nous rappelons dans l'annexe les principaux éléments sur lesquels elle est basée, une bonne compréhension de la théorie mathématique sous-jacente n'étant possible que par la lecture de la thèse de JAFFRENOU (1978). Remarquons seulement qu'elle se conforme au schéma classique des analyses multitableaux, résumé dans l'article d'ESCOUFIER (1980). Interstructure, compromis et intrastructure permettent en effet de définir trois étapes essentielles, tant sur le plan de la théorie mathématique, que sur celui du calcul et de l'interprétation des résultats (PERNIN, 1986). Si les termes utilisés peuvent paraître nouveaux à celui qui n'est pas habitué à la pratique des méthodes

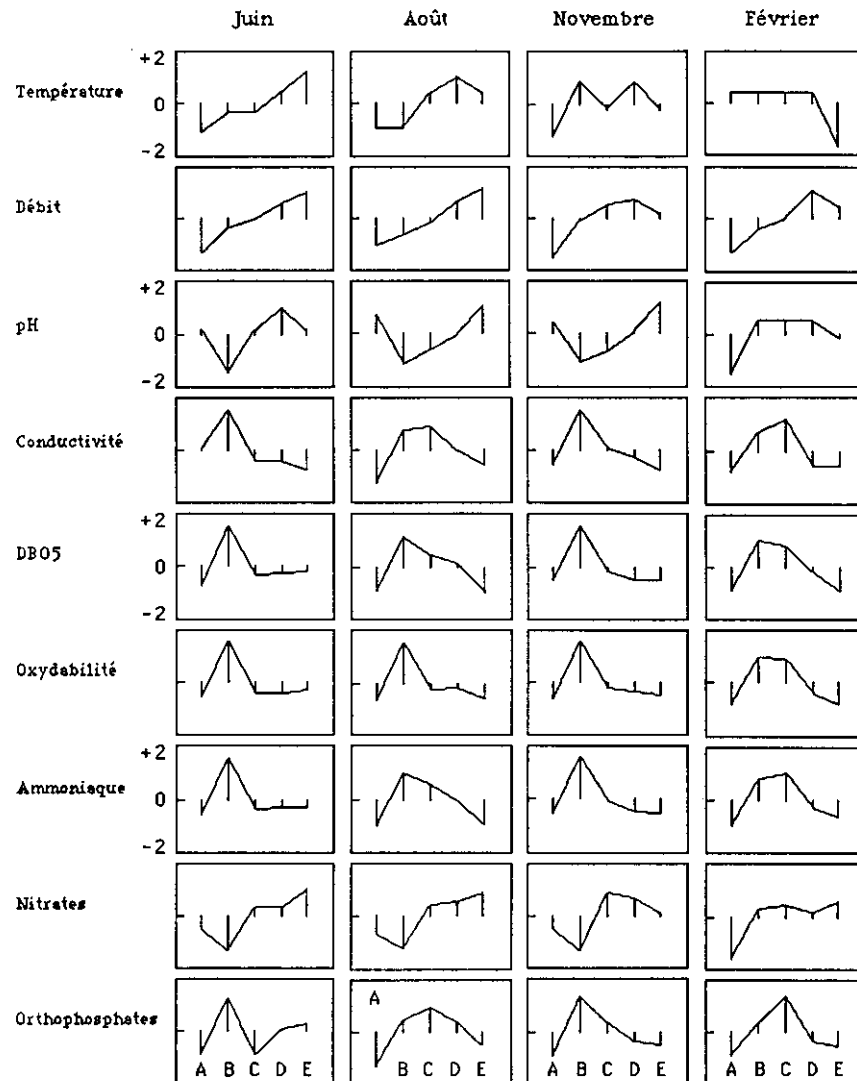


FIG. 1. — Représentation graphique des neuf variables (en lignes) mesurées à quatre dates (en colonnes) dans les cinq stations (A, B, C, D et E). La hauteur des bâtons est proportionnelle à la valeur de la variable après centrage et réduction (par date et par variable). Cette figure permet en particulier de comparer les profils (sur les cinq stations) des diverses variables, et ceci aux quatre dates d'échantillonnage.

multitableaux, ils sont néanmoins consacrés par l'usage, et recouvrent des notions non ambiguës, dont nous espérons faciliter la mise en pratique dans le domaine de l'écologie factorielle.

L'AT présente, par rapport aux autres analyses mettant en œuvre la notion d'opérateur (PAGES et coll., 1976, ROBERT et ESCOUFIER, 1976), la contrainte de ne s'appliquer qu'au cas où les données sont constituées par la mesure des *mêmes variables* sur les *mêmes individus* (ici les stations) à plusieurs reprises (c'est-à-dire, en général, à plusieurs dates).

*Acta Œcologica/Œcologia Generalis*

Dans le cas où les variables ou les stations diffèrent d'une date à l'autre (par exemple simplement lorsque leurs effectifs sont différents), la méthode STATIS (L'HERMIER DES PLANTES, 1976) permet, au prix d'une plus grande complexité théorique et pratique, d'obtenir le même type de résultats. Ces résultats sont généralement présentés en trois étapes, qui reflètent les étapes des calculs (GLAÇON, 1981) :

— la première étape, appelée *interstructure*, propose une description globale des dates d'échantillonnage. Dans le cas particulier de l'AT, on obtient de plus pour chaque variable, une description des stations relative à la typologie des dates d'échantillonnage,

— la seconde étape permet de définir un « compromis » entre les diverses dates d'échantillonnage, c'est-à-dire un résumé le plus proche possible des situations à chaque date. Dans le cas des méthodes basées sur les opérateurs, ce compromis est lui-même un opérateur; dans le cas de l'AT, il s'agit d'un tableau [stations × variables] fictif, présentant certaines propriétés d'optimalité. En particulier, les variables de ce tableau maximisent la moyenne des corrélations avec les variables correspondantes de chacun des tableaux (cf. annexe). L'analyse de ce compromis (par une ACP) fournit, dans l'AT, une représentation simultanée des stations et des variables (ou des dates et des variables) et non pas seulement une représentation des stations (ou des dates) comme dans les méthodes à opérateurs « classiques ». Un autre avantage réside dans la possibilité de construire plusieurs compromis successifs, associés aux facteurs successifs de l'interstructure, alors que dans les analyses pour lesquelles les individus (ou les variables) diffèrent à chaque date, on n'obtient qu'un seul compromis, associé au premier vecteur propre de l'interstructure.

— la troisième étape, ou *intrastructure*, consiste à représenter simultanément chaque station à toutes les dates, et, de façon symétriques chaque variable à toutes les dates. L'intérêt de ces représentations est de permettre une analyse fine de l'évolution de la typologie des stations ou des variables au cours du temps. Divers procédés mathématiques permettent d'atteindre cet objectif. Le plus simple est celui que nous utiliserons ici; il consiste à projeter en éléments supplémentaires les variables et les stations de tous les tableaux initiaux dans l'ACP du compromis (GLAÇON, 1981).

L'AT offre en fait à l'utilisateur la possibilité de considérer ses données sous deux aspects différents :

— soit une suite chronologique de tableaux [stations × variables] (un tableau par date d'échantillonnage),

— soit une série de tableaux [dates × variables] (un tableau par station).

