Published in ICCAT Coll. Vol. Sci. Pop. (1998) 47 : 289-300.

Multivariate analysis of relationships between tuna catches and fishing strategies. Application to the Venezuelan purse seiners in the Caribbean sea

Daniel Gaertner[1], Jean-Claude Gaertner[2],
Jesus Marcano[3] and Mauricio Pagavino[4]

Incorporation of the components of a fishery system into stock assessment studies, and especially harvesting practices and fishermen's behaviour at sea, requires the coupling of the traditional abundance data with several auxilliary data sets. These sets contain different types of information collected at different times and locations (e.g., fishing strategies, prices of the tuna species, environmental factors, etc.). To illustrate the usefulness of some multivariate analyses techniques in handling such diverse information, multiple factorial analysis, between-class and within-class principal component analyses, and principal component analysis with instrumental variables are applied to the Venezuelan purse seine tuna fishery operating in the southern Caribbean Sea. These examples show how multivariate analyses can be used (a) to compare the fishing strategies developed by two classes of vessels, relative to the detection of schools associated with floating objects, (b) to break out spatial and temporal components in the relationships between fish size category CPUEs, and finally (c) to take into account these fishing strategies when carrying out the statistical analysis of the CPUE data.

Dans les études sur l'évaluation de la ressource, la prise en considération des composantes d'un système pêche, et en particulier des modalités de capture de la ressource et du comportement des pêcheurs, implique le couplage de données traditionnelles sur l'abondance avec plusieurs ensembles de données auxilliaires. Ces tableaux regroupent des données collectées à divers endroits et à différents moments (par ex., stratégies de pêche, prix des espèces de thons, facteurs environnementaux). Pour illustrer l'utilité de quelques méthodes d'analyse multivariée dans le traitement de données aussi diverses, une analyse factorielle multiple, des analyses en composantes principales inter et intra-classes et une analyse en composante principales avec variables instrumentales ont été appliquées aux senneurs thoniers vénézuèliens qui travaillent dans la partie sud de la mer des Caraïbes. Ces exemples montrent comment les méthodes multivariées peuvent être utilisées pour (a) comparer les stratégies de pêche développées par deux classes de bateaux, par rapport à la détection de bancs associés à des objets flottants, (b) décomposer les liaisons entre les prises par unité d'effort (PUE) des différentes catégories de taille en un effet temporel et en un effet spatial, (c) tenir compte de ces stratégies de pêche dans l'analyse statistique des données sur les PUE.

En el estudio de las evaluaciones de recursos, el tener en cuenta los componentes de un sistema de pesca, en particular los que conciernen a las formas du utilizacion del recurso y el comportamiento de los pescadores, implica acoplar los datos tradicionales de abundancia con otros conjuntos de datos adicionales. Estos conjuntos contienen informacion diversa, recogida en diferentes estratos espacio-temporales (por ejemplo, estrategias de pesca, precio de las especies de tunidos, factores medioambientales, etc.). Para ilustrar la utilidad de alguna analisis factoriales, se ha aplicado un analisis factorial multiple, un analisis en componentes principales inter y intra clases y un analisis en componentes principales con variables instrumentales, a los cerqueros venezolanos que faenen en la zona sur del Mar Caribe. Estos ejemplos demuestran que los analisis factoriales pueden aplicarse para (a) comparar las estrategias de pesca desarrolladas por dos grupos de barcos, en relacion con la deteccion de cardumenes asociados con objetos flotantes, (b) desglossar los componentes espaciales y temporales en las relaciones entre las CPUEs de diferentes categorias comerciales, y finalmente, (c) tener en cuenta estas estrategias al hacer el analisis estadistico de los datos de CPUE.

[1] IRD (Institut de Recherche pour le Développement), HEA, BP 5045, 34032 Montpellier Cedex 1, France
[2] Ifremer, 1 rue Jean Vilar, 34200 Sète, France
[3] FONAIAP Apdo. 236 Cumana 6101 (Sucre) Venezuela
[4] IOV-UDO, Apdo 245, Cumana 6101, (Sucre) Venezuela

1 - INTRODUCTION

In past years, the analysis of population dynamics was limited to the study of fluctuations in abundance, and attempts to explain changes in stock size on the basis of environmental fluctuations (limited, in fact, to the physical environment) or the evolution of fishing gears. The results of stock assessments were limited in their application due to the fact that they were conducted and interpreted within a limited conceptual context, i. e., traditionally on a single-species basis, while not accounting for the fleet interactions, shifts between target species as a function of price, etc. The transfer of knowledge between scientists and administrators must often overcome differences in definition of scales and of units of observation. Indeed, fishery managers are confronted with local practical problems, concerning the utilisation of different fishing gears in a given area, while scientists provide global assessment of the overall production (which represents the sum of all harvests). In the same way the decision makers have to transpose general management decisions (such as may be recommended by scientists) into meaningful regulations (e.g., by reconverting effective effort into nominal effort) and, as a result, they need more information on the harvesting practices.

Nevertheless, it is recognised today that to be effective, fishery management must (a) be based on good regulations, and also (b) be able to anticipate the dynamic responses of fishermen to changes in stock size and in the management itself (Hilborn and Walters, 1992). Achievement of the objective of sustainable fisheries requires the scientific study of the totality of the components of fishery systems including the inter-relations between fishermen, and, for example, oceanographic, biological, technological and socio-economic factors (Charles, 1995). For these reasons, studies on the environment and on its utilisation by fishermen must concentrate more on anticipating and monitoring future developments than on finding an hypothetical equilibrium population level, which in reality is unachievable due to continuous fluctuation in the ecosystem components.

Analysis of fishing activities, in attempts to answer questions such as where, when, how and why, require the coupling of traditional catch rate data for the different species with auxiliary data, such as those on environmental factors, fishing tactics, prices by commercial categories, etc. All of these data are often broken down into spatial-temporal strata for the analysis. The visual analysis of such large data bases becomes impossible and it is necessary to resort to the use of multivariate statistical methods. These methods allow identification of underlying structures within the data and synthesis of the information that they contain. Thus, in studies of fishery biology, diversity of catches and variability of CPUEs can be related to environmental factors or to the behaviour of fishermen at sea.

Given that one of the main objectives of the ICCAT Tuna Symposium has been to explore methods used in stock assessment and to discuss new approaches, this paper gives an overview on the usefulness of some factorial analyses, specifically: multiple factorial analyses (MFA), principal component analysis with respect to instrumental variables (PCAIV), and between-class and within-class principal component analysis.

