

Analyse de la composante chronologique dans les tableaux croisés: une approche descriptive

GÉRARD ANTILLE*

1. INTRODUCTION

L'analyse des données, qui a connu et connaît encore un rapide développement grâce à la microinformatique en particulier, s'est imposée comme un excellent outil de description de grands ensembles de données. Les différentes méthodes de réduction de dimension ont fait leur preuve tant en sciences expérimentales, qu'en sciences sociales, qu'en psychologie entre autres, pour l'analyse de tableaux à double entrée.

Les techniques d'analyse des suites temporelles de tableaux à double entrée, développées dans les années 80, voir KROONENBERG (1983), ESCOUFIER (1985), ne sont que trop peu utilisées. Dans cet article nous désirons contribuer à la diffusion de quelques approches de cette problématique et présenter une méthode originale d'étude descriptive de tableaux croisant individus et variables et indicés par le temps.

Deux ensembles de données concernant les cantons suisses vont servir de support à l'illustration de notre propos. Nous nous sommes intéressés à l'évolution des dépenses cantonales ventilées selon la division des tâches des institutions publiques. La seconde illustration décrit l'évolution des taux de différentes causes de décès au cours du temps. Dans ces deux contextes les unités statistiques sont donc les cantons suisses pour lesquels deux ensembles de variables ont été observées au cours du temps.

Dans cette étude un ensemble de données est un tableau tridimensionnel formé par une suite des tableaux de données croisant les unités et les variables. On notera par x_{ijt} l'observation de la j -ème variable de l'unité statistique i à l'époque t , ainsi un tableau tridimensionnel de données se note $X = \{x_{ijt} | i = 1, \dots, n; j = 1, \dots, p; t = 1, \dots, T\}$. Quant aux données de l'année t elles sont représentées par la matrice X_t dont l'élément générique est x_{ijt} . Le tableau X peut donc aussi s'écrire $X = \{X_t | t = 1, \dots, T\}$.

Les méthodes d'analyse de données de la famille des analyses en composantes principales décrivent les données de chaque matrice X_t indépendamment les unes des autres et proposent une description simplifiée de la structure de corrélation de chaque tableau X_t . L'analyse canonique généralisée ainsi que les méthodes dérivées permettent de met-

* Chargé de cours à l'Université de Genève et à l'Université de Neuchâtel.

Adresse: Département d'économétrie, Université de Genève, Uni-mail, 1211 Genève 4.

L'auteur remercie Mesdames Mie Inoue et Jeanine Munyeshuli qui ont contribué, grâce aux logiciels MATLAB et S-PLUS, à la réalisation numérique et graphique de cette communication lors de la rédaction de leur mémoire de diplôme au département d'économétrie de l'Université de Genève.

tre en évidence les éléments communs à tous les tableaux ou à une partie d'entre eux. L'analyse en composantes principales généralisées décrit simultanément les matrices X_t en exprimant les éléments communs aux différents tableaux ainsi que les caractéristiques propres à chacun d'entre eux. La méthode STATIS permet la représentation graphique des individus au cours du temps et ainsi de comparer leurs évolutions respectives.

La difficulté majeure dans la construction des méthodes de réduction de dimension réside dans le choix de directions de projection, choix basés sur un ou des critères d'optimalité construits à l'aide, essentiellement, de mesures de dispersion. Une autre approche qui conduit souvent à des résultats analogues consiste à déterminer des variables synthétiques, résumant les tableaux de départ, et vérifiant des critères d'optimalité basés sur des mesures de liaison entre ces variables synthétiques. Dans notre approche les directions de projection sont définies par une analyse en composantes principales effectuées sur X_c , une matrice compromis résumant les T matrices X_t composant X le tableau tridimensionnel des données. L'évolution temporelle des unités statistiques est visualisée le long des axes principaux de la matrice compromis et définit ainsi une trajectoire par unité statistique. L'application de méthodes de classification aux trajectoires permet d'effectuer des regroupements d'unités statistiques à composantes temporelles semblables.

Dans cet article nous présentons, dans la section 2, quelques méthodes importantes d'analyse de tableaux tridimensionnels afin de préciser les différences avec la technique d'analyse que nous proposons et décrivons dans la section 3. Dans la section 4 nous développons deux exemples d'application, l'un autour de la méthode STATIS, l'autre selon notre approche. Dans les deux cas nous nous intéressons à l'évolution temporelle des cantons suisses par rapport, d'une part aux dépenses cantonales ventilées selon les tâches des institutions et d'autre part à différentes causes de mortalité. Des groupes de cantons à comportement similaires sont mis en évidence.