Les résultats obtenus différeront selon la solution adoptée, et seront aussi différents de ceux obtenus par la méthode STATIS. Le choix entre les deux possibilités est ici strictement lié aux objectifs de l'étude. Nous avons préféré privilégier l'étude de la structure (pour les différentes variables) et de l'évolution (chronologique) du gradient amont-aval défini par la localisation géographique des cinq stations (BOURNAUD et coll., 1983). Cet objectif correspond à considérer les données comme une suite chronologique de tableaux [stations × variables], stratégie qui est celle de la représentation graphique de la figure 1. L'autre point de vue (un tableau [dates × variables] par station) aurait privilégié l'étude du cycle annuel des mesures des différentes variables et sa variabilité spatiale. Nous avons choisi la première stratégie et nous désirons de cette manière approcher les éventuelles modifications du gradient amont-aval dans la perspective de l'écologie fonctionnelle : la précision des descriptions obtenues par les méthodes multitableaux est telle qu'il convient de définir sans ambiguïté le but de l'analyse.

## PROCÉDURE DE CALCUL

### *L'interstructure*

La première étape des calculs consiste à centrer et réduire les données par date et par variable. On réorganise ensuite les dates et variables de façon à obtenir le tableau Z de l'annexe, qui présente en colonne les quatre dates d'échantillonnage et en ligne les neuf variables, pour chacune des cinq stations (cinq paquets de neuf lignes). Le tableau 2 permet de vérifier les calculs. L'annexe

TABLEAU 2. — Ce tableau contient les mêmes valeurs que le tableau 1, mais centrées et réduites par variable à chaque date, [et réorganisées de façon à obtenir les dates d'échantillonnage en colonnes et les variables] (à l'intérieur des stations) en lignes. C'est l'ACP de ce tableau qui constitue la première étape des calculs (interstructure).

	Juin	Août	Novembre	Février	
Température	-1.37	-1.17	-1.60	+0.50	
Débit	-1.62	-1.27	-1.88	-1.54	
pH	+0.20	+0.94	+0.56	-1.88	
Conductivité	+0.00	-1.53	-0.70	-0.98	
DBO5	-0.96	-1.09	-0.60	-1.09	Station A
Oxydabilité	-0.64	-0.77	-0.74	-0.98	
Ammoniaque	-0.74	-1.14	-0.65	-1.18	
Nitrates	-0.53	-0.86	-0.57	-1.97	
Orthophosphates	-1.05	-1.64	-1.14	-1.09	
	-0.39	-1.17	+1.07	+0.50	
	-0.43	-0.78	-0.08	-0.49	
	-1.83	-1.40	-1.31	+0.68	
	+1.91	+0.94	+1.87	+0.93	
	+1.94	+1.42	+1.98	+1.27	Station B
	+1.99	+1.95	+1.97	+1.27	
	+1.98	+1.33	+1.96	+1.04	
	-1.62	-1.49	-1.62	+0.41	
	+1.62	+0.61	+1.71	+0.41	
	-0.39	+0.50	-0.27	+0.50	
	-0.02	-0.23	+0.67	+0.00	
	+0.20	-0.73	-0.84	+0.68	
	-0.48	+1.17	+0.12	+1.47	
	-0.42	+0.62	-0.19	+1.03	Station C
	-0.50	-0.28	-0.23	+1.13	
	-0.46	+0.84	-0.10	+1.31	
	+0.40	+0.49	+1.14	+0.58	
	-1.07	+1.15	+0.48	+1.74	
	+0.59	+1.33	+1.07	+0.50	
	+0.82	+0.86	+0.95	+1.43	
	+1.22	-0.07	+0.09	+0.68	
	-0.48	+0.04	-0.35	-0.71	
	-0.31	+0.21	-0.57	-0.14	Station D
	-0.50	-0.21	-0.41	-0.45	
	-0.41	+0.08	-0.57	-0.36	
	+0.40	+0.69	+0.84	+0.24	
	+0.14	+0.48	-0.45	-0.42	
	+1.57	+0.50	-0.27	-2.00	
	+1.24	+1.41	+0.34	+0.59	
	+0.20	+1.27	+1.50	-0.17	
	-0.95	-0.63	-0.94	-0.71	
	-0.25	-1.17	-0.60	-1.06	Station E
	-0.36	-0.70	-0.59	-0.98	
	-0.38	-1.11	-0.64	-0.81	
	+1.34	+1.15	+0.21	+0.75	
	+0.35	-0.61	-0.60	-0.65	

statistique rappelle que l'interstructure des quatre tableaux est obtenue par l'ACP (*non normée*) du tableau Z.

#### Le compromis

Le compromis peut être considéré comme une solution au problème de la comparaison des ACP faites sur les tableaux de chacune des dates. Il est en effet tentant d'exécuter séparément

*Acta Oecologica/Oecologia Generalis*

les quatre ACP et de comparer les plans factoriels obtenus. L'analyse du compromis se présente alors comme « la meilleure ACP possible » par rapport à ces quatre ACP, au sens que l'inertie qu'elle prend en compte est maximale (car les variables du tableau compromis maximisent les corrélations avec les variables des tableaux à chaque date). Le compromis permet d'une part d'exprimer les *structures communes aux diverses dates*, et d'autre part de juger des *écarts à ce modèle* grâce à la projection des stations et des variables à chaque date dans un espace commun (étape de l'intrastructure).

Sur le plan mathématique, le compromis de l'AT est une combinaison linéaire des tableaux initiaux, dont les coefficients sont les composantes d'un des vecteurs propres de l'analyse de l'interstructure (annexe). En pratique, il suffira donc de prendre les coordonnées factorielles des lignes sur un des facteurs obtenus par l'ACP du tableau Z, et de les considérer comme un tableau [stations × variables] fictif (à cinq lignes et neuf colonnes), dont on effectuera l'ACP (non normée). Le tableau 3 permettra de vérifier les calculs; il présente le compromis correspondant au premier vecteur propre de l'interstructure (et donc à la première coordonnée factorielle des lignes de Z).

TABLEAU 3. — Premier compromis de l'analyse triadique. Il s'agit des coordonnées factorielles des lignes du tableau 2 sur le premier axe de l'ACP, réorganisées de façon à obtenir un tableau [stations × variables] fictif. Les cinq lignes correspondent aux cinq stations et les neuf colonnes correspondent aux neuf variables mesurées à chaque date : 1 : Température de l'eau. 2 : Débit. 3 : pH. 4 : Conductivité. 5 : DBO<sub>5</sub>. 6 : Oxydabilité. 7 : Ammoniaque. 8 : Nitrates. 9 : Orthophosphates.

	1	2	3	4	5	6	7	8	9
Station A	-2.02	-3.13	+0.20	-1.59	-1.83	-1.52	-1.80	-1.80	-2.46
Station B	-0.05	-0.87	-2.15	+2.86	+3.32	+3.64	+3.20	-2.36	+2.25
Station C	+0.12	+0.24	-0.50	+1.02	+0.41	-0.09	+0.67	+1.32	+1.02
Station D	+1.80	+1.96	+0.90	-0.68	-0.41	-0.76	-0.62	+1.13	-0.07
Station E	+0.15	+1.81	+1.55	-1.61	-1.50	-1.26	-1.46	+1.70	-0.73

Comme nous l'avons remarqué, il est possible d'utiliser les autres vecteurs propres pour constituer une série de compromis. Nous nous limiterons cependant ici à l'étude du premier compromis.

*L'intrastructure.*

L'étape de l'intrastructure autorise deux types de représentation : représentation des stations aux diverses dates et représentation des variables aux diverses dates. Pour obtenir la représentation des stations aux diverses dates, il suffit de projeter en éléments supplémentaires les 20 lignes du tableau 1 (après centrage et réduction par variable à chaque date) dans l'ACP du compromis. On obtient ainsi les coordonnées factorielles des 20 points correspondant au cinq stations pour les quatre dates de mesure.