2 - ORIGIN OF DATA AND PRELIMINARY PROCESSING

Data were obtained from logbooks, introduced into the Venezuelan surface fishery in 1987 (Gaertner *et al.*, 1988). Purse seiners were classified into two categories according to their carrying capacity: small and medium seiners (PSS < = 600 t.); large seiners (PSL > 600 t.).

The aggregative behaviour of tunas with respect to surface floating objects is well known. In the southern Caribbean Sea, tunas are often associated with whales, whale-sharks (*Rhiniodon typus*) or sometime with flotsam, all of which are easier for fishermen to spot than the tuna themselves. In the absence of these, schools are generally found by detecting flocks of birds that are feeding on the same prey as the tunas. For this study, purse seine sets have been identified as beingon schools that were associated with whales, or with whale-sharks or with neither. The latter are described as “school” sets, or “non-associated” sets. Compared with other tropical areas, the percentage of sets made on flotsam-associated schools is very low (Gaertner *et al.*, 1996); hence, this type of fishing mode was not included in the present analysis.

Previous studies have shown that it is more appropriate to consider the association of tunas with floating objects on the basis of fish size rather than tuna species (Gaertner *et al.*, 1996). Therefore, three categories of fish size were chosen, based on the classification used commercially: small (< = 3 kg); medium (from > 3 kg to < = 15 kg); large (> 15 kg).

Data on catch per unit effort (CPUE) for 1989 to 1994, in metric tons caught per fishing days, were tabulated for these three size grouping of fish by class of purse seiner and by spatio-temporal strata of 1° x 1° square x quarter, and in order to homogenise the variance, CPUE was expressed in Log (X+1) form. The basic idea in the analysis was to associate this table of CPUE (strata x CPUE) with an auxiliary array (same strata x types of floating object). The total number of sets made on each type of floating object was divided by the effort exerted in the stratum, to provide a frequency index.

In accordance with the objectives of the study (reconstitution of a pooled CPUE index for both fleets, during a standard year, by 1° x 1° square x quarter, etc.) CPUE was standardized each year in the following manner:

- calculation of a relative index for each fleet and for each year;

α_{cfysq} = $CPUE_{cfysq}$ / aver ($CPUE_{cfy..}$), with

aver ($CPUE_{cfy..}$) = $\Sigma^s \Sigma^q C_{cfy..} / \Sigma^s \Sigma^q E_{cfy..}$

For c = commercial category (1, 2, 3)

f = fleet (1, 2)

y = year (1989, ..., 1994)

s = 1° x 1° square (1, 2,..., n)

q = quarter (1, 2, 3, 4) and

$\Sigma^s \Sigma^q C_{cfy..}$ = sum of catches for all squares and quarters for c, f, y,

$\Sigma^s \Sigma^q E_{cfy..}$ = sum of efforts for all squares and quarters for c, f, y.

- the pooled relative index corresponds to the average of relative indices of the two fleets, weighted by their respective effort:

$\mu_{c.ysq} = \Sigma^f (w_{cfysq} \times \infty_{cfysq})$,

with $w_{cfysq} = E_{cfysq} / [\Sigma^f E_{cfysq}]$

The standard year CPUE was obtained by averaging the pooled indices within each strata. An abundance index expressed in a more usual unit was obtained by multiplying this standardised index by the average 1989-1994 CPUE for large purse seiners.

The central procedure of calculation in linear factorial analyses is based on the diagonalisation of a square matrix obtained from a statistical triplet. This triplet is composed by a transformed matrix and of one row weight and one column weight matrices (Fig. 1). Depending on the multivariate method used, the nature of the transformation and the two weight matrices are different.

The multivariate data analyses were made with the ADE-4 software (Thioulouse *et al.*, 1995). The ADE-4 package is freely available on the Internet by anonymous FTP to biom3.univ-lyon1.fr (into the directory /pub/mac/ADE) or by http://biomserv.univ-lyon1.fr/ADE-4.html (or ADE-4F.html for the French version). ADE-4 is now available for Windows 95.

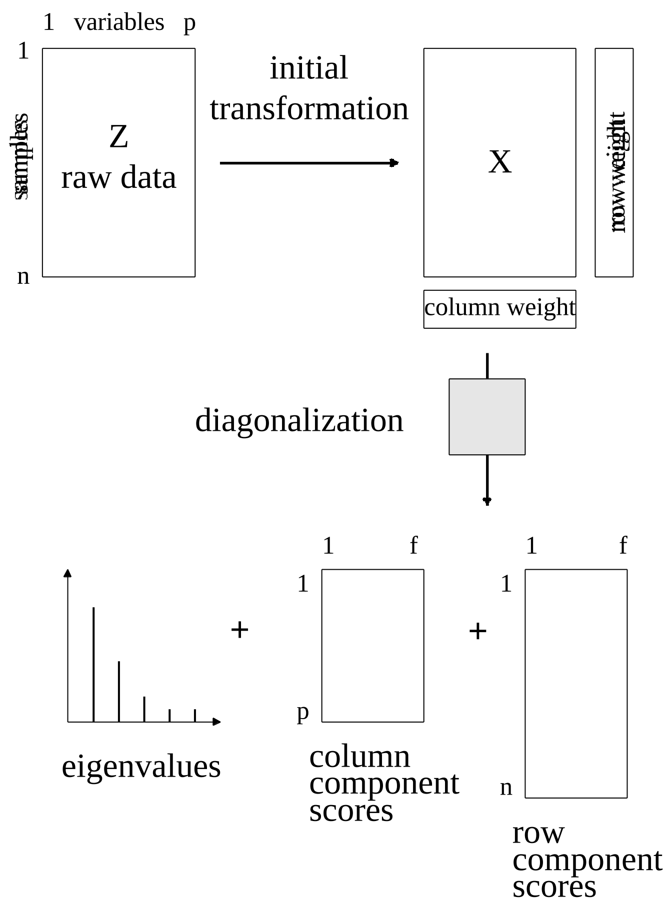


Figure 1 - Core procedure for calculations in linear multivariate data analysis (after Dolédec and Chessel, 1991, simplified).

3 - MULTIPLE FACTORIAL ANALYSIS

3.1 Basic concepts of the method

Multiple Factorial Analysis (MFA) (Escofier and Pages, 1994) belongs in the general framework of multiple tables analyses, such as the generalised procustes analysis (Gower, 1975), the Tuckals3 method (Kroonenberg and Brouwer, 1985) or the Statis method (Lavit *et al.*, 1994). The aim of these procedures is to make evident (a) the underlying common structure present on several tables, and (b) the variability of each of them as compared to an "average solution". The objective of MFA is to describe data in which the same observations (rows) are described by different tables of variables. These tables can be built up of quantitative as well as categorical groups of variables (see Escofier and Pages, 1990, 1994; Lebart *et al.*, 1995). The implementation of MFA can be broken into two distinct phases (Fig. 2).