2. QUELQUES METHODES CLASSIQUES

2.1 *L'analyse canonique généralisée et l'analyse discriminante de tableaux*

L'analyse canonique généralisée recherche T variables synthétiques $Z_t = X_t b_t$, chacune d'elle résumant un tableau X_t , et une variable Z de telle manière que $\sum R^2(Z, Z_t)$ soit maximale où R^2 est le coefficient de détermination entre Z et Z_t . La variable Z est appelée la première composante canonique. En posant $H_t = X_t(X_t'X_t)^{-1}X_t'$ et $H = \sum H_t$ on montre que $\sum R^2(Z, Z_t) = Z'HZ$ et par conséquent Z est le vecteur propre de H correspondant à la plus grande valeur propre de H , voir par exemple CASIN (1995).

Puis le procédé est réitéré sous la contrainte d'orthogonalité entre les variables canoniques.

L'analyse discriminante de tableaux diffère de l'analyse canonique généralisée par la contrainte d'orthogonalité imposée dans ce cas, à chaque étape, entre les variables syn-

thétiques. On montre que, à l'étape i , les variables synthétiques étant données, la variable canonique Z^i est la première composante principale des variables synthétiques $Z_t^i = X_t b_t^i$. De plus la variable Z^i est un multiple de la somme des variables synthétiques de l'itération i .

Une présentation complète des deux méthodes est donnée dans les articles de CASIN et THURLOT (1986), PONTIER et NORMAND (1992) et CASIN (1995).

Dans les deux cas on représente, sur les axes des graphiques, les variables synthétiques et la variable canonique associée qui est leur centre de gravité. Comme CASIN (1995) le fait très justement remarquer les graphiques ne sont pas très lisibles et il lui paraît plus judicieux de représenter les unités statistiques dans chacun des sous-espaces engendrés par les colonnes de chaque tableau et d'effectuer des interprétations tableau par tableau. Bien que les projections soient solutions de l'optimisation de la somme des coefficients de détermination, ces méthodes ne permettent pas d'obtenir une vision temporelle de ces tableaux.

2.2 *L'analyse en composantes principales généralisée*

L'analyse en composantes principales généralisée détermine une variable auxiliaire Z et T variables synthétiques Z_t , projections orthogonales de Z sur les espaces engendrés par les colonnes des différentes matrices X_t telles que $\sum R^2(Z, Z_t)$ et $\sum \text{cov}^2(Z, X_{j,t})$ soient maximales sous la contrainte de normalisation de Z et où $X_{j,t}$ représente la variable j du tableau t . Puis le procédé est itéré sous la contrainte d'orthogonalité des Z_t^i . On reconnaît la combinaison des critères définissant l'analyse canonique généralisée, sa généralisation à l'analyse discriminante des tableaux et les critères définissant l'analyse en composantes principales de chaque tableau.

Une présentation de cette méthode est donnée dans CASIN (1996).

L'analyse en composantes principales généralisée s'apparente aux techniques de régression PLS (TENNELHAUS, GAUCHI et MÉNARDO, 1995) qui s'appliquent au cas de deux tableaux. Dans cette optique on recherche une variable auxiliaire Z et T variables synthétiques $Z_t = X_t b_t$ maximisant $\sum \text{cov}^2(Z, Z_t)$ sous contrainte de normalisation de Z . Puis le procédé est réitéré sous la contrainte d'orthogonalité des variables synthétiques. L'analyse en composantes principales généralisée se distingue des techniques PLS par la propriété supplémentaire de corrélation entre les variables Z^i .

Les représentations graphiques sont les mêmes qu'en analyse en composantes principales. Il est aussi possible d'effectuer des représentations communes qui permettent de repérer des unités atypiques ainsi que des comparaisons d'évolution par rapport au temps mais l'interprétation des axes de projection est délicate car les corrélations avec les variables sont différentes d'une unité statistique à l'autre (CASIN, 1996).

2.3 La méthode STATIS

La méthode STATIS, pour Structuration des Tableaux à Trois Indices de la Statistique, est construite à partir des vecteurs propres et valeurs propres de la matrice des corrélations vectorielles entre tableaux; notons C cette matrice.

