

MONIQUE LE GUEN

Séries chronologiques et analyse de données, des images de synthèse pour répondre à la question : ex-post, la classification COB des sicav est-elle pertinente ?

Journal de la société statistique de Paris, tome 135, n° 1 (1994), p. 37-71

http://www.numdam.org/item?id=JSFS_1994__135_1_37_0

© Société de statistique de Paris, 1994, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

III

ARTICLES

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES, DES IMAGES DE SYNTHÈSE POUR RÉPONDRE À LA QUESTION :

EX-POST, LA CLASSIFICATION COB DES SICAV EST-ELLE PERTINENTE ?

Par Monique LE GUEN
CNRS, Institut Orléanais de Finance

Mots clés : Séries chronologiques, Analyses de données, Méthodologie de traitement, Visualisation, Finance, Sicav, Interdisciplinarité.

Résumé

Dans les données d'enquêtes, les techniques d'analyses descriptives multidimensionnelles ont bouleversé l'enchaînement des méthodes traditionnelles.

En appliquant ces techniques à des séries chronologiques de rentabilités des sicav, nous avons pu révéler des formes stables montrant l'homogénéité des catégories de sicav, analysées ex-post, et mettant en évidence la régularité des gestionnaires de sicav.

En complément, les anomalies de classement nous ont permis de repérer les sicav au comportement atypique par rapport à leur catégorie d'origine.

*L'apport des techniques d'analyse de données a été de mettre en évidence une organisation qui ne pouvait être appréhendée par des techniques statistiques classiques.**

* Je remercie très sincèrement le Professeur CAZES pour ses encouragements et ses commentaires.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

INTRODUCTION

Dans le domaine des sciences économiques et sociales, il existe deux grands types de données qui donnent prise à des traitements statistiques : les données d'enquêtes, enquêtes d'opinion ou enquêtes économiques, et les séries chronologiques.

Pour les premières, les traitements ont été profondément modifiés par l'analyse des données ou analyses descriptives multidimensionnelles. « Ces techniques ont bouleversé l'enchaînement des méthodes traditionnelles » (Lebart, 1989)¹. Dans l'ordonnement traditionnel du dépouillement d'enquêtes : tris à plat et croisées avec calculs de pourcentages et de moyennes, suivis de méthodes plus sophistiquées : régression, analyse de la variance ou covariance, modèles log-linéaire, est venue s'insérer une étape de structuration des données, typologies et sélection de tableaux croisés pertinents.

Cette nouvelle approche du traitement de l'information repose sur les outils complémentaires que sont l'analyse factorielle et la classification².

Dans le traitement des séries chronologiques, que ce soit dans un but cognitif, c'est-à-dire porteur de connaissance, ou un but de prévision, la modélisation a toujours été retenue comme premier objectif.

Il nous a semblé que des analyses descriptives multidimensionnelles appliquées à des séries chronologiques pourraient apporter un autre point de vue sur la structure de l'information, et suggérer, comme en dépouillement d'enquêtes, des orientations pour une modélisation postérieure.

Notre activité actuelle sur une base de données de sicav nous a permis de tester cette nouvelle stratégie.

Cette étude se décomposera en 4 chapitres.

Après une présentation, au chapitre 1, du domaine d'analyse, à savoir ce que sont les sicav, de quelles informations et mesures on dispose, quelle est la question que l'on se pose, nous survolerons au chapitre 2 les différentes approches de traitements des séries chronologiques utilisées en économie et finance.

Au chapitre 3, nous présenterons les premiers traitements que l'on peut effectuer sur des séries chronologiques, graphiques, analyse de variance et étude des rangs, pour aborder, au chapitre 4, l'apport de l'analyse de données avec la complémentarité entre les analyses factorielles (analyse en composantes principales et analyse factorielle des correspondances) et la classification ascendante hiérarchique.

1. Lebart L., *Stratégie du traitement des données d'enquêtes*. Revue de MODULAD-INRIA, juin 1989, n° 3, p. 21-29.

2. Les programmes informatiques et l'enseignement sont largement développés dans la communauté française des statisticiens. Citons au hasard ADDAD, MODULAD, SICLA, SPAD.