- First, a principal component analysis (PCA) is carried out for each of the K tables (K separate PCAs). The inverse of the first eigenvalue (i.e., the variance explained by the first component) obtained in each separate PCA, serves to weight each table's variables during the second phase. This weighing operation (a) balances the influence of each table in the global PCA, and (b) maintains the internal structure of each table (because the weighing factor is the same for each variable into a given table).
- During the second step, a global PCA is completed for all the weighted variables of the K tables. Then a typology of the variables (or a typology of the observations) can be made.

The MFA method is relatively similar to canonical analysis since it tries to identify common features across K groups of variables (the K tables). An application of canonical analysis to the comparison of the fishing strategies of two fleets of longliners was given by Ehrhardt (1994). However, canonical analysis is not easy to interpret and its main objective is limited to the search for the best correlations between groups of principal components.

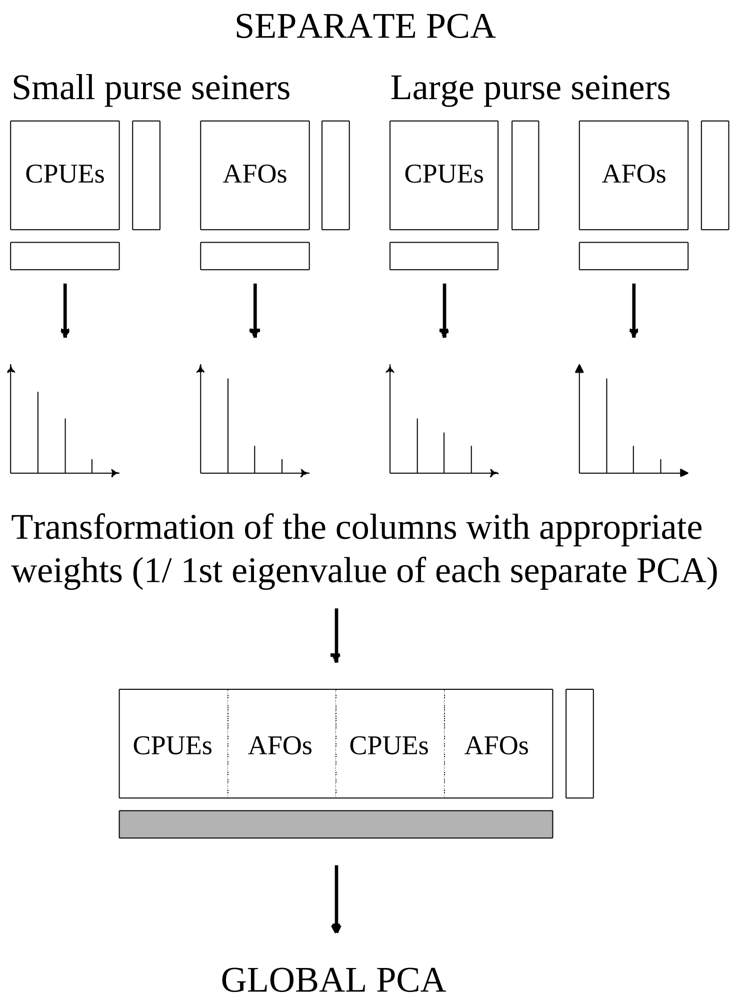


Figure 2 - Schematic representation of the two distinct phases of the MFA.

Published in J.S. Beckett (Ed.) Proc. of the ICCAT Tuna symp. (1998) Coll. Vol. Sci. Pap. ICCAT 50 (2) : 628-636

3.2 Application

MFA was used to compare the performance of the two classes of Venezuelan purse seiners, relative to their respective CPUEs of each of the three size categories of tuna and to their strategy with respect to associated floating objects (AFO). A minimum threshold of five fishing days in a strata by each of the two fleets was chosen so as to eliminate strata with low effort, or strata visited solely by one class of purse seiner. Strata were also eliminated when information on floating objects was incomplete (catches but no observations as to the presence or not of AFO). The aim of this study was to compare tactics of the two fleets across the same spatio-temporal strata, from 1989 to 1994.

This analysis examined four tables: two of CPUE (variables: CPUE1, CPUE2, CPUE3; for each fleet) and two of fishing modes (variables: school sets, whale sets, whale-shark sets; for each fleet). It is obvious (Table 1) that it is difficult to make a simultaneous analysis of these four tables on a visual basis.

Correlation analyses of the projections of the global space of observations and each of the four subspaces generated by each table indicate common features across these tables. Correlations (Table 2) are roughly comparable between (a) the two tables of AFOs and the table of CPUEs for the large purse seiners, on the first factor in the MFA, and (b) tables both of CPUE on the second factor. With the exception of the table of CPUEs for small purse seiners, the third factor is correlated poorly and is thus of less interest.

The decomposition of the three first factors for each group of variables confirms that the first component is linked to both groups of AFOs and to the CPUEs for large purse seiners (Table 3). This latter group contributes to a similar extent to the inertia of the second axis. The CPUEs for small purse seiners are linked strongly to the third factor. It is necessary to keep in mind that the maximum inertia of each group of variables was limited to 1 by the weighting procedure.

The representation of the groups of variables (Fig. 3) (presented here separately for each fleet to facilitate the interpretation) shows a difference between the CPUE of the larger fish size category (respectively U3S for small vessels and U3L for large purse seiners) and the CPUEs of the two other fish size categories. This suggests that the spatio-temporal strata where small tunas are abundant are not the same as for large tunas. The same conclusion can be drawn for the AFO variables: school sets (respectively SCHS and SCHL) are different in comparison to the other two

CATCH PER UNIT EFFORT						ASSOCIATED FLOATING OBJECTS					
SMALL PS			LARGE PS			SMALL PS			LARGE PS		
U1P	U2P	U3P	U1G	U2G	U3G	SCH.	WHA.	W-SHA.	SCH.	WHA.	W-SHA.
1.95	3.14	1.10	1.61	0.00	1.39	4	1	14	0	0	0
0.00	2.20	2.20	0.00	2.71	3.30	1	2	11	0	3	11
0.69	1.61	2.08	0.00	2.30	2.56	16	8	8	0	14	7
3.56	3.50	2.83	0.00	1.79	0.69	3	12	1	0	1	1
1.39	1.79	2.89	0.00	1.10	2.71	13	7	24	0	6	8
2.08	1.39	1.95	0.00	1.10	2.08	5	2	2	0	1	1
1.95	1.79	2.20	0.00	2.89	2.56	10	9	6	0	11	7
1.79	2.64	2.77	0.00	2.64	2.71	14	7	4	0	2	1
1.61	2.30	1.39	0.00	2.71	1.61	6	3	4	1	1	6
2.08	1.79	0.00	0.69	0.69	0.00	2	0	2	0	0	1
1.10	1.10	1.61	0.00	2.30	1.39	0	0	2	3	2	9
1.10	2.20	1.10	0.00	3.58	2.56	4	2	5	2	3	0
1.61	2.30	0.69	0.00	1.10	0.00	12	9	4	0	1	0
0.69	2.08	0.69	0.00	3.00	0.00	1	4	1	0	5	0
1.79	1.95	1.39	1.79	2.89	0.69	8	1	19	1	5	11
" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "
" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "
" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "
" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "	" " "