Rappelons que le coefficient de corrélation vectorielle RV est une mesure de lien entre deux tableaux définie par $RV(X_i, X_j) = \text{tr}(X_i X_i' X_j X_j') / (\text{tr}(X_i X_i') \text{tr}(X_j X_j'))$. De plus on dit que deux tableaux sont équivalents si leur $RV = 1$ et dans ce cas on montre que la structure des individus d'un tableau est, à un scalaire près, similaire à la structure de l'autre.

En notant $\nu' = (\nu_1, \nu_2, \dots, \nu_T)$ le premier vecteur propre de la matrice C , on définit une matrice compromis W , résumé des matrices $X_t X_t'$, par l'égalité $W = \sum \nu_t X_t X_t'$. Cette matrice compromis est vectoriellement la plus corrélée, parmi toutes les combinaisons linéaires des matrices $X_t X_t'$, avec chacune des $X_t X_t'$. Les premiers vecteurs propres de la matrice compromis déterminent les axes de projections des variables

La représentation graphique des unités statistiques des tableaux X_t , s'effectue par projections sur les vecteurs propres de la matrice compromis des T matrices de covariances, compromis construit selon le schéma ci-dessus. Il est ainsi possible de visualiser l'ensemble des positions au cours du temps de chaque unité. L'ensemble de ces positions définit la trajectoire d'une unité statistique. L'interprétation des axes du compromis est basée sur la corrélation entre chacune des variables et l'axe en question et suppose, par soucis de cohérence, que les vecteurs propres du compromis W soient égaux à ceux de $X_t X_t'$ pour tout t , contrainte irréaliste. Cependant la comparaison des trajectoires est très informative quant à l'évolution temporelle des différentes unités statistiques.

La méthode STATIS est largement décrite dans LAVIT (1988).

3. COMPOSANTES PRINCIPALES DES MOYENNES ET CLASSIFICATION

Les descriptions ci-dessus ont mis en évidence quelques propriétés d'optimalité des différentes méthodes mais aussi certains points faibles tels que la non lisibilité des graphiques ou la difficulté d'interprétation des axes de projection ou encore la signification des matrices compromis. Ces méthodes présentent toutes un intérêt théorique évident comme solution de différents problèmes d'optimisation mais mettent en œuvre des constructions parfois sophistiquées qui ne permettent pas d'obtenir une vision simultanée des différents tableaux.

Dans notre approche nous considérons le temps comme la variable expliquant le passage d'un tableau à un autre et avons comme objectif de faire ressortir des structures inter-temporelles parmi les unités statistiques. Par soucis de simplicité et surtout de compréhension et d'interprétation nous avons choisi comme matrice compromis le tableau M , moyenne des X_t . Ce choix est arbitraire. On définit ainsi n unités statistiques fictives ayant pour valeurs les moyennes par rapport au temps des variables de tous les

tableaux X_t . Ainsi m_{ij} l'élément générique de la matrice M est donné par $m_{ij} = (1/T) \sum x_{ijt}$ et l'unité compromise k , élément fictif, est définie par la $k^{\text{ème}}$ ligne de M .

Il s'agit ensuite d'effectuer une analyse en composantes principales de la matrice M centrée et réduite afin de déterminer les vecteurs et valeurs propres de la matrice des corrélations associées à M . Les premiers vecteurs propres servant à la construction du sous espace des projections des individus fictifs, sous espace s'interprétant selon les principes classiques de l'analyse en composantes principales.

L'idée élémentaire suivante consiste à considérer chaque tableau X_t comme un nouvel ensemble d'individus à localiser dans l'espace de représentation des individus compromis moyens. Cette localisation n'ayant de sens que si les corrélations entre les variables et les axes de projection demeurent stables au cours du temps.

Pour une unité donnée les projections successives de ces valeurs au cours du temps déterminent sa trajectoire dans l'espace de représentation des individus compromis. La lisibilité de tels graphiques n'est possible, dès que le nombre d'individus est supérieur à quelques unités, qu'en dimension un ou deux. En visualisant les trajectoires de toutes les unités le long d'un seul axe principal, il devient plus aisé d'effectuer des comparaisons inter-unités. Cependant même dans cette situation le problème de lisibilité demeure.