1. Les informations et leur contexte

1.1. Que sont les sicav ?

La gestion collective de l'épargne assurée par les Organismes de Placement Collectif en Valeurs Mobilières est apparue en France en 1964. Ces OPCVM constituent et gèrent pour le compte de tiers des portefeuilles de valeurs mobilières¹.

Parmi les OPCVM les plus populaires, les Sociétés d'Investissement à Capital Variable, ou sicav, ont connu ces dernières années un très vif succès auprès du public. Ce succès pourrait même dans un proche avenir conduire les autorités compétentes à modifier les règles du jeu. En effet, l'importance de la collecte et en particulier le gonflement continu des sicav monétaires provoquent des déséquilibres entre l'épargne courte et l'épargne longue : il y a beaucoup d'épargne fugace et peu d'épargne investie. D'où les désordres que nous subissons.

Selon Europerformance², filiale de Paribas, chargée de collecter, traiter et diffuser l'information relative aux OPCVM, la France compte aujourd'hui 11 millions de porteurs de parts de sicav. 1 000 produits se partagent le marché et le total des capitaux gérés a atteint 1 760 milliards de francs en 1992. A titre de comparaison, moins de 6 millions de Français possèdent des actions.

C'est dire l'importance des sommes en jeu, l'importance du nombre des porteurs et l'importance des informations relatives aux sicav que le porteur en titre et potentiel doit connaître, eu égard à la diversité des produits offerts.

Pour faciliter le repérage des sicav, la Commission des Opérations de Bourse classe les sicav selon des critères de répartition de portefeuilles en actions et obligations, et selon des critères de placement à court terme, moyen ou long terme. L'objectif de ce classement est d'assurer une bonne information des épargnants et de définir en quelques mots le risque qu'ils prennent.

Ainsi quatre éléments identifient une sicav : la caractérisation sommaire, la durée minimale de placement recommandée, le type et les objectifs de gestion, la dominante fiscale. Cependant un certain nombre de professionnels dénoncent actuellement les imperfections de la classification en vigueur. Une des principales critiques est relative aux sicav court terme régulières et sensibles qui constituent de fait un fourre-tout très mal défini. Parmi les sicav dites diversifiées en actions et obligations, le choix d'une sicav investie totalement sur la Bourse de Paris plutôt qu'une autre investie à 50 % sur des marchés étrangers, ne fait pas courir le même risque à l'épargnant.

1. Pour une information plus détaillée, voir : Gallais-Hamonno G., *Sicav et Fonds Communs de Placement, les OPCVM en France*, QSJ n° 2654, PUF, 1992.

2. CDSCOPE n° 115, mars 1993.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

1.2. Une nouvelle classification COB

Pour contraindre les gestionnaires à afficher plus clairement leurs objectifs, la COB a initié un projet de classification qui devrait aboutir début 1994 à une nouvelle classification¹.

En conséquence de cette nouvelle grille, les agences d'évaluation de performances vont devoir modifier leurs critères.

Ces agences², de par leur indépendance vis-à-vis des établissements gestionnaires, apportent aux souscripteurs des compléments d'information qui renforcent l'action de la COB. Elles produisent des informations nominatives sur les sicav, et effectuent des classements de sicav sur la base de leurs résultats passés. Ces renseignements donnent aux épargnants des critères de choix selon leurs objectifs : recherche de plus-values, horizon de placement, niveau de risque accepté, etc.

1.3. Mesures des résultats et performances des sicav

Pour apprécier l'évolution du portefeuille d'une sicav sur une année, les financiers définissent une mesure appelée *rentabilité* qui est le taux de croissance de sa valeur liquidative entre les 31 décembre de deux années consécutives $t + 1$ et t . La valeur liquidative au 31 décembre $t + 1$ est majorée de la valeur du dividende versé au cours de l'année $t + 1$ (période de 1 an).

$$R_{t+1,t} = \frac{(V_{t+1} + D_{t+1} - V_t)}{V_t}$$

Pour calculer la rentabilité sur plusieurs années (multipériode), il suffit de composer les rentabilités annuelles intercalaires et d'en calculer la moyenne géométrique :

$$R_{t+n,t} + 1 = \sqrt[n]{(1 + R_{t+1,t})(1 + R_{t+2,t+1}) \dots (1 + R_{t+n,t+n-1})}$$

A partir de cette mesure de rentabilité³, il est possible de comparer des sicav entre elles en se basant sur l'écart type de leurs rentabilités calculé a posteriori, tel est le cas des études menées par Albouy (1982), ou d'analyser leurs performances en les comparant à des indices de référence.