Table 1 - Sample data bases used in the tuna fishery studies. The four tables used for the Multiple Factorial Analysis are: the Catch Per Unit of Effort tables (three fish size categories (U1, U2, U3)) and the fishing mode tables (school sets (SCH), whale sets (WHA) and whale-shark(W-SHA) sets) for the 2 classes of Venezuelan purse seiners. Data were compiled by spatio-temporal strata (1° x 1° square x quarter), from 1989 to 1994.

Tables	Factor 1	Factor 2	Factor 3
CPUEs Small PS	0.59	0.72	0.81
CPUEs Large PS	0.74	0.83	0.25
A.F.O.s Small PS	0.79	0.48	0.33
A.F.O.s Large PS	0.83	0.56	0.41

Table 2 - Correlation coefficients between total space of observations and subspace of each table, for the main factors of the MFA.

Tables	Factor 1	Factor 2	Factor 3
CPUEs Small PS	0.32	0.35	0.64
CPUEs Large PS	0.54	0.56	0.06
A.F.O.s Small PS	0.62	0.10	0.08
A.F.O.s Large PS	0.66	0.26	0.17
Eigenvalues	2.14	1.27	0.95
% Cumulated Inertia	0.26	0.41	0.53

Table 3 - Contribution of each table to the total inertia of the main factors of the MFA.

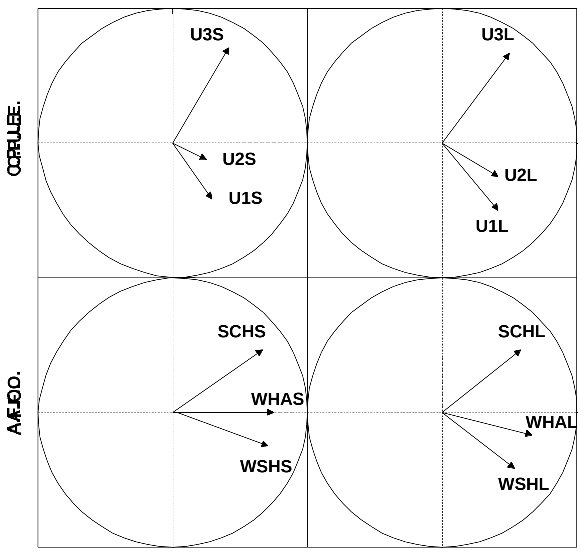


Figure 3 - Representation of variables in the first two components of MFA. Partial graphs of CPUE categories are in the upper part and partial graphs of associated floating objects are in the lower part, for small and medium PS (left) and large PS (right).

fishing modes. These figures indicate that large fish are caught in school sets, whereas small and medium tunas are associated more with whale-sharks and with whales. The next step with this method would be to analyse the projection of the observations (the strata), because in fishery studies this projection can be useful to identify spatio-temporal effort clusters. Given, however, that some locations were visited during only a single quarter of the year, it was difficult to find a clear typology of these strata and hence such analyses are not presented in this study.

In conclusion, within the same spatio-temporal strata, the two classes of purse seiners behave comparably with respect to AFOs. However the CPUEs for the two classes of vessels (fishing power) are less similar (Table 2 and 3). The CPUEs for small purse seiners are linked less to the various fishing modes than are the CPUEs for large purse seiners. This may be explained by the fact that small and medium purse seiners more frequently seek the help of baitboats to hold tuna schools stationary (Gaertner *et al.*, 1996). This additional factor could be introduced into future analyses (one of the advantages of MFA is the possibility to couple such different data sets).

4 - BETWEEN-CLASS AND WITHIN-CLASS FACTORIAL ANALYSIS

4.1 Basic concepts of the method

Bearing in mind that the fishery systems are dynamic, their study involves the description of a three-dimensional data table (species abundance x times x areas), i.e., CPUEs are sampled in different locations on different occasions. The aim of between-class and within-class PCA is to use location and time as two qualitative variables In other words, when analysing a spatio-temporal influence on the data, the choice is whether to focus on a given effect (e.g., space or time) or to eliminate this effect. In practice, a global PCA is calculated initially for the whole data, then successive PCAs on a between-date matrix, on a within-date matrix, on a between-location matrix and on a within-location matrix. As in an ANOVA, total variability can be broken down into spatial and temporal effects (Dolédec and Chessel, 1987, 1989). This is made possible by comparing the first eigenvalue (the variance explained by the first principal component) obtained during each analysis (Fig. 4a).

In order to consider this, let Z be a table with n rows (areas x dates) and q columns (CPUEs). After the initial transformation (e.g. log or centering) from Z to X, b groups of rows are defined (each row is allocated to a group, for instance, by dates). Two new arrays are then created: (a) the "X+" table (a between-date matrix with b rows x q columns) in which data are cumulated by groups, and (b) the "X-" table (a within-date matrix with n rows and q columns) in which data are centred in a manner to eliminate the temporal effect. Between-class analysis is the PCA of the "X+" table, namely the comparison of the mean distributions of areas by dates (Fig.4b). Within-class analysis corresponds to PCA of the "X-" table (Fig.