Afin de palier à cette difficulté, nous proposons d'appliquer des méthodes de classifications aux différentes trajectoires ce qui permet de regrouper des unités statistiques et de mettre en évidence une structure inter-temporelle entre les individus, objectif de cet article. Remarquons cependant, que les résultats sont fonction de la mesure de similarité ou de distance choisie entre les trajectoires ainsi que de l'algorithme de classification de celles-ci. Par absence de fondement théorique nous utiliserons la distance euclidienne qui correspond à la description usuelle de notre environnement physique. Les trajectoires les plus proches permettront de regrouper les unités statistiques correspondantes et ainsi de constituer des groupes dont les tailles vont en s'accroissant pour finalement regrouper tous les individus. Dans ce cas on parle de méthodes de classification hiérarchique ascendante.

Signalons encore que l'étude de l'évolution des variables s'effectue par dualité comme nous le montrerons dans un des exemples traités dans la section suivante.

Comme mentionné précédemment le choix de la matrice compromise est arbitraire et comme toutes les moyennes notre compromis souffre de non robustesse mais possède la propriété d'optimalité liées aux moindres carrés. Le choix de la mesure de proximité entre les trajectoires des individus est lui aussi arbitraire et de plus il en est de même pour l'algorithme de classification. En conséquence, il est indispensable de conserver un œil critique face aux résultats et de ne les considérer avant tout que comme une description possible de la réalité.

L'avantage principal de cette méthode en est sa simplicité tant dans sa réalisation que dans les interprétations des résultats comme les exemples l'illustreront.

4. APPLICATIONS

Dans les deux exemples que nous allons traiter les unités statistiques sont les vingt-six cantons et demi-cantons suisses pour lesquels deux ensembles distincts de variables ont été observées au cours du temps. Notre objectif est de décrire l'évolution temporelle des cantons par rapport aux variables les caractérisant.

4.1 Dépenses des cantons selon la répartition des tâches des institutions publiques

Dix-huit variables mesurant la dépense annuelle cantonale définie selon une classification fonctionnelle ont été observées entre 1980 et 1989. On est donc en présence d'un tableau tridimensionnel $26 \times 18 \times 9$. La liste des variables est donnée en annexe.

Les disparités, entre les dépenses totales des cantons, rendent évident l'usage des profils des cantons et non celui des valeurs absolues des dépenses dans la suite de cette illustration. Ainsi pour chaque unité statistique nous prenons en compte la part dépensée dans les différents secteurs.

La lecture de la matrice des corrélations vectorielles, présentées dans la Table 1, montre que plus les tableaux sont éloignés dans le temps plus leur corrélation diminue et que, entre deux années consécutives, les *RV* indiquent que les structures se ressemblent.

Table 1: Corrélations vectorielles

	1981	1982	1983	1984	1985	1986	1987	1988	1989
1980	0.9666	0.8599	0.8563	0.7898	0.7730	0.7794	0.7023	0.6372	0.6158
1981		0.9238	0.9177	0.8204	0.7906	0.7904	0.7206	0.6288	0.6079
1982			0.9784	0.9205	0.8678	0.8693	0.8360	0.7260	0.7200
1983				0.9205	0.8821	0.8730	0.8434	0.7332	0.7280
1984					0.9679	0.9452	0.9455	0.8684	0.8523
1985						0.9574	0.9468	0.8935	0.8682
1986							0.9664	0.9111	0.8891
1987								0.9549	0.9425
1988									0.9666

Le plan principal, associé à la matrice des corrélations vectorielles, contient 67 % de l'inertie totale et est caractérisé par les variables „santé“, „enseignement et recherche“ opposé à „trafic et énergie“ pour le premier axe. Quant au deuxième axe il oppose „santé“ et „enseignement et recherche“. On observe aussi une très grande stabilité dans le temps des covariances entre les variables et les axes du plan principal.