1.4. Comparaison des sicav sur séries chronologiques

Pour comparer les résultats des sicav sur une année, un des critères peut être leur classement selon les valeurs décroissantes des rentabilités. Pour étudier la stabilité des classements sur 2 années, il est possible de construire la matrice de transition

1. « Les nouvelles règles du jeu, une classification rajeunie », les Dossiers du *Figaro Économie*, spécial sicav, 30 novembre 1992, p. 13-18.

2. Les plus connues sont Europerformance déjà citée, qui fait partie de la Caisse des Dépôts et Consignations, Micropal et l'Association des Sociétés et Fonds Français d'Investissement (ASFFI).

3. Cette mesure de rentabilité est souvent dénommée ROR à partir de l'acronyme anglais « Rate Of Return ».

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

croisant les sicav, réparties en un petit nombre de classes de performances, des plus rentables aux moins rentables, et de voir si les classements de l'année $t + 1$ sont équivalents à ceux de l'année t . Dans ce cas la diagonale de la matrice doit être plus lourdement chargée.

Pour un grand nombre d'années, les séries chronologiques des sicav, suite de rentabilités indicées par le temps, deviennent intellectuellement difficiles à comparer si, de plus, le nombre de sicav devient important. Il faut avoir recours à des méthodes statistiques *résumé* comme les études de distribution. Ces études, comme l'ont montré De Drouas, De Vitry et Boulier (1988), aboutissent toujours à la conclusion que l'hypothèse de normalité des distributions ne peut être retenue car les *queues* des distributions sont trop épaisses, c'est-à-dire qu'elles présentent un trop grand nombre de points. De même les indicateurs standard de Skewness – test d'asymétrie de la distribution empirique –, et de Kurtosis – test de l'aplatissement de la distribution empirique par rapport à la loi normale –, rejettent la loi normale.

Il semble que jusqu'à présent la communauté financière n'ait pas encore trouvé de solutions *ad hoc* pour apprécier sur un long horizon les performances des sicav, comme le note Poncet (1990) dans la lettre de l'AFFI :

« Face [...] à l'impossibilité pratique d'établir un classement ordinal exhaustif de tous les OPCVM quelles que soient leurs caractéristiques, il incombe à la communauté financière de rechercher des solutions ».

La difficulté de la tâche vient non de la complexité de la matière mais de l'importance des informations que l'on veut chercher à réduire.

« Comment prendre une option en présence d'une pléthore d'informations qu'on ne parvient pas à lire »¹.

C'est la démarche de l'activité statistique que d'apporter des éléments de réponse.

2. L'activité statistique et l'analyse de données

L'activité statistique en sciences économiques et sociales repose sur une collecte massive d'informations suivie de calculs, pour ne livrer qu'un petit nombre de résultats.

Pour qu'à partir d'un amas d'informations, la Statistique livre des traits saillants, il faut que celle-ci soit sélective, que des tris soient opérés, que les hypothèses liées à la théorie mathématique soient les moins contraignantes possible, et que, de plus, la robustesse et la stabilité des résultats obtenus soient éprouvées. Il n'existe sur le « marché » que peu de méthodes scientifiques capables d'effectuer de tels tris et de

1. Breton T., *La dimension invisible, le défi du temps et de l'information*, Éditions Odile Jacob, 1991.

dégager une structure dans une surabondance d'informations en dehors des méthodes descriptives multidimensionnelles stimulées et développées par l'École française d'analyse de données¹.

Parmi ces méthodes, deux techniques², l'analyse factorielle des correspondances et la classification automatique, sont couramment utilisées de manière complémentaire. La première technique permet de positionner spatialement les objets les uns par rapport aux autres, la seconde cherche à les regrouper, de manière à faire émerger des formes facilitant les interprétations.