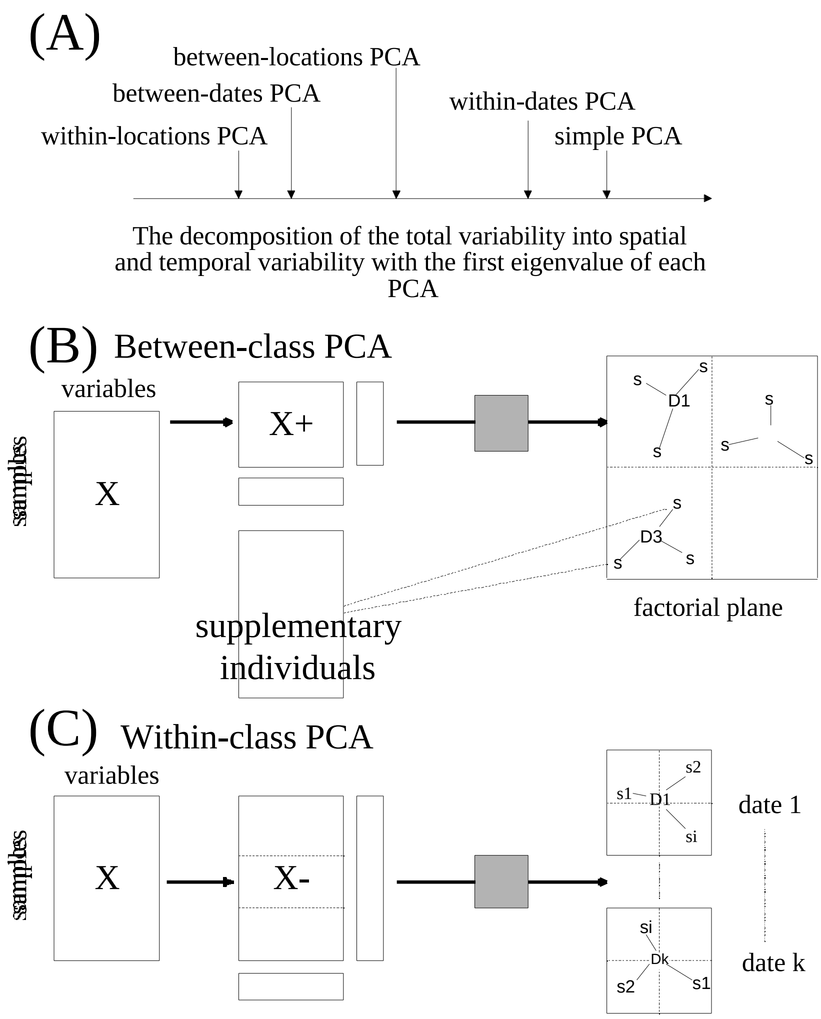


Figure 4 - (a) The decomposition of the total variability into spatial and temporal variability by representation of the first eigenvalue of each between-class and within-class PCAs. (b) Core procedure of the between-class PCA (i.e., for instance, focusing on a temporal effect, in the array noted X+, data are cumulated by dates). (c) Core procedure of the within-class PCA (i.e., for instance removing a temporal effect, in the array noted X-lay the residuals by dates); after Dolédec and Chessel, 1991.

4C), namely the analysis of the residuals. For more technical details on these methods, see, e.g., Dolédec and Chessel (1989, 1991). An a priori group effect can also be taken into account using discriminant analysis. Nevertheless, discriminant analysis use the Mahalanobis distance whereas between-class analysis use the Euclidian distance. It is important to point out that discriminant analysis performs best with a limited number of variables and is not well suited for large arrays (Dolédec and Chessel, 1991).

4.2 Application

In this example the spatio-temporal variability of the fish-size-specific CPUEs pooled for both fleets during a standard year (as explained in the section on preliminary processing) is analyzed (Table 4). The data base corresponds therefore to an array of three categories of standardized CPUEs (columns) by 1° square - quarter (rows). As mentioned, given the small size of these strata (1° lat. x 1° long. square), some locations were not visited every quarter. Hence, only the study of the temporal structure is presented in this document.

The between-class PCA focusses on the temporal effect by comparing the cumulative profiles of 1° squares by quarters. This analysis looks for maximum dispersions across the set of centres of gravity of the subspaces defined by dates (quarters). The centres of gravity are scattered along the first factor that explains 90% of the total inertia (Fig. 5). The second

Strata			Pooled CPUE			Floating objects		
Quarter	Lat.	Lon.	U1	U2	U3	SCH.	WHA.	W-SHA.
1	9	56	2.814	0.000	0.727	0.000	0.143	0.000
1	10	67	1.788	2.744	2.320	0.352	0.198	0.788
1	10	68	1.637	2.396	0.000	0.200	0.000	0.800
1	11	64	1.582	1.884	2.063	0.081	0.086	0.822
1	11	65	1.171	1.863	1.711	0.118	0.113	0.747
1	11	66	1.096	1.839	1.279	0.151	0.070	0.444
1	11	67	1.862	2.494	1.678	0.200	0.119	0.552
1	11	68	1.282	2.177	1.410	0.272	0.089	0.362
1	11	69	1.207	2.297	1.512	0.212	0.120	0.597
1	12	64	1.744	2.244	1.434	0.667	0.500	0.333
1	12	65	1.514	0.338	0.088	0.215	0.000	0.357
1	12	66	1.057	1.846	0.000	0.238	0.111	0.667
1	12	67	0.939	0.894	1.111	0.239	0.092	0.136
1	12	68	1.771	2.255	2.120	0.330	0.273	0.084
1	12	69	2.075	2.723	2.481	0.354	0.111	0.580
1	12	70	2.702	3.460	2.917	0.429	1.143	0.286
1	13	67	0.000	1.488	1.187	0.000	0.200	0.000
2	8	54	2.784	0.000	2.670	0.167	0.167	0.000
2	8	55	0.000	0.000	0.000	0.000	0.400	0.000
2	8	56	0.881	0.000	0.000	0.200	0.000	0.000
2	9	56	1.394	0.000	0.000	0.286	0.000	0.000
2	10	59	1.851	0.676	1.162	0.000	0.333	0.000
"	"	"	" " "	" " "	" " "	" " "	" " "	" " "
"	"	"	" " "	" " "	" " "	" " "	" " "	" " "

Table 4 - Sample data bases used in the between-class and within-class PCAanalyses (only CPUEs table), and in PCAIV (CPUEs as the dependant variables, and associated floating objects as the independant variables). Data were pooled for both fleets and by spatio-temporal strata (1° x 1° x quarter for an average "standard" year).