La représentation des variables aux diverses dates est obtenue en projetant en éléments supplémentaires dans l'ACP du compromis les 36 colonnes du tableau centré et réduit par variable et par date comportant en lignes les 5 stations et en colonnes les quatre groupes de neuf variables. Du point de vue de la mise en œuvre informatique, remarquons que, si on utilise des fichiers séquentiels le même fichier peut être considéré comme contenant un tableau à cinq lignes (stations) et 36 colonnes (dates et variables) ou 20 lignes (station et dates) et neuf colonnes (variables).

**RESULTATS**

*L'interstructure*

La figure 2 présente divers graphiques permettant d'interpréter l'ACP du tableau Z. Le premier axe (70 % de l'inertie) représente classiquement un effet taille



(mais inter-tableaux) qui prend en compte la structure permanente des données, caractérisée ici par une pollution maximale dans la station *B*.

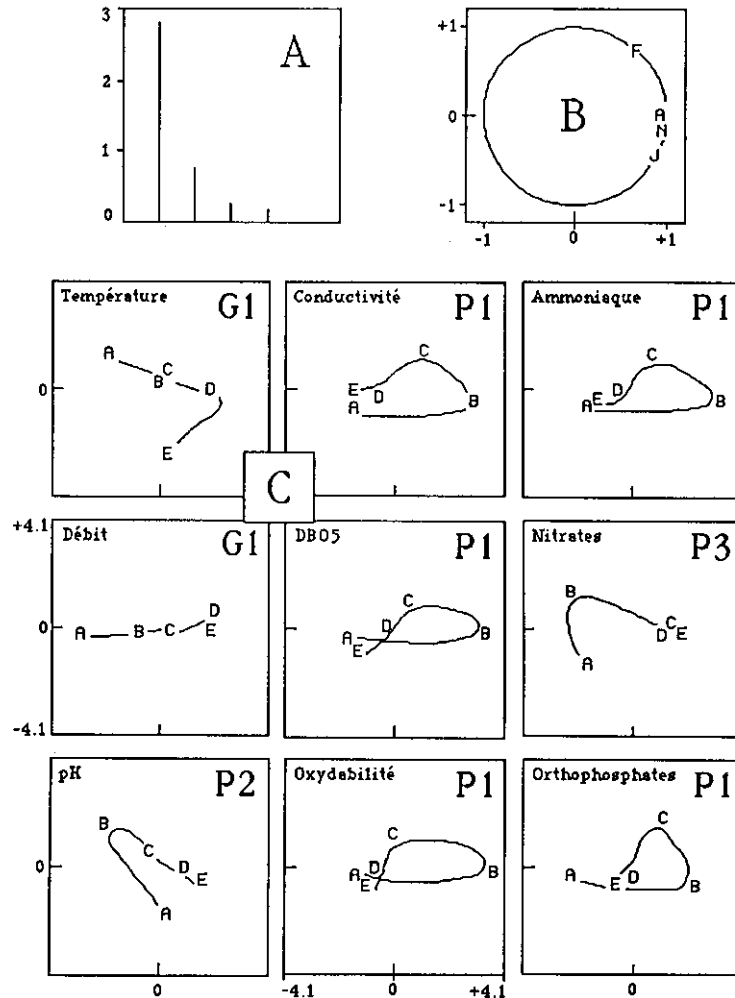


FIG. 2. — Analyse de l'interstructure. *A* : diagramme des quatre valeurs propres. *B* : cercle des corrélations des quatre dates d'échantillonnage (juin, août, novembre et février) avec les deux premiers facteurs de l'interstructure. *C* : plan factoriel  $F1 \times F2$ ; les 45 lignes du tableau 2 ont été représentées en neuf groupes de cinq, correspondant aux neuf variables mesurées dans chacune des cinq stations (*A*, *B*, *C*, *D* et *E*). Les quatre types de processus spatio-temporels analysés dans le texte sont notés ici *P1* (pollution-restauration), *P2* (acidification puis remontée du pH), *P3* (oxydation de l'azote ammoniacal en nitrates) et *G1* (gradient amont-aval de débit et température).

Le deuxième axe (19 % de l'inertie) rend compte de la spécificité des phénomènes hivernaux. La représentation des stations *pour chaque variable* (fig. 2 *C*)

*Acta Œcologica/Œcologia Generalis*

est particulièrement riche du point de vue de l'interprétation de l'analyse. Elle permet en effet d'établir une typologie des processus spatio-temporels sous-jacents aux observations. Quatre types de processus nettement différenciés apparaissent.

— Le plus simple correspond au *gradient amont-aval de température et de débit*. Il est pris en compte presque exclusivement par le premier axe; seule la température dans la station *E* s'en écarte, et se reporte sur le second axe. Conformément à l'interprétation de ce second axe, on peut observer (fig. 1) que la température est en effet minimale dans la station *E* en février. Il s'agit là d'un phénomène sans intérêt réel, lié au fait que la température, mesurée seulement au degré près, est la même dans toutes les stations en février sauf dans la station *E*, où elle est inférieure de 1° à celle des autres stations.

— Le second type de processus est celui de pollution-restauration du cours d'eau, caractérisé par les cinq variables suivantes : conductivité,  $DBO_5$ , oxydabilité, teneur en ammoniacque et en orthophosphates. La station *A*, non polluée, s'oppose sur le premier axe à la station *B*, dont la pollution est maximale. Le processus de restauration se traduit ensuite par un retour des stations *C*, *D* et *E* vers les valeurs négatives de  $F1$ . La restauration est en général assurée au niveau de la station *E*, sauf pour la teneur en orthophosphates (le retour aux données du tableau 1 permet de vérifier cette particularité). L'axe 1 rend bien compte du processus de pollution-restauration, mais l'axe 2 souligne les caractéristiques particulières des stations *C*, *D* et *E*, et révèle ainsi l'importance de la pollution hivernale (février). Cette dernière est même plus importante dans la station *C* que dans la station *B* pour la conductivité et les teneurs en ammoniacque et orthophosphates (tableau 1). Le processus de restauration du cours d'eau ne peut donc pas être considéré comme un simple gradient amont-aval d'auto-épuration mais fait intervenir une interaction spatio-temporelle entre pollutions estivale et hivernale et position des stations le long du cours d'eau.

— Le troisième type de processus est celui de l'acidification puis de la remontée du pH de l'eau. Il diffère nettement du précédent car ce sont ici les stations *A* et *B* qui s'opposent sur l'axe 2. Le premier axe correspond au gradient amont-aval du pH annuel moyen, avec une exception pour la station *A*, qui occupe une position intermédiaire entre les stations *C* et *D*. L'opposition des stations *A* et *B* sur l'axe 2 provient du fait que le pH chute en février dans la station *A* (ce qui explique la faiblesse de sa moyenne annuelle), alors que dans la station *B*, il est maximal à cette même date. L'interprétation de cette interaction spatio-temporelle dans les variations du pH reste à discuter.

— Le processus d'oxydation de l'azote ammoniacal en nitrates ressemble au précédent par son opposition entre les stations *A* et *B*, mais en diffère par le fait que les stations *C*, *D* et *E* sont très proches les unes des autres. L'axe 1 est ici aussi très lié aux moyennes annuelles des mesures par station, et l'opposition entre les stations *A* et *B* est due au fait que la teneur en nitrates est minimale en février dans la station *A*, alors qu'elle est maximale à la même date dans la station *B* (tableau 1). L'interprétation de ces observations en termes de physico-chimie du milieu aquatique revient au spécialiste de ce domaine.