Cette complémentarité a fait ses preuves dans beaucoup de domaines de la connaissance³, sciences humaines – linguistique, sociologie –, traitements d'enquêtes, marketing, médecine, sciences naturelles, documentation automatique, principalement en bibliométrie, etc., mais elle est encore peu employée dans le traitement des données chronologiques.

2.1. Séries chronologiques, les approches

Le traitement des séries chronologiques dans un but cognitif ou prévisionnel a donné prise à une importante littérature depuis une vingtaine d'années.

On peut citer chronologiquement quelques ouvrages français complets pour leur époque : *Méthodes du traitement des chroniques* par J.-C. Laloire paru en 1972, *Choix et valeurs des méthodes de prévision* par S.C. Whellwright et S. Makridakis paru en 1974, *Les techniques de la prévision à court terme* par H. Kaufman et J.-C. Groboillot, paru également en 1974, *L'analyse chronologique* par L. Philips, R. Blomme et C. Van den Berghe, paru en 1981, *Séries chronologiques* par J.-J. Drosbeke, B. Fichet et Ph. Tassi, paru en 1989, *Séries temporelles et modèles mathématiques* par C. Gouriéroux et A. Montfort paru en 1990.

On note qu'avec le temps, se complexifie la culture et la formulation mathématique, au prix d'hypothèses de plus en plus contraignantes qui restreignent la portée d'usage de ces méthodes. De plus, le bon sens et le sens critique, occultés par la technique, ne peuvent plus s'exercer.

Si les modèles ARMA et tous les modèles dérivés ne sont plus en vogue, les hypothèses de linéarité et de stationnarité n'étant pas validées en économie et les applications n'ayant pas été à la hauteur des espérances et encore moins des investissements, par contre les modèles stochastiques ont encore de beaux jours devant eux, pour peu qu'ils s'associent aux équations différentielles⁴.

1. Et peut-être la voie prometteuse des réseaux de neurones.

2. Lebart, Morineau et Fénelon, 1979.

3. Pour une vue d'ensemble, voir les *Cahiers de l'Analyse des Données*.

4. « Marchés financiers et gestion de portefeuilles : une mise en perspective des nouveaux outils », *JSSP* 1992 – 4.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

A moins que le chaos¹ ne vienne tout bouleverser, ou que la physique du désordre vienne au secours des désordres boursiers² en faisant appel à la psychologie sociale, lorsqu'elle s'intéresse aux comportements collectifs et phénomènes mimétiques.

Ce qui se dégage de la littérature dominante, que l'on raisonne dans l'espace des fréquences ou dans l'espace du temps, le passage de l'un à l'autre s'effectuant au moyen de la Transformée de Fourier, est que l'objet de base des deux édifices est le concept de **corrélation des séries décalées**.

Avant de chercher à modéliser à partir de ces corrélations décalées, il semble logique d'effectuer des analyses descriptives multidimensionnelles en utilisant les méthodes d'analyses de données afin de découvrir d'éventuelles structures stables.

A. Quaqahez³ et Kh. Aludaat⁴ ont étudié, par l'analyse factorielle des correspondances, des séries chronologiques théoriques décalées, en simulant de vrais mouvements périodiques sinusoïdaux, des polynômes ou encore des produits ou combinaisons de fonctions classiques, afin de dégager des règles d'interprétation. Cette méthode permet d'apprendre *au géomètre à reconnaître de quelles relations mathématiques approchées résultent les formes de graphiques issues de l'analyse de données réelles*.

La démarche d'analyse des séries décalées sur des données réelles, a été prise notamment par Benzecri⁵ ou encore Wong Lai Kuen et S.Y.L. Cheung⁶, avec succès. Cette technique a permis aux auteurs de repérer des mouvements généraux ainsi que des accidents ou périodes atypiques, de révéler des hiérarchies d'effets saisonniers en séparant ces effets des tendances à long terme, ce qui évite de recourir à des opérations préalables douteuses de corrections des variations saisonnières. Cette technique permet également de diagnostiquer des glissements des centres de cycles annuels qui sont interprétables, et tout ceci **sans** nécessité de définition de **modèles a priori**.