Figure 1: Image du compris

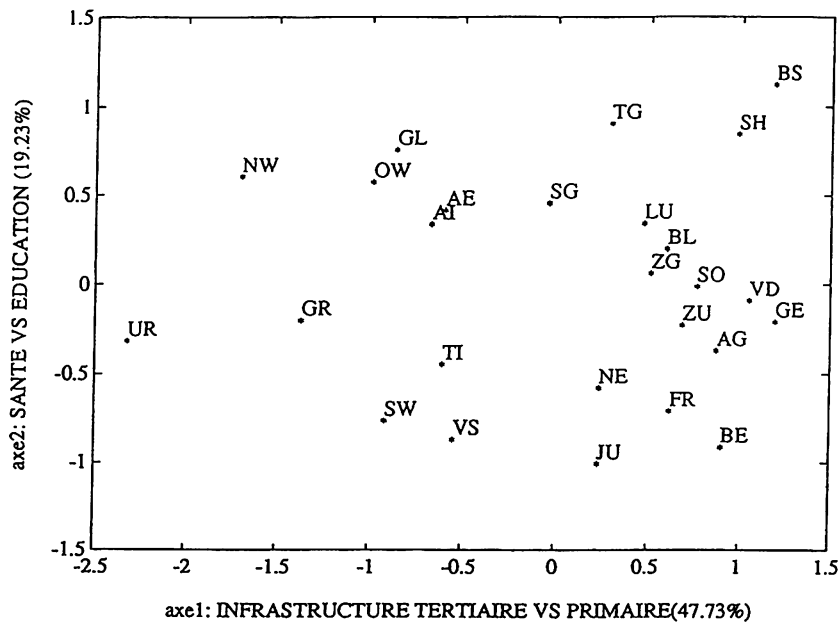
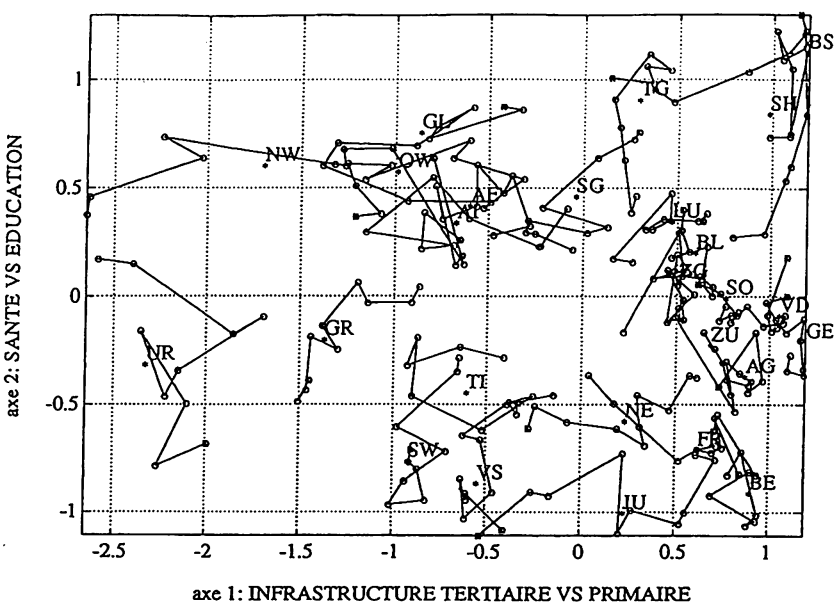


Figure 2: Trajectoires des cantons



L'image du compromis des individus, associé à la méthode STATIS, Figure 1, montre l'opposition entre les petits cantons de la Suisse centrale plus les deux Appenzell, le Valais, les Grisons ainsi que le Tessin d'une part et Genève, Bâle Ville et Vaud d'autre part, opposition qui s'exprime en termes d'infrastructure primaire et tertiaire selon le premier axe.

Comme on peut le constater dans la Figure 2, la représentation des trajectoires des cantons autour de leur représentant dans la matrice compromis est extrêmement délicate et ne permet que difficilement des comparaisons dans le temps.

4.2 Causes de décès dans les cantons suisses

Dans cette seconde illustration l'ensemble de données contient les observations de sept causes de décès (infections, tumeurs, maladies cardio-vasculaires, maladies respiratoires, accidents, suicides, autres), observées entre les années 1969 et 1995. On est donc en présence d'un tableau tridimensionnel $26 \times 7 \times 26$. Comme dans l'exemple précédent, les différences de taille des populations des cantons nous amènent à considérer les profils des cantons.