La figure 3, qui superpose les neuf graphiques de la figure 2 C fournit une synthèse des relations entre les variables des cinq stations, et permet d'observer la superposition des quatre processus qui viennent d'être décrits. On notera en particulier l'évolution conjointe du groupe de variables caractérisant la pollution

d'une station à l'autre. Ajoutons qu'il serait possible de poursuivre la discussion en étudiant les représentations des diverses variables regroupées par station, au lieu de regrouper les stations pour chaque variable (fig. 2 C).

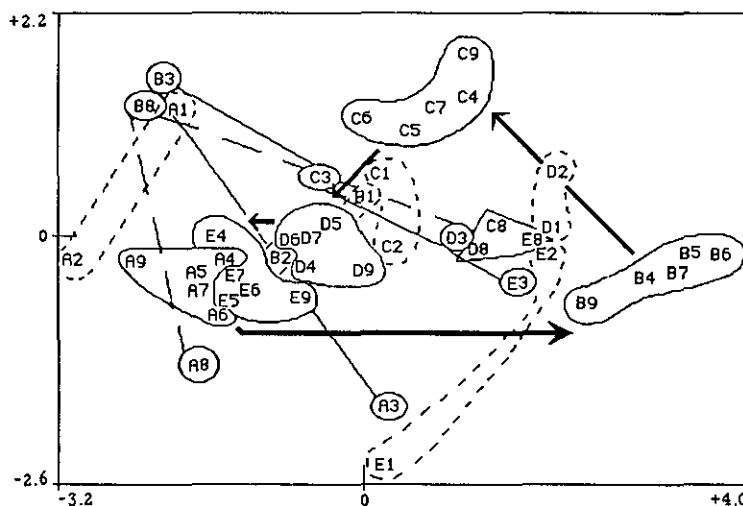


FIG. 3. — Superposition des neuf graphiques de la figure 2 C. Cette figure propose une synthèse des relations entre les variables des cinq stations et permet d'observer la superposition des quatre types de processus caractérisant la physico-chimie du cours d'eau. Chaque ligne du tableau 2 est représentée ici grâce à un code à deux caractères : le premier est la lettre correspondant à chaque station (A, B, C, D et E), et le second est le numéro de la variable (de 1 à 9).

L'analyse de l'interstructure, pour résumer, confronte une *ordination des dates d'échantillonnage* et une *typologie des processus spatiaux univariés*. Elle est une *étude de fonctionnement longitudinal*, vue à travers les différentes variables mesurées.

#### *Le compromis.*

L'ACP usuelle (qui renouvelle inutilement un centrage des colonnes) appliquée au tableau 3 permet d'obtenir la figure 4, sur laquelle sont superposées les stations et les variables du compromis.

Cette figure propose donc un résumé synthétique des structures inter-stationnelles, porteur de sa propre légende elle-même constituée par la typologie des variables artificielles créées par l'analyse. L'interprétation de ces structures et de cette typologie est ici immédiate, et l'intérêt de ce type de représentation n'apparaît en fait que dans les cas réels, où la complexité des phénomènes sous-jacents rendent effectivement nécessaire l'élaboration d'un résumé succinct et directement lisible. La comparaison avec la figure 5, qui représente les ACP réalisées à chaque date, montre bien qu'on obtient un résumé *à la fois* des relations inter-stationnelles *et* inter-variables sur les quatre dates.

On remarquera simplement, dans le but de faciliter l'interprétation des représentations de l'intrastructure, que l'axe 1 correspond à un gradient de pollution croissante (opposition des stations A et B par les variables conductivité,  $DBO_5$ ,

*Acta Oecologica/Oecologia Generalis*

est particulièrement riche du point de vue de l'interprétation de l'analyse. Elle permet en effet d'établir une typologie des processus spatio-temporels sous-jacents aux observations. Quatre types de processus nettement différenciés apparaissent.

— Le plus simple correspond au *gradient amont-aval de température et de débit*. Il est pris en compte presque exclusivement par le premier axe; seule la température dans la station *E* s'en écarte, et se reporte sur le second axe. Conformément à l'interprétation de ce second axe, on peut observer (fig. 1) que la température est en effet minimale dans la station *E* en février. Il s'agit là d'un phénomène sans intérêt réel, lié au fait que la température, mesurée seulement au degré près, est la même dans toutes les stations en février sauf dans la station *E*, où elle est inférieure de 1° à celle des autres stations.

— Le second type de processus est celui de pollution-restauration du cours d'eau, caractérisé par les cinq variables suivantes : conductivité,  $DBO_5$ , oxydabilité, teneur en ammoniacque et en orthophosphates. La station *A*, non polluée, s'oppose sur le premier axe à la station *B*, dont la pollution est maximale. Le processus de restauration se traduit ensuite par un retour des stations *C*, *D* et *E* vers les valeurs négatives de  $F1$ . La restauration est en général assurée au niveau de la station *E*, sauf pour la teneur en orthophosphates (le retour aux données du tableau 1 permet de vérifier cette particularité). L'axe 1 rend bien compte du processus de pollution-restauration, mais l'axe 2 souligne les caractéristiques particulières des stations *C*, *D* et *E*, et révèle ainsi l'importance de la pollution hivernale (février). Cette dernière est même plus importante dans la station *C* que dans la station *B* pour la conductivité et les teneurs en ammoniacque et orthophosphates (tableau 1). Le processus de restauration du cours d'eau ne peut donc pas être considéré comme un simple gradient amont-aval d'auto-épuration mais fait intervenir une interaction spatio-temporelle entre pollutions estivale et hivernale et position des stations le long du cours d'eau.

— Le troisième type de processus est celui de l'acidification puis de la remontée du pH de l'eau. Il diffère nettement du précédent car ce sont ici les stations *A* et *B* qui s'opposent sur l'axe 2. Le premier axe correspond au gradient amont-aval du pH annuel moyen, avec une exception pour la station *A*, qui occupe une position intermédiaire entre les stations *C* et *D*. L'opposition des stations *A* et *B* sur l'axe 2 provient du fait que le pH chute en février dans la station *A* (ce qui explique la faiblesse de sa moyenne annuelle), alors que dans la station *B*, il est maximal à cette même date. L'interprétation de cette interaction spatio-temporelle dans les variations du pH reste à discuter.

— Le processus d'oxydation de l'azote ammoniacal en nitrates ressemble au précédent par son opposition entre les stations *A* et *B*, mais en diffère par le fait que les stations *C*, *D* et *E* sont très proches les unes des autres. L'axe 1 est ici aussi très lié aux moyennes annuelles des mesures par station, et l'opposition entre les stations *A* et *B* est due au fait que la teneur en nitrates est minimale en février dans la station *A*, alors qu'elle est maximale à la même date dans la station *B* (tableau 1). L'interprétation de ces observations en termes de physico-chimie du milieu aquatique revient au spécialiste de ce domaine.

La figure 3, qui superpose les neuf graphiques de la figure 2 C fournit une synthèse des relations entre les variables des cinq stations, et permet d'observer la superposition des quatre processus qui viennent d'être décrits. On notera en particulier l'évolution conjointe du groupe de variables caractérisant la pollution

d'une station à l'autre. Ajoutons qu'il serait possible de poursuivre la discussion en étudiant les représentations des diverses variables regroupées par station, au lieu de regrouper les stations pour chaque variable (fig. 2 C).