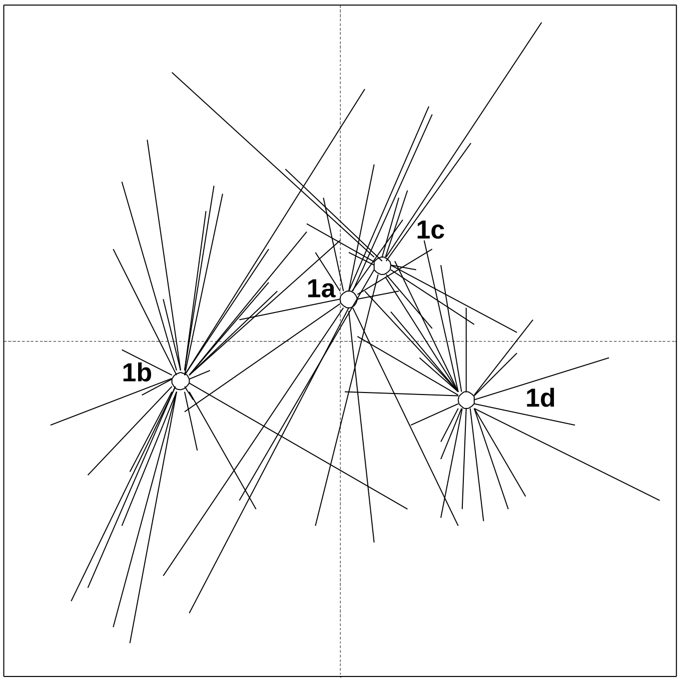


Figure 5 - Projection of the gravity centers of the four quarters onto the factorial plane (1-2) of the between-class PCA. Plots of individual strata are linked to their own gravity center (1a = 1 st quarter, 1b = 2 nd quarter, 1c = 3 rd quarter, 1d = 4 th quarter).

and the fourth quarter show the greatest separation. In order to evaluate the variability of each sampling unit around each centre of gravity, the initial data are plotted as supplementary individuals. To aid interpretation, each strata was linked with its mean date. This figure suggests that the spatial variability is lower during the fourth quarter than during the other quarters. Consequently, it is likely that the purse seining strategy does not depend on a spatial component during this quarter. In contrast, the heterogeneity between the strata during the other quarters of the year suggests that the decision to visit a strata (thus the distribution of the fishing effort) could be linked with its relative productivity.

Projection of variables indicates that CPUE2 is linked with the first axis, in the opposite direction to that of CPUE3 (Fig. 6). It appears that CPUE1 is not well projected into the plane formed by axes 1 and 2. The strong heterogeneity between quarters (confirmed by a permutation test) can be explained by the variation in the abundance of the target fish groups contributing to CPUE2; large skipjack (*Katsuwonus pelamis*), blackfin (*Thunnus atlanticus*) and intermediate yellowfin (*T. albacares*) and in a minor way, of CPUE3, i. e., mainly large yellowfin.

A within-class analysis can be made to remove the seasonal (quarterly) effect. The objective is to determine the axes of common direction across the subspaces of observations. In order to enable simultaneous study of the spatial typologies all centres of groups are plotted at the origin of the factorial maps and the strata are scattered with the maximal variance around the origin. The projection of the four subspaces in the within-class analysis (one per quarter) shows a geographical pattern. During the first two

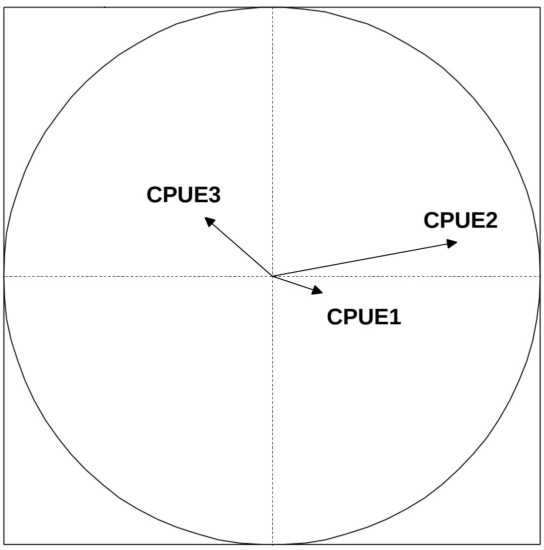
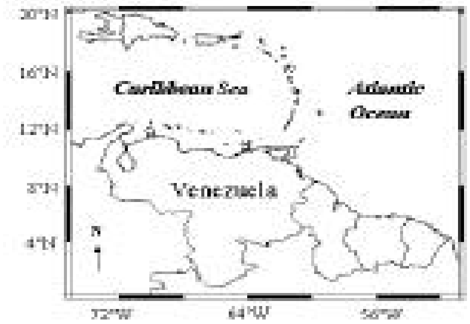


Figure 6 - Projections of the size categories CPUEs onto the factorial plane (1-2) of the between-class PCA.

quarters of the year strata from East of Venezuela (offshore of the Guyanas) are projected to the left and above. Western Venezuelan strata, close to the islands of Curaçao and Aruba, are projected to the right (Fig.7; see the two upper figures). This gradient is not as evident for the third and fourth quarters, probably because there is no information for the eastern strata.

It is interesting to note that this analysis does not show a clear inshore-offshore gradient in fish abundance as would generally be expected for juveniles (small tunas are more coastal than adults). With respect to projections of the variables, it will be noted that variables CPUE2 and CPUE3 appear to be linked positively, whereas they are the opposite in the between-class analysis (Fig. 8). This point indicates clearly that taking into account, or by contrast eliminating, a temporal (or a spatial) effect can change the perception of the relationship between the variables. In the present analysis, it is suggested that CPUE2 and CPUE3 show different seasonal patterns (especially during the second and fourth quarters) but have similar geographical structures.



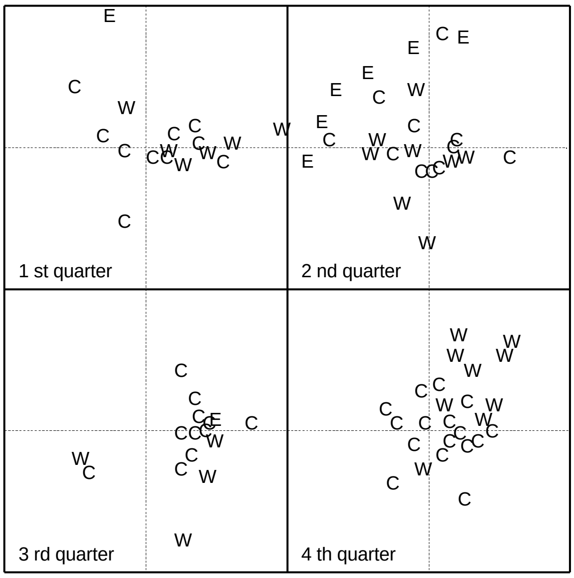


Figure 7 - Projections of the strata onto the factorial plane (1-2) of the within-class PCA. Eastern strata (E)= long. < 60 W, Western strata (W)= long. > 67 W, Central strata (C)= intermediate locations.

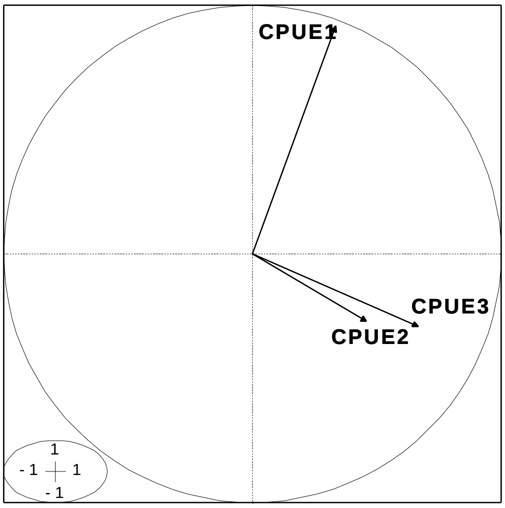


Figure 8 - Projections of the size categories CPUEs onto the factorial plane (1-2) of the within-class PCA.