Indépendamment de l'analyse des séries décalées, mais dans le même esprit, P. Louvet et O. Taramasco⁷ ont utilisé l'AFC sur des séries chronologiques pour étudier la dépendance sérielle sur des rentabilités journalières à la Bourse de Paris. Sur la base d'un tableau de contingence croisant la rentabilité d'un jour, codée selon les déciles, avec la rentabilité d'une bourse suivante toujours codée selon les déciles,

1. O. Taramasco et I. Girerd-Potin, *Les rentabilités à la Bourse de Paris sont-elles chaotiques ?*, Les entretiens de la finance, AFFI, décembre 1992.

2. A. Orléan, « Les désordres boursiers », *La Recherche*, mai 1991, p. 668-672.

3. A. Quaqahez, « L'Analyse des séries chronologiques décalées. Principes d'interprétation sur des cas modèles », *Cahiers de l'analyse de données*, vol. XII, 1987, n° 4, p. 407-418.

4. Kh. Aludaat et A. Quaqahez, « Corrélation entre séries chronologiques : méthode des séries décalées », *Cahiers de l'analyse de données*, vol. XIII, 1988, n° 4, p. 495-502.

5. J.-P. Benzecri, « L'Analyse des séries chronologiques, Méthode des séries décalées », *Cahiers de l'analyse de données*, vol. XII, 1987, n° 4, p. 289-309.

6. Wong Lai Kuen et S.Y.L. Cheung, « La Bourse de Hong Kong : Analyse des cours en février-mars 1987 par la méthode des séries décalées », *Cahiers de l'analyse de données*, vol. XIV, 1989, n° 2, p. 133-142.

7. P. Louvet et O. Taramasco, *Analyse descriptive de la dépendance sérielle sur les rentabilités journalières à la Bourse de Paris*, décembre 1992, Journées AFFI.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

les auteurs ont pu caractériser un *effet d'entraînement sur la tendance* qui amène les cours à poursuivre les mouvements amorcés la veille et un *effet d'entraînement* sur la volatilité et ceci **sans** avoir besoin de spécifier d'hypothèse sur les **distributions** marginales.

Dans d'autres domaines que celui de la finance, et sur des séries temporelles longues, des tentatives sont faites pour associer analyse spectrale et analyse de données.

Citons pour exemple, Clairambault et Celeux¹ qui ont étudié la variabilité du rythme cardiaque de nouveau-nés en stades de sommeil, à partir de la Transformée de Fourier en 3 bandes de fréquences (haute, moyenne et basse), d'un enregistrement cardiaque. En appliquant des techniques d'analyses de données (analyses en composantes principales et analyse factorielle discriminante), ils ont pu conclure à des comportements physiologiques différenciés selon les stades de prématurité des nouveau-nés.

2.2. Objectif

Après cette présentation du contexte dans lequel nous nous trouvons – une mesure non discutée de la rentabilité mais pas de méthodologie d'analyse définie, l'importance des informations à traiter et une méthodologie de traitement des grandes données maintenant bien rodée – il nous a semblé que les méthodes d'analyse de données pourraient apporter un éclairage nouveau non sur un classement exhaustif de tous les OPCVM, mais sur l'organisation, si elle existe, des OPCVM.

Plus précisément, notre objectif était de chercher à confirmer ou infirmer par les chiffres, que les catégories de sicav étaient gérées de manière homogène par la profession des gérants de sicav, et ce de manière différenciée selon les catégories.

Nous avons déjà utilisé l'analyse des correspondances et l'analyse en composantes principales sur des données chronologiques², et le repérage des années atypiques grâce à la conjonction des méthodes de classification, nous paraissent bien adaptés à cet objectif.

3. Les données et les premiers traitements descriptifs

Pour ses recherches l'IOF³ a constitué une base de données sur les sicav ayant au moins 5 années d'existence. Les informations élémentaires sont les rentabilités

1. J. Clairambault et G. Celeux, « Analyse discriminante appliquée à l'étude du rythme cardiaque », *La revue de MODULAD*, n° 8, décembre 1981, INRIA.

2. M. Le Guen et M. Saint-Marc, *Recherche sur les méthodes économétriques de la datation des cycles : application aux cycles monétaires du XIX^e et XX^e siècles*, colloque Structures économiques et économétrie, Aix-en-Provence, 1984.