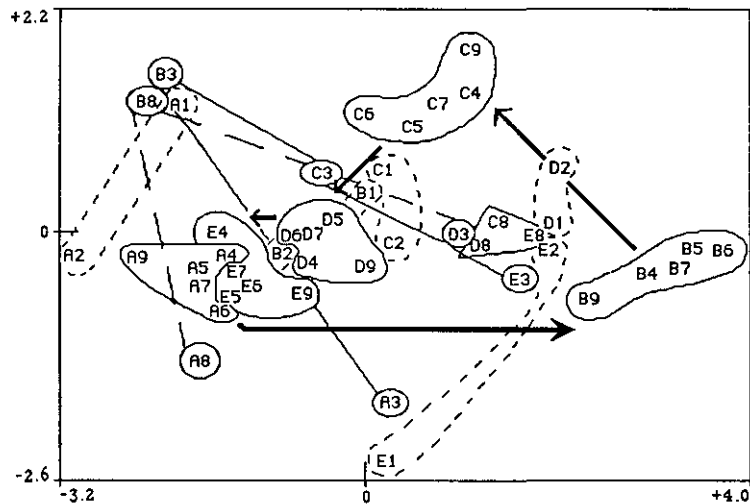


FIG. 3. — Superposition des neuf graphiques de la figure 2 C. Cette figure propose une synthèse des relations entre les variables des cinq stations et permet d'observer la superposition des quatre types de processus caractérisant la physico-chimie du cours d'eau. Chaque ligne du tableau 2 est représentée ici grâce à un code à deux caractères : le premier est la lettre correspondant à chaque station (A, B, C, D et E), et le second est le numéro de la variable (de 1 à 9).

L'analyse de l'interstructure, pour résumer, confronte une *ordination des dates d'échantillonnage* et une *typologie des processus spatiaux univariés*. Elle est une *étude de fonctionnement longitudinal*, vue à travers les différentes variables mesurées.

#### Le compromis.

L'ACP usuelle (qui renouvelle inutilement un centrage des colonnes) appliquée au tableau 3 permet d'obtenir la figure 4, sur laquelle sont superposées les stations et les variables du compromis.

Cette figure propose donc un résumé synthétique des structures inter-stationnelles, porteur de sa propre légende elle-même constituée par la typologie des variables artificielles créées par l'analyse. L'interprétation de ces structures et de cette typologie est ici immédiate, et l'intérêt de ce type de représentation n'apparaît en fait que dans les cas réels, où la complexité des phénomènes sous-jacents rendent effectivement nécessaire l'élaboration d'un résumé succinct et directement lisible. La comparaison avec la figure 5, qui représente les ACP réalisées à chaque date, montre bien qu'on obtient un résumé *à la fois* des relations inter-stationnelles *et* inter-variables sur les quatre dates.

On remarquera simplement, dans le but de faciliter l'interprétation des représentations de l'intrastructure, que l'axe 1 correspond à un gradient de pollution croissante (opposition des stations A et B par les variables conductivité,  $DBO_5$ ,

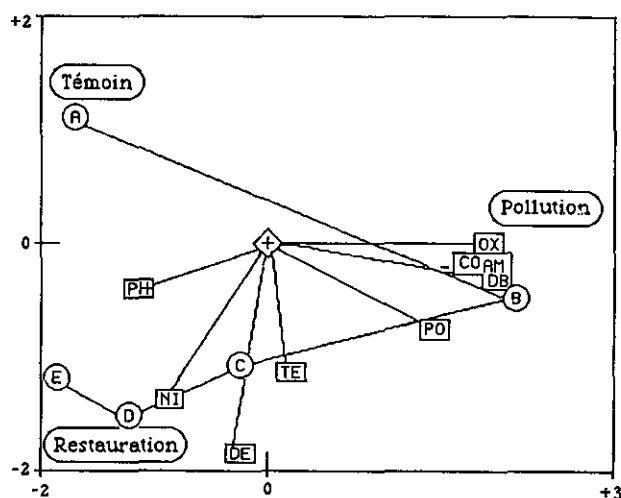


FIG. 4. — Représentation simultanée des lignes (cinq stations : A, B, C, D et E) et des colonnes (neuf variables représentées par deux lettres de leur nom) de l'analyse du compromis. Cette analyse propose une typologie moyenne des stations et des variables. La structure commune clairement exprimée ici est la succession témoin-pollution-restauration pour les stations, et l'opposition entre les variables caractérisant la pollution (ammoniaque, DBO<sub>5</sub>, conductivité oxydabilité) et celles du gradient amont-aval de température et débit. C'est dans cet espace de référence que l'étape de l'intrastructure va projeter les stations (et les variables) à chaque date, permettant ainsi de juger des écarts à ce modèle commun (fig. 6 et 8).

oxydabilité, ammoniaque, orthophosphates), alors que l'axe 2 correspond à un gradient amont-aval, principalement caractérisé par l'augmentation de température et de débit. Le pH est rejeté vers les valeurs négatives de F1 car il diminue lorsque la pollution augmente. Les variables pH, nitrates et orthophosphates prennent des valeurs négatives sur F2 car, bien que participant à la définition du niveau de pollution, elles présentent cependant aussi un gradient amont-aval net (dû à la restauration du cours d'eau), par opposition à l'oxydabilité par exemple, dont les valeurs ne sont réellement élevées qu'à la station B (en valeurs relatives puisque les données sont centrées et réduites par date). C'est sur le fond de ce double gradient pollution-restauration et amont-aval que va se situer l'interprétation des résultats de l'intrastructure.

L'analyse du compromis est donc une *typologie stations-variables moyenne*. Elle est assez proche des ACP à chaque date, mais permettra d'identifier les *écarts au modèle commun* pour les diverses dates.

#### L'intrastructure.

Il est possible de représenter les *stations aux diverses dates* de plusieurs manières : on peut relier chaque suite de quatre points correspondant aux quatre dates d'échantillonnage d'une station, ou bien relier entre eux les cinq points correspondant aux cinq stations échantillonnées à une date. Nous avons préféré cette seconde solution dans la figure 6, car elle privilégie l'aspect « typologie de processus spatiaux », que nous avons choisi *a priori*. De plus, cette présentation est

ainsi directement comparable aux quatre plans factoriels des *ACP* réalisées à chaque date (fig. 5). La représentation de l'intrastructure par rapport à la collection des