Table 2: Profils du compromis

	infection	tumeur	cardiovasc	respiratoire	accident	suicide	autre
ZH	0.014	0.112	0.390	0.042	0.051	0.030	0.361
BE	0.010	0.114	0.388	0.062	0.051	0.029	0.346
LU	0.009	0.122	0.386	0.050	0.062	0.027	0.345
UR	0.010	0.122	0.398	0.048	0.077	0.025	0.319
SZ	0.010	0.120	0.383	0.046	0.075	0.024	0.342
OW	0.011	0.108	0.405	0.052	0.075	0.023	0.326
NW	0.013	0.121	0.390	0.039	0.087	0.022	0.329
GL	0.013	0.110	0.394	0.052	0.060	0.023	0.349
ZG	0.015	0.118	0.376	0.046	0.063	0.032	0.349
FR	0.011	0.127	0.365	0.045	0.076	0.031	0.344
SO	0.009	0.125	0.377	0.053	0.053	0.030	0.353
BS	0.013	0.115	0.347	0.063	0.035	0.026	0.400
BL	0.010	0.134	0.365	0.054	0.049	0.035	0.352
SH	0.010	0.113	0.402	0.053	0.044	0.027	0.350
AR	0.008	0.097	0.425	0.057	0.049	0.032	0.332
AI	0.009	0.093	0.467	0.047	0.063	0.039	0.282
SG	0.010	0.108	0.411	0.049	0.061	0.028	0.333
GR	0.012	0.112	0.387	0.059	0.066	0.028	0.336
AG	0.010	0.119	0.378	0.053	0.065	0.032	0.343
TG	0.008	0.104	0.425	0.048	0.065	0.030	0.319
TI	0.011	0.130	0.394	0.036	0.048	0.017	0.364
VD	0.013	0.115	0.355	0.050	0.057	0.032	0.379

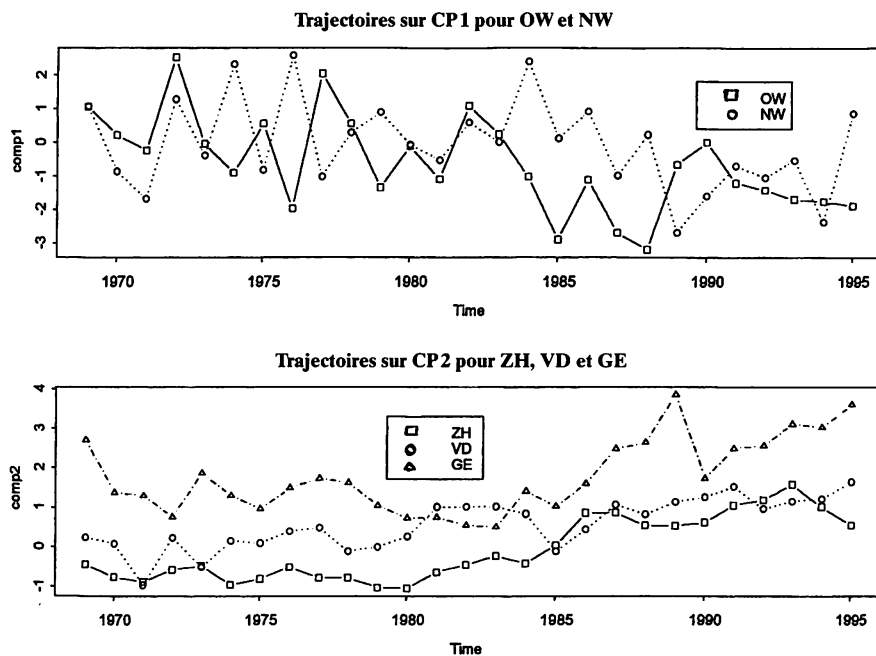
	infection	tumeur	cardiovasc	respiratoire	accident	suicide	autre
VS	0.015	0.123	0.333	0.047	0.094	0.027	0.362
NE	0.011	0.115	0.358	0.047	0.057	0.032	0.381
GE	0.017	0.115	0.326	0.046	0.051	0.030	0.415
JU	0.011	0.116	0.352	0.056	0.069	0.025	0.371

Le tableau compromis M est donc une matrice 26×7 dont la somme de chaque ligne est égale à un.

L'analyse en composantes principales appliquée aux individus du compromis donne au plan principal un pouvoir explicatif égal à 65 %. Ce plan est caractérisé par les variables „maladies cardio-vasculaires“ et „suicides“ opposées aux autres causes de décès selon le premier axe. Quant au deuxième axe, il oppose „maladies respiratoires“ à „accidents“. Remarquons que les quatre axes principaux expliquent 90 % de l'inertie et contiennent l'information pertinente. Ainsi l'axe trois oppose la variable „tumeurs“ aux autres causes de décès et le quatrième axe la variable „suicides“ aux autres causes.