3. A l'initiative et sous la direction de G. Gallais-Hamonno, directeur de l'Institut Orléanais de Finance.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

annuelles qui permettent de calculer les rentabilités composées sur différentes sous-périodes de 5 ans, 10 ans, 15 ans, etc. Chaque année la base de données est mise à jour et les nouvelles sicav ayant 5 ans d'existence sont introduites.

Lors des traitements ici rapportés, la base contenait 239 sicav réparties en 12 catégories (cf. tableau 1). Ces catégories ont été définies par G. Gallais-Hammonno sur la base des catégories COB : répartition selon le portefeuille en actions et/ou obligations, selon les marchés français et/ou étrangers. Certaines catégories ont été affinées en cas de spécialisation géographique ou sectorielle. On notera que les sicav monétaires sont totalement exclues.

Tableau 1. Ventilation des 239 sicav selon 12 catégories

N ^{os}	Catégories de sicav	Fréquence	Pourcentage
1	Obligations françaises	39	16,3
2	Obligations françaises de 1 ^{re} catégorie	25	10,5
3	Obligations étrangères ou françaises et étrangères	34	14,2
4	Actions à vocation française	31	13,0
5	Actions à vocation étrangère	58	24,3
6	Spécialisation géographique : Amérique du Nord	8	3,3
7	Spécialisation géographique : Japon	13	5,4
8	Spécialisation géographique : Europe	7	2,9
9	Spécialisation géographique : Régions françaises	4	1,7
10	Spécialisation sectorielle : Or	3	1,3
11	Spécialisation sectorielle : Technologie	3	1,3
12	Spécialisation sectorielle : Immobilier et foncier	14	5,8
Total		239	100,0

L'horizon total disponible porte sur les années 1964 (date de création des premières sicav) à 1991. Au cours de cette période le nombre de sicav évolue chaque année. Certaines sicav disparaissent, d'autres fusionnent, d'autres sont créées. Ainsi en 1971, seulement 43 sicav existent dans la base, tandis que 239 sicav sont disponibles en 1991.

La répartition des sicav selon les catégories évolue aussi au cours des années. En 1991 les sicav obligataires (catégories 1 à 3) sont nettement majoritaires puisqu'elles représentent 41,0 % du stock, les sicav actions 37,2 %. Parmi les sicav spécialisées seules 3 catégories, la n° 6 (Amérique du Nord), la n° 7 (Japon) et la n° 12 (secteur

immobilier et foncier), permettent des comparaisons statistiques, les effectifs des autres sicav spécialisées étant trop faibles (cf. tableau 1).

3.1. Premier aperçu graphique

En analysant (cf. graphique 1) la représentation graphique de l'évolution des rentabilités moyennes¹ de 7 catégories les plus importantes en effectif, au cours de la période du 31/12/71 au 31/12/91, soit 20 années, on note :

- un très net mouvement d'ensemble, avec des chutes de rentabilité importantes, ponctuelles et systématiques en 1974, 1981, 1987 et 1990, suivi d'une remontée de taux équivalent l'année suivante ;
- les amplitudes ou volatilités sont plus faibles pour les sicav obligataires que pour les sicav actions ;
- les évolutions des 3 catégories obligataires sont très similaires.

Il y a bien une hétérogénéité dans les résultats des catégories représentées par leurs centres de gravité, mais est-ce que les catégories sont homogènes ?

3.2. Inadaptation de l'analyse de la variance

Pour comparer les catégories entre elles on peut comparer la moyenne de chaque catégorie par l'intermédiaire de l'analyse de la variance, et étudier les distributions empiriques calculées sur une même période.

Au vu du graphique 2 représentant les distributions des ROR par catégories sur 20 ans – du 31/12/71 au 31/12/91 – on remarque que les variances de chaque catégorie sont très éloignées les unes des autres. L'hypothèse d'égalité requise en Anova n'est donc pas satisfaite. En conséquence, l'écart type sur les moyennes (0,033 80) n'a aucune valeur.

Cette représentation graphique permet ainsi d'éviter d'effectuer d'autres calculs statistiques sans pertinence.