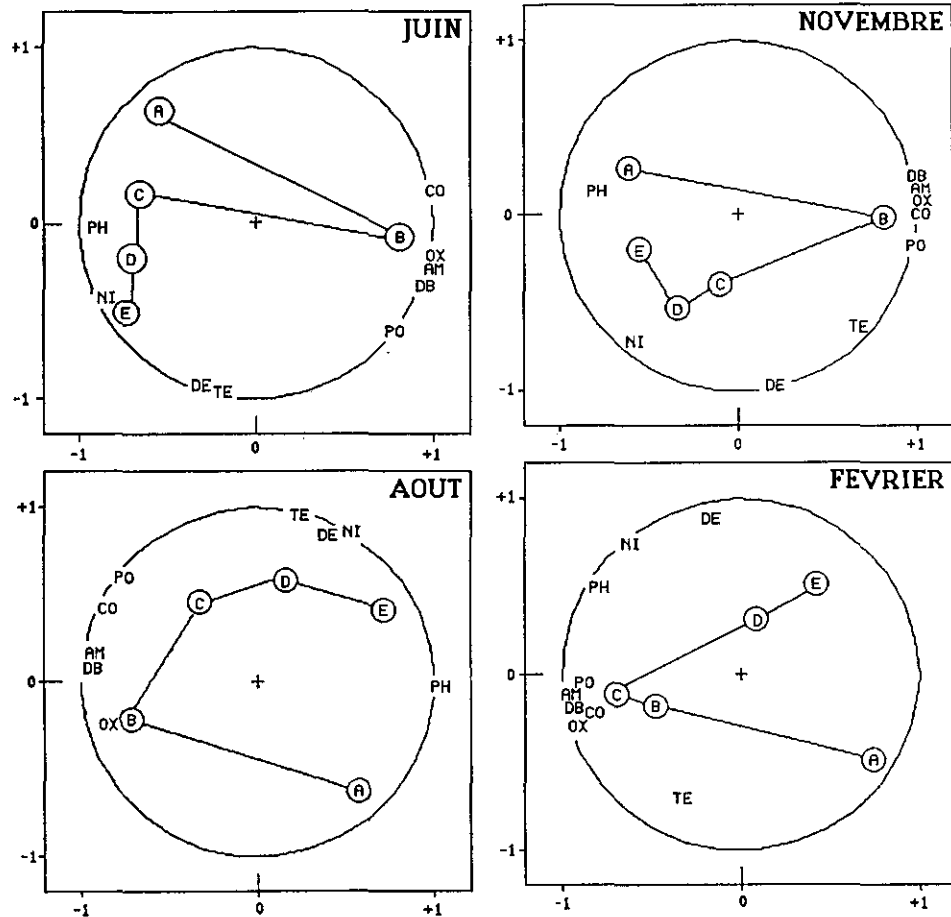


FIG. 5. — Comparaison des *ACP* faites à chaque date. Chacun des quatre graphiques correspond à la représentation simultanée des lignes (cinq stations : A, B, C, D et E) et du cercle des corrélations des neuf variables (représentés par deux lettres de leur nom) avec les deux facteurs. Les cinq stations ont été reliées entre elles afin de souligner les structures inter-stationnelles à chaque date, et de faciliter leur comparaison.

quatre plans factoriels d'*ACP* permet d'exprimer à une même échelle les relations inter-stationnelles aux quatre dates, tout en maximisant l'inertie du nuage global. On notera, en comparant les figures 5 et 6, que les trajectoires obtenues en reliant les cinq stations pour chaque date ont des formes très comparables entre l'intrastructure et les quatre *ACP* (à une inversion de signe près sur *F1* et *F2* en août et février). L'interprétation, commune aux deux figures, est la suivante :

— en juin, seule la station B est polluée, les stations C, D et E restent pratiquement indemnes. La capacité de régénération du cours d'eau semble donc suffi-

*Acta Oecologica|Oecologia Generalis*

sante pour assurer un retour à la normale dès la station *C* (voir tableau 1 et figure 1).

— En août, la pollution, toujours maximale à la station *B*, a nettement atteint les stations *C* et *D*, et le retour à la normale n'est réalisé qu'à la station *E*;

— En novembre, la pollution est moins importante qu'en août (sauf dans la station *B*), et la restauration n'est ici aussi complète qu'à la station *E*. Les stations *C* et *D* sont cependant déjà plus proches de la normale qu'en août.

— En février, la pollution est par contre maximale dans la station *C* (surtout pour la teneur en orthophosphates), et, même à la station *E*, la restauration n'est pas complète.

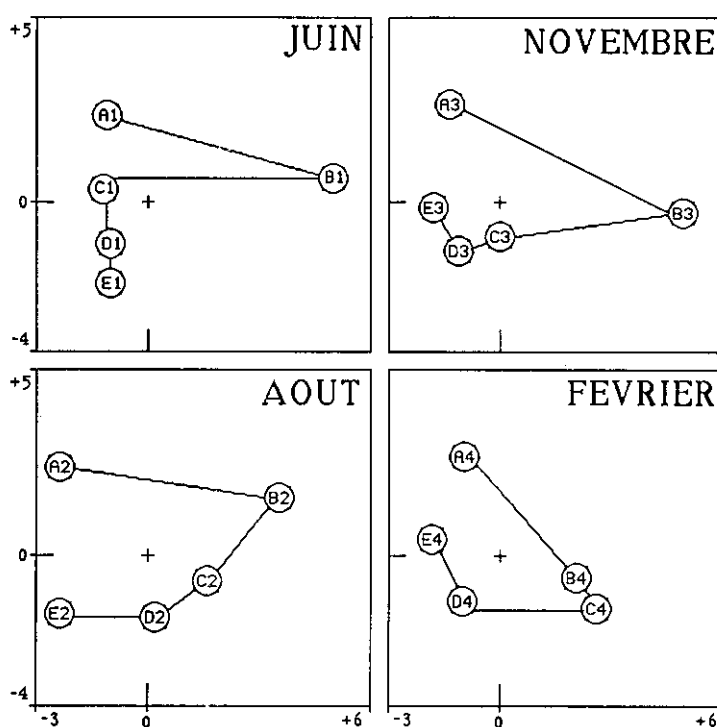


FIG. 6. — Représentation des cinq stations aux quatre dates (infrastructure). Les 20 lignes du tableau 1 (après centrage et réduction par date et par variable) sont projetées en éléments supplémentaires dans le plan  $F1 \times F2$  de l'analyse du compromis (fig. 4). Chaque ligne est représentée par un code à deux caractères dont le premier est la lettre de la station (*A*, *B*, *C*, *D* et *E*), et le second numéro de la date (juin = 1, août = 2, novembre = 3, février = 4). Nous avons relié entre elles les cinq stations à chaque date, ce qui privilégie l'aspect « typologie de processus spatiaux » de l'analyse, et permet de la comparer directement à la collection des plans factoriels des *ACP* réalisées à chaque date (fig. 5). Les trajectoires obtenues sont très comparables, mais l'avantage de l'infrastructure est qu'elle permet d'exprimer à une même échelle les relations inter-stationnelles à chaque date.

La typologie fine des processus spatiaux proposée par l'infrastructure met donc en évidence l'impact des deux périodes touristiques (août et février) sur la pollution



du cours d'eau, séparées par deux saisons durant lesquelles la régénération est relativement plus rapide (juin et novembre). La figure 7 permet de résumer ces conclusions et montre, sur le premier facteur, la déformation du gradient de pollution-restauration au cours du temps (fig. 7 A), et sur le second facteur l'évolution du gradient amont-aval de débit et température (fig. 7 B). Ce gradient, très fort en juin, est moins net en février : par exemple le débit à la station E en février est inférieur à celui de la station D.

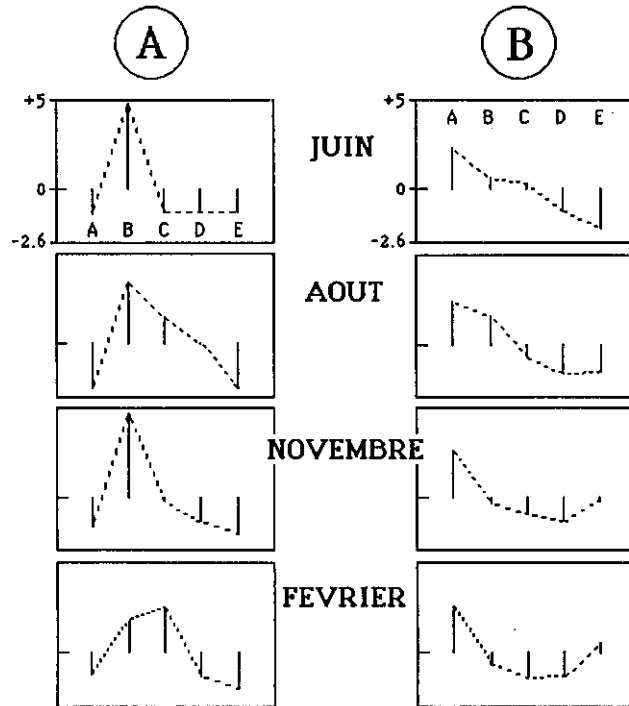


FIG. 7. — Coordonnées factorielles des stations à chaque date sur le premier (7 A) et le second (7 B) axe (intrastructure). La hauteur des bâtons est proportionnelle à la valeur de la coordonnée factorielle. Cette figure résume en A la déformation du gradient de pollution-restauration et en B l'évolution du gradient amont-aval de débit et température au cours du temps.