5 - THE PRINCIPAL COMPONENT ANALYSIS WITH RESPECT TO INSTRUMENTAL VARIABLES

5.1 Basic concepts of the method

The aim of factorial analysis with respect to instrumental variables is to consider the influence of a set of auxiliary variables on the structure of a table of information. Depending upon the objectives of the study and the nature of data, this method can be applied to PCA (PCAIV used by Sabatier, 1984, 1987; Lebreton *et al.*, 1991; Prodon and Lebreton, 1994; Pech and Laloë, 1997) or to correspondence analysis (canonical correspondence analysis, or CCA, used by Ter Braak, 1986; 1987, 1988; Chessel *et al.*, 1987; Lebreton *et al.*, 1988 a, 1988 b). The former method was applied to tuna-dolphin associations in the East Pacific by Fiedler and Reilly (1993) and by Reilly and Fiedler (1993).

The study of relationships between the variables of interest (e.g., the fish size specific CPUEs) and an auxiliary table (incorporating different kinds of information such as associated floating objects, environmental factors, prices of tunas) can be approached in various manners: (a) by doing a simple PCA of the first table and then attempting to link factorial coordinates of observations with the variables of the second table, or (b) by directly coupling both tables in a manner that constrains the factorial coordinates to become a linear combination of this second table. The PCAIV, known equally as "PCA under linear constraints" (Lebreton *et al.*, 1991; Prodon and Lebreton, 1994), uses this last procedure.

Furthermore, combining linear regression and ordination gives PCAIV a certain robustness, especially with respect to the Guttman "horse-shoe" effect, frequently found in many traditional PCAs. Indeed, this parabolic relationship between the first two axes can be an artefact as well as due to a lack of linearity between variables (Prodon and Lebreton, op. cit.).

5.2 Application

The working matrix remains that of pooled CPUEs aggregated on an average year, but here AFOs constitute auxiliary variables (i.e., the fishing modes are considered as reflecting the strategies adopted by the tuna fishermen; Table 4). Results of multiple regression steps showed that instrumental variables explai-

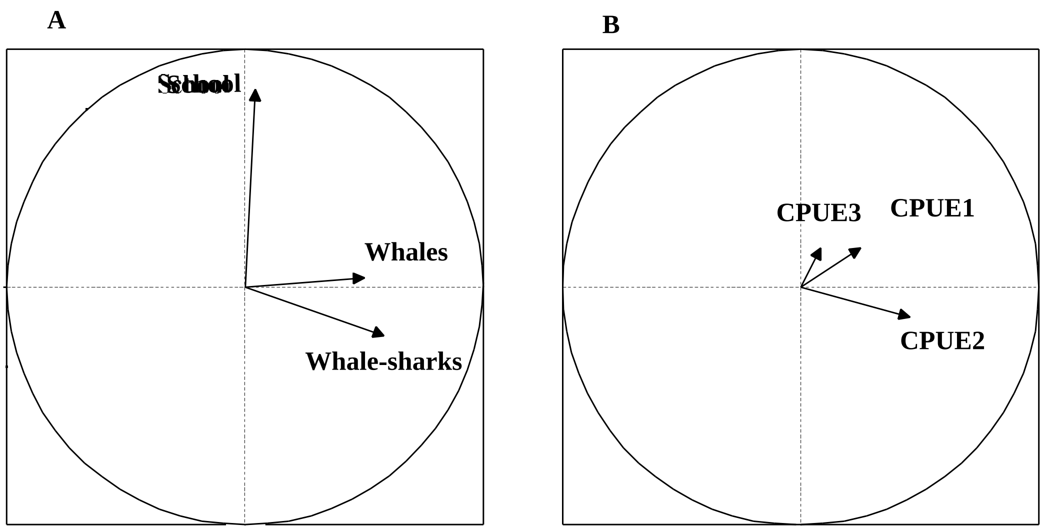


Figure 9 - (A) Projection of the size categories CPUEs onto the factorial plane (1-2) of the PCAIV and (B) projection of instrumental variables (associated floating objects), as supplementary variables.

ned 25% of the variability of the CPUEs (significance was confirmed by a test of permutation). Consequently, PCAIV was applied to this data set (in other words the PCA is made on the fitted CPUEs).

The major part of the information is given by the first two factors of the PCAIV (respectively 83 and 15% of the total inertia). The second size specific CPUE contributes most of the inertia of the first axis (lower part of fig. 9). This result reflects the favourable adjustment obtained for this size specific CPUE (nearly 40% of its variability was explained by the fishing modes). The projection of the instrumental variables in the first graph of the PCAIV shows that CPUE2 is associated with sets made on whales (upper part of fig. 9). As mentioned, the fish in this size category comprise several species of tunas: large skipjack, large blackfin and medium size yellowfin. Although their projections are close, it appears that whalesharks and whales do not attract exactly the same sizes of tunas. Finally, sets made on non-associated schools (namely, school sets) remain isolated, and could be linked to the CPUE3. A similar hypothesis was drawn from the MFA results.

6 - CONCLUSION

This paper provides an overview of some modern multivariate methods available to fishery biologist in order to highlight the links between the resources and the features of the fishery system. Given the complexity of the information available to the study tuna fisheries, the synthesis or comparison between data tables is difficult without the use of an appropriate structuring tool. The presentation of several examples shows that various types of factorial analysis can be useful in exploratory analysis and in fishery monitoring. This is especially true for (a) identification of spatio-temporal effort clusters (typology of spatio-temporal strata), (b) taking into account changes in target species and/or in the resulting distribution of fishing effort, and (c) determining which component of a fishery system responds the fastest to a regulation. These mutivariate analyses are not claimed to be substitutes for stock assessment models, but they can help to better understand the differences in the harvesting methods and practices.