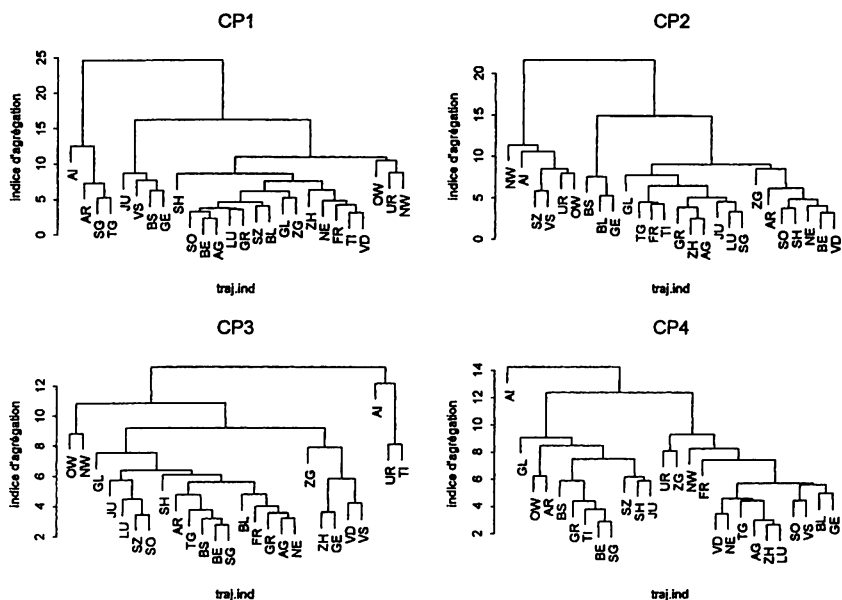
Les projections des vingt-six cantons le long de chacun de ces axes donnent vingt-six trajectoires qui se superposent et ne permettent que difficilement de mettre en évidence des structures inter-temporelles parmi ces cantons. Quelques trajectoires de paires de cantons sont présentées dans la Figure 3.

Figure 3: Quelques trajectoires



Un algorithme de classification hiérarchique ascendante appliqué à la matrice des distances euclidiennes entre trajectoires a été utilisé pour construire les arbres de classification permettant de visualiser et ainsi de découvrir les groupes de cantons à „ comportement similaire „ au cours du temps. Le long du premier axe on notera que les cantons SG, TG, AR et AI composent un groupe très distinct des autres cantons, GE, BS, VS, et JU forment aussi un groupe homogène de même que NW, UR, et OW. A l'intérieur de chacun de ces groupes la mortalité due aux maladies cardio-vasculaires et dans une moindre mesure aux suicides a évolué de manière analogue mais, entre ces groupes l'évolution est différente ainsi que par rapport aux autres cantons. Une analyse plus approfondie s'obtient par une lecture détaillée des données initiales. Les arbres correspondant aux classifications des trajectoires le long des axes 2, 3 et 4 sont aussi donnés dans la Figure 4.

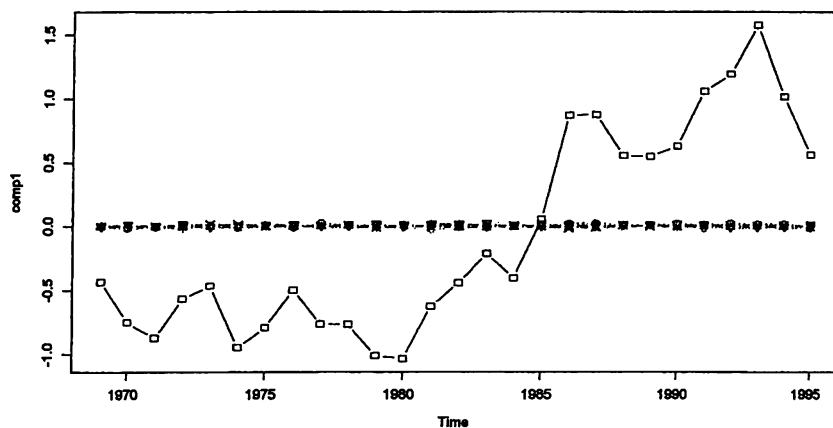
Figure 4: Arbres de classification



Globalement on constate que AI a un comportement que l'on peut qualifier d'atypique, mais il n'est pas impossible que cela soit simplement dû au petit nombre d'habitants de ce canton. On notera aussi la proximité de ZH et des cantons romands le long des axes 1, 3 et 4.