Pour représenter les *variables aux diverses dates*, on peut, comme précédemment, regrouper les points correspondant aux différentes variables pour une même date (quatre groupes de neuf points), ou au contraire regrouper les points correspondant aux quatre dates pour chaque variable (neuf groupes de quatre points). Cette seconde solution a été choisie dans la figure 8. Elle a ici l'avantage de mieux faire ressortir la participation de chaque variable à la définition des deux gradients qui induisent les typologies inter-stationnelles précédemment décrites.

Pour la température, l'axe 1 sépare période estivale et période hivernale, et l'axe 2 isole le mois de février pour lequel le gradient amont-aval d'augmentation de température n'est pas sensible (la température est même plus faible dans la station E).

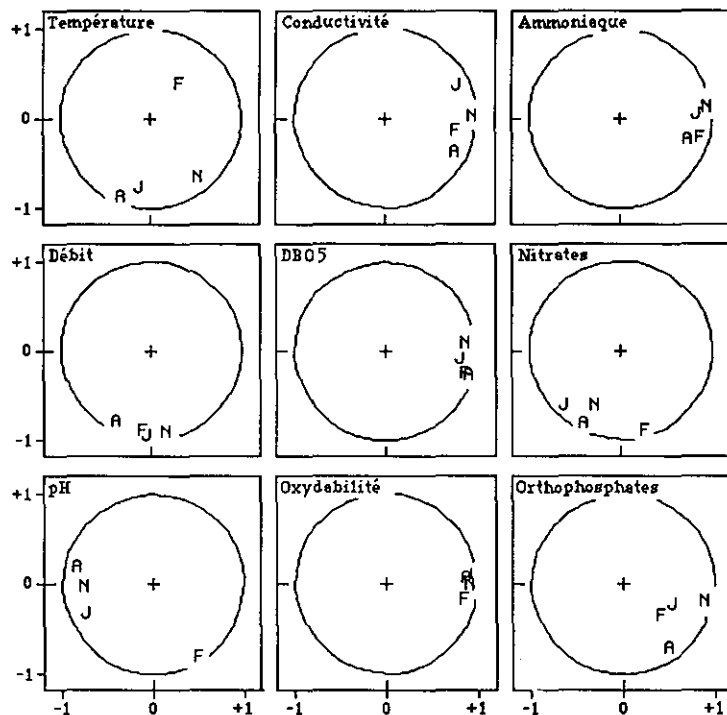


FIG. 8. — Représentation des neuf variables aux quatre dates (intrastructure). La projection des neuf variables mesurées aux quatre dates d'échantillonnage (après centrage et réduction par date et par variable) dans l'espace du compromis (fig. 4) permet de résumer l'évolution de chaque variable au cours du temps.

Le débit constitue la principale composante du gradient amont-aval, et les quatre dates d'échantillonnage y participent de façon équivalente : elles ont toutes les quatre des coordonnées très négatives sur F2.

Pour le pH, les valeurs négatives sur F1 en juin, août et novembre traduisent la *baisse* de pH due à la pollution. La position exceptionnelle du mois de février est due à un pH faible à cette date dans la station A.

Les cinq variables conductivité, DBO<sub>5</sub>, oxydabilité, ammoniaque et orthophosphates, caractéristiques du phénomène de pollution, sont bien, pour les quatre dates d'échantillonnage, orientées vers les fortes valeurs de F1. L'oxydabilité est la variable la moins influencée par le gradient amont-aval (le retour à la normale se fait dès la station C). A l'opposé, la conductivité est très marquée par ce gradient. En juin la conductivité présente en effet un maximum à la station B et un minimum à la station E, alors qu'en août, on observe un minimum à la station A et un maximum à la station C, ce qui explique l'opposition de ces deux dates sur l'axe 2 (fig. 8). La teneur en orthophosphates subit un gradient amont-aval net en août. A cette date en effet, la concentration augmente de la station A jusqu'à la station C. Par contre, en novembre, le maximum est à nouveau atteint à la station B, et la restauration incomplète ne permet pas d'obtenir un gradient amont-aval net. Ces deux dates s'opposent donc sur l'axe 2.

Enfin, la concentration en nitrates possède un profil particulier. En effet, du fait que cette concentration est liée à la teneur en oxygène, et donc inversement proportionnelle au degré de pollution, on observe sur la figure 8 que les mesures de juin, août et novembre prennent des valeurs négatives sur F1 (donc caractéristiques des stations non polluées). En février par contre, on observe une concentration en nitrates élevée dans la station *B* (par exemple supérieure à celle de la station *D*), ce qui se traduit sur la figure 8 par le fait que le point représentatif de cette date a une coordonnée positive sur F1. De plus, aux quatre dates d'échantillonnage, l'augmentation de la teneur en nitrates d'une station à l'autre suit le gradient amont-aval de débit et température, et les quatre points ont donc des coordonnées négatives sur F2.

L'intrastructure autorise donc une *typologie de processus spatiaux résumée sur les variables*, ou une *typologie de variables résumée sur leur évolution spatio-temporelle*.

### CONCLUSION

La décomposition de l'étude des structures internes des données en trois étapes successives et complémentaires (interstructure, compromis, intrastructure) est caractéristique des méthodes d'analyses multitableaux. Elle présente l'avantage décisif de permettre une approche progressive de la complexité de ces structures.

Par rapport aux méthodes « classiques » d'analyse multitableaux (méthode STATIS), l'AT présente de nombreux avantages, qui contribuent à faciliter à la fois sa mise en œuvre informatique (elle n'a nécessité, dans le cadre de la programmation de notre unité, l'écriture d'aucun programme particulier) et l'interprétation de ses résultats (en particulier pour l'interstructure et le compromis). Ces avantages ne sont cependant acquis qu'au prix d'une contrainte dont le poids est fonction des pratiques expérimentales de l'utilisateur : elle impose d'utiliser les *mêmes variables* mesurées sur les *mêmes individus* à plusieurs reprises. Cette contrainte explique peut-être le peu de développements qu'a connus la méthode depuis les travaux de JAFFRENOU (1978), mais ne semble pas prohibitive dans de nombreux domaines de l'écologie et de la biologie des populations.

Enfin, la simplicité de mise en œuvre des outils théoriques manipulés dans le cas particulier de l'AT (*i. e.* le fait que les objets caractéristiques associés à chaque date soient des tableaux) permet de transposer aux méthodes multitableaux les résultats classiques des diverses analyses à un tableau. Il devient ainsi possible de définir de façon simple une analyse discriminante inter-tableaux, une analyse des correspondances sur une suite de tables de contingence, ou l'analyse canonique de plusieurs tableaux de variables quantitatives.