The other multivariate methods, such as discriminate analysis, canonical analysis, etc., which have been used in the past, are less robust in their application or are difficult to interpret. The major interest of the methods presented in this paper are the following:

- MFA and other multi-table analysis are new and promising ways to analyse several tables simultaneously. The same observations (spatio-temporal strata as well as fishing units, fleets, etc.) can be described by various groups of variables (CPUEs, environmental factors, tactics, etc.). These methods could be used also to compare the abundance into fishing areas between dates (e.g., for designing impact assessment of a given regulation, a change in fishing gear, etc.). In this case several tables (strata x CPUEs), corresponding respectively to the situation before and after the modification, could be analysed with a multi-table technique. When each table describes yearly fishing activities, MFA is also a useful tool for detecting and removing an unusual year before examining the results over the period analysed.
- Between-class PCA and within-class PCA appear to be an interesting tool for spatio-temporal studies addressing the questions "where" (spatial effect) and "when" (temporal effect). The same methods can be applied to examine the influence of specific factors, such as fleet effects, fishing gear effects, etc., by successively removing others and focusing on the factor of interest.
- In this study, PCAIV was used to link an abundance matrix with an auxiliary information table (associated floating objects). Coupling regression analysis with the PCA enabled the relationships between the CPUEs to be examined while taking into account the different fishing strategies adopted for the different size classes of purse seiners. Other instrumental variables, such as environmental factors, prices for the different size categories of tunas etc., must be investigated. PCAIV is also a useful way for analysing separately the effects of the main factors and the interactions of a saturated linear model (Pech and Laloë, 1997). After partitioning the initial table into several fitted matrices, obtained by regressing the factors of the model (i. e., the influential variables), a PCA of each of these arrays allows analysis of each factor independent of other sources of variation.

Acknowledgments

We thank the two anonymous referees for helpful comments given and the editor for grammatical corrections.

References

CHARLES A. T., 1995 - Fishery science: the study of fishery systems. Aquat. Living Resour. 8, 233-239.

DOLÉDEC S. and D. Chessel, 1987 - Rythmes saisonniers et composantes stationnelles en milieu aquatique. I. Description d'un plan d'observation complet par projection de variables. Acta Oecologia, Oecol. Gener. 8, 3, 403-426.

DOLÉDEC S. and D. Chessel, 1989 - Rythmes saisonniers et composantes stationnelles en milieu aquatique. II. Prise en compte et élimination d'effets dans un tableau faunistique. Acta Oecologia, Oecol. Gener. 10, 3, 207-232.

DOLÉDEC S. and D. Chessel, 1991 - Recent developments in linear ordination methods for environmental sciences. Advances in ecology. India, 1, 133-155.

CHESSEL D., J.-D. Lebreton and N. Yoccoz, 1987 - Propriétés de l'analyse canonique des correspondances: une illustration en hydrobiologie. Revue de statist. appl. 35, 4, 55-72.

EHRHARDT N. M., 1994 - Analysis of the correlation among indices of relative abundance of tunas and billfish species caught in the taiwanese and japanese atlantic longline fisheries. In: Report of the second ICCAT billfish workshop. Inter. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41, 458-472.

ESCOFIER B. and J. Pages, 1990 - Analyses factorielles simples et multiples. Objectifs, méthodes et interprétation. Dunod Eds., Paris, 267 p.

ESCOFIER B. and J. Pages, 1994 - Multiple factor analysis (AFMULT package). Computational statistics and data analysis. 18, 121-140.

FIEDLER P. C. and S. B. Reilly, 1993 - Interannual variability of dolphin habitats in the eastern tropical Pacific. II Effects on abundances estimated from tuna vessel sightings. 1975-1990. Fish. Bull. U.S., 92, 451-463.

GAERTNER D., M. Medina-Gaertner, C. Castillo and L. Martinez, 1988 - Instalacion de un sistema de bitacoras para la pesca atunera de superficie en Venezuela y analisis de los primeros resultados. Inter. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 28, 130-140.

GAERTNER D., M. Pagavino and J. Marcano, 1996 - Utilisation de modèles linéaires généralisés pour évaluer les stratégies de pêche thonière à la senne en présence d'espèces associées dans l'Atlantique ouest. Aquat. Living Resour. 9, 305-323.

GOWER J. C., 1975 - Generalized procustes analysis. Psychometrika, 40, 1, 33-51.

HILBORN R. and C. J. Walters, 1992 - Quantitative fisheries stock assessment. Choice, Dynamics and uncertainty. Chapman and Hall, New York. 570 p.

KROONENBERG P. M. and P. Brouwer, 1985 - User's guide to TUCKALS3, version 4.0. Department of education. Leiden University, Leiden. Netherlands.

LAVIT Ch., Y. Escoufier, R. Sabatier and P. Traissac, 1994 - The ACT (Statis method). Computational statistic and data analysis, 18, 97-119.

LEBART L., A. Morineau and M. Piron, 1995 - Statistique exploratoire multidimensionnelle. Dunod Edns. Paris. 439 p.

LEBRETON J.-D., D. Chessel, R. Prodon and N. Yoccoz, 1988a - L'analyse des relations espèces-milieu par analyse canonique des correspondances. I Variables de milieu qualitatives. Acta Oecologica, Oecol. Gener., 9, 53-67.

LEBRETON J.-D., D. Chessel, M. Richardot-Coulet and N. Yoccoz, 1988b - L'analyse des relations espèces-milieu par analyse canonique des correspondances. II Variables de milieu quantitatives. Acta Oecologica, Oecol. Gener., 9, 137-151.

LEBRETON J.-D., R. Sabatier, G. Banco and 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. p. 85-114. In: J. Devillers and W. Karcher (Eds.). Applied Multivariate Analysis in SAR and Environmental studies. Kluwer.

PRODON R. and J.-D. Lebreton, 1994 - Analyses multivariées des relations Espèces-milieu: structure et interprétation écologique. Vie et milieu, 44, 1, 69-91.

PECH N. and F. Laloë, 1997 - Use of principal component analysis with instrumental variables (PCAIV) to analyse fisheries catch data. ICES Journ. of Marine Science, 54, 32-47.

REILLY S. B. and P. C. Fiedler, 1993 - Interannual variability of dolphin habitats in the eastern tropical Pacific. I. Research vessel surveys. 1986-1990. Fish. Bull. U.S., 92, 434-450.

SABATIER R., 1984 - Quelques généralisations de l'analyes en composantes principales de variables instrumentales. Statis. et anal. des données, 9, 3, 75-103.

SABATIER R., 1987 - Analyses factorielles de données structurées et métriques. Statis. et anal. des données. 12, 3, 75-96.

TER BRAAK C. J. F., 1986 - Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology, 67, 5, 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 correspondence analysis. p. 551-558 In: H. H. Bock (ed.) Classification and related methods of data analysis. North Holland, Amsterdam.

THIOULOUSE J., S. Dolédec, D. Chessel, and J. M. Olivier, 1995 - ADE Software: Multivariate analysis and graphical display of environmental data. Proceeding of the 4th International Software Exibition for Environmental Science and Engineering ; 57-62.