L'analyse des variables montre une très grande stabilité par rapport au temps à l'exception de la variable Infections qui possède une trajectoire particulière (Figure 5). Remarquons que cette modification de trajectoire intervient autour de 1985, période qui correspond à la prise en compte du Sida parmi les maladies infectieuses.

Figure 5: Trajectoires des variables



5. CONCLUSIONS

Dans cette communication nous avons présenté une méthode descriptive d'analyse de tableaux tridimensionnels unités x variables x temps. Cette méthode est basée sur l'analyse en composantes principales d'une matrice compromise, en l'occurrence la matrice moyenne par rapport au temps. Cette ACP définit les directions de projections des différentes unités statistiques et permet de visualiser leur trajectoire dans le temps. Une méthode de classification de ces trajectoires construit des groupes d'unités statistiques à comportement temporel similaire. L'illustration présentée montre la simplicité de la méthode et la facilité d'interprétation des résultats.

ANNEXE

Liste des variables utilisées dans la première illustration.

- | | |
|-------------------------------|---|
| a) Administration publique | j) Prévoyance sociale |
| b) Justice | k) Aménagement du territoire |
| c) Police et service du feu | l) Trafic, énergie |
| d) Défense nationale | m) Agriculture |
| e) Enseignement et recherche | n) Sylviculture, chasse, pêche |
| f) Culture, loisirs et sports | o) Cours d'eau, avalanche |
| g) Culte | p) Tourisme, industrie, artisanat, commerce |
| h) Santé | q) Intérêts passifs |
| i) Hygiène du milieu | r) Dépenses du service financier. |

BIBLIOGRAPHIE

- CASIN, PH. (1995), „L'analyse discriminante des tableaux évolutifs“, *Revue de Statistique Appliquée*, XLIII, 3, 73–91.
- CASIN, PH. (1996), L'analyse en composantes principales généralisée, *Revue de Statistique Appliquée*, XLIV, 3, 63–81.
- CASIN, PH. et J. C. THURLOT (1986), „Une présentation de l'analyse canonique généralisée dans l'espace des individus“, *Revue de Statistique Appliquée*, XXXIV, 3, 65–75.
- ESCOUFIER Y. (1985), „Objectifs et procédures de l'analyse conjointe de plusieurs tableaux de données“, *Statistique et analyse de données*, 10, 1–10.
- KROONENBERG P. M. (1983), *Three-mode principal component analysis. Theory and applications*, Leiden, Reprint 1989.
- LAVIT, CH. (1988), *Analyse conjointe de tableaux quantitatifs*, Masson.
- PONTIER, J. et NORMAND M. (1992), „A propos de généralisation de l'analyse canonique“, *Revue de Statistique Appliquée*, XL, 1, 57–75.
- TENENHAUS, M., J. P. GAUCHI et C. MÉNARDO (1995), „Régression PLS et Applications“, *Revue de Statistique Appliquée*, XLIII, 1, 7–63.

RESUME

Dans cet article nous développons une nouvelle technique d'analyse de données appliquées à des tableaux croisés indicés par le temps. L'objectif de cette technique est la description des évolutions temporelles des unités statistiques. Cette méthode est basée sur les principes d'analyses des tableaux tridimensionnels, telles que analyse en composantes principales, analyse canonique et analyse discriminante généralisées ainsi que la méthode STATIS et est complétée par une classification automatique des trajectoires dans le temps des unités statistiques.

SUMMARY

In this article we develop a new technique of data analysis applied to three dimensional tables. The main goal of that method is to describe the evolution of the statistical units with respect to the time. Inspired by the generalized principal components analysis, canonical analysis and discriminant analysis as well as the STATIS method, our technique exhibits similar statistical units provided by a classification of their trajectories.

ZUSAMMENFASSUNG

In diesem Aufsatz wird eine neue Methode zur Analyse von Dreiwegtabellen entwickelt. Das Hauptziel der Methode ist die Beschreibung der Dynamik der untersuchten Einheiten. Basierend auf Verfahren der verallgemeinerten Hauptkomponentenanalyse, der kanonischen Korrelationsanalyse und der Diskriminanzanalyse sowie der STATIS-Methode reproduziert unser Verfahren Trajektorien, die den zugrundeliegenden Daten ähnlich sind.