C'est cependant au plan proprement écologique que les perspectives sont les plus nettes. Après avoir été le domaine des typologies d'état, l'analyse des écosystèmes est devenue celui des typologies de fonctionnement (BOURNAUD et AMOROS, 1984) : on y compare moins des enregistrements, de toute manière multivariés, que des évolutions (typologie de biotopes) ou des stratégies spatiales (typologie de périodes). Les notions d'*instabilité des biotopes* et de *stabilité des processus*, de *pérennité des mécanismes* et de *réversibilité des évolutions* développées par ROUX et coll. (1986) disent bien aux biométriciens que les objets de base sont, dans cette optique, les tableaux complets de mesure (dates-variables ou stations-variables). Comme outils au service de cette problématique, les méthodes multitableaux devraient jouer un rôle considérable.

*Acta Oecologica/Oecologia Generalis*

ANNEXE STATISTIQUE

La thèse de JAFFRENOU (1978) donne des indications théoriques reprises dans celle de GLAÇON (1981) que nous avons transcrites en termes de procédure utilisant un programme d'ACP. Les remarques qui suivent précisent le passage du général au particulier qui justifie les propositions ci-dessus.

Soient  $X_{k|k=1,\dots,t}$   $t$  tableaux portant sur les mêmes individus et les mêmes variables. Soient  $n$  l'effectif d'individus et  $p$  celui des variables. Les tableaux sont normés (chaque variable  $y$  a une moyenne nulle et une variance unité). On appelle  $E$  l'espace euclidien  $R^p$  muni de la métrique des poids uniformes, de matrice :

$$D_p = (1/p)I_p$$

dans la base canonique ( $e_i$ ). Soit  $E^*$  le dual de  $E$ . De même  $F$  est l'espace euclidien  $R^n$  muni de la métrique des poids de matrice :

$$D_n = (1/n)I_n$$

dans la base canonique ( $f_j$ ). L'espace des tableaux  $L(E^*, F)$ , canoniquement isomorphe à  $E^* \otimes F$ , est muni de la base des vecteurs  $e_i^* \otimes f_j$ , qui permet de récrire un tableau comme un vecteur ligne par ligne sur une seule colonne. Appelons  $Z$  le tableau à  $np$  lignes et  $t$  colonnes constitué par la juxtaposition des  $t$  tableaux  $X_{k|k=1,\dots,t}$  ainsi réécrits.  $L(E^*, F)$  est alors muni du produit scalaire  $Tr$ , noté  $\circ$ , qui vaut pour un couple de tableaux étudiés  $X_k$  et  $X_l$  :

$$(X_k \circ X_l) = Tr(D_p {}^t X_k D_n X_l)$$

où  $Tr$  désigne la trace de l'endomorphisme et  ${}^t X_k$  l'application transposée de  $X_k$ . Dans la base considérée la matrice de ce produit scalaire est  $D_p \otimes D_n$  (produit de Kronecker, voir GLAÇON, 1981, p. 8), soit tout simplement la matrice diagonale :

$$D_{np} = (1/np)I_{np}$$

Le produit scalaire des tableaux  $X_k$  et  $X_l$  vaut la moyenne des coefficients de corrélation entre les valeurs d'une même variable prises dans chacun des deux tableaux. L'ACP dans l'espace des tableaux (interstructure) muni de la pondération uniforme  $I_t$  et de la métrique  $Tr$  se trouve donc être par rapport à la base canonique de  $G = R_t$  et la base ci-dessus, exactement l'ACP de  $Z$ , qu'il est inutile de recentrer, par rapport aux métriques  $I_t$  et  $D_{np}$ , soit l'ACP ordinaire de  $Z$ .

En utilisant les théorèmes généraux du schéma de dualité (cf. par exemple CAILLIEZ, 1984, chapitre 3) on vérifie que cette ACP donne comme premier facteur un vecteur  $A' = (a_1, a_2, \dots, a_t)$  qui, sous la contrainte :

$$\sum a_k^2 = 1$$

maximise :

$$\| ZA \|_{D_p \otimes D_n}^2 = ZA \circ ZA = Tr(D_p (\sum a_k X_k) D_n (\sum a_k X_k))$$

soit, au facteur  $1/p$  près, l'inertie associée à l'ACP du tableau  $\sum a_k X_k$  (compromis), ou encore la somme des variances de ce tableau compromis.

Dans ce point de vue l'interstructure appelle d'elle-même l'ACP du compromis comme optimale. Ces remarques justifient la procédure employée dont l'accessibilité ne demande qu'un programme d'ACP classique.

RÉFÉRENCES

BOURNAUD M. & AMOROS C., 1984. — Des indicateurs biologiques aux descripteurs de fonctionnement : quelques exemples dans un système fluvial. *Bull. Ecol.*, 15, 1. 57-66.

- BOURNAUD M., MAUCET D. & CHAVANON G., 1983. — Structure des peuplements de macro-invertébrés benthiques en place et en dérive. *Acta Oecologica, Oecol. Gener.*, 4, 1, 39-57.
- CAILLIEZ F., 1984. — *Analyse des données*. Les Presses de l'Université de Montréal, Montréal, 103 p.
- DOLEDEC S. & CHESSEL D., 1987. — Rythmes saisonniers et composantes stationnelles en milieu aquatique. I. — Description d'un plan d'observation complet par projection de variables. *Acta Oecologica Oecol. Gener.*, 8, 3, 403-406.
- ESCOUFIER Y., 1980. — L'analyse conjointe de plusieurs matrices de données. In : JOLIVET et al. eds., *Biométrie et Temps*, 59-76.
- GLAÇON F., 1981. — *Analyse conjointe de plusieurs matrices de données. Comparaison de différentes méthodes*. Thèse de 3<sup>e</sup> cycle, Université Scientifique et Médicale de Grenoble, 73 p.
- JAFFRENOU P. A., 1978. — *Sur l'analyse des familles finies de variables vectorielles. Bases algébriques et application à la description statistique*. Thèse de 3<sup>e</sup> cycle, Université de Lyon 1, 97 p.
- L'HERMIER DES PLANTES H., 1976. — *Structuration des tableaux à trois indices de la statistique. Théorie et application d'une méthode d'analyse conjointe*. Thèse de 3<sup>e</sup> cycle, Université des Sciences et Techniques du Languedoc, Montpellier II.
- PAGES J. P., ESCOUFIER Y. & CAZES P., 1976. — Opérateurs et analyse des tableaux à plus de deux dimensions. *Cahiers du Bureau Universitaire Recherche Opérationnelle*, n° 25, Université Pierre-et-Marie-Curie, Paris, 62-89.
- PERNIN M. O., 1986. — *Contribution à la méthodologie d'analyse de données longitudinales. Exemple de la croissance chez l'être humain (Auxologie)*. Thèse de Doctorat, Université de Lyon 1, 264 p.
- PONTIER J. & PERNIN M. O., 1986. — Multivariate and longitudinal data on growing children solved by LONGI. *Proceedings of the Third International Symposium on Data Analysis*, Brussels. Plenum Press (sous presse).
- ROBERT P. & ESCOUFIER Y., 1976. — A unifying tool for linear multivariate statistical methods: the *RV* coefficient. *Appl. Statist.*, 25, 3, 257-265.
- ROUX A. L. et Coll., 1986. — Recherches interdisciplinaires sur les écosystèmes de la basse-plaine de l'Ain (France) : potentialités évolutives et gestion. *Documents de Cartographie Écologique*, XXIX, 166 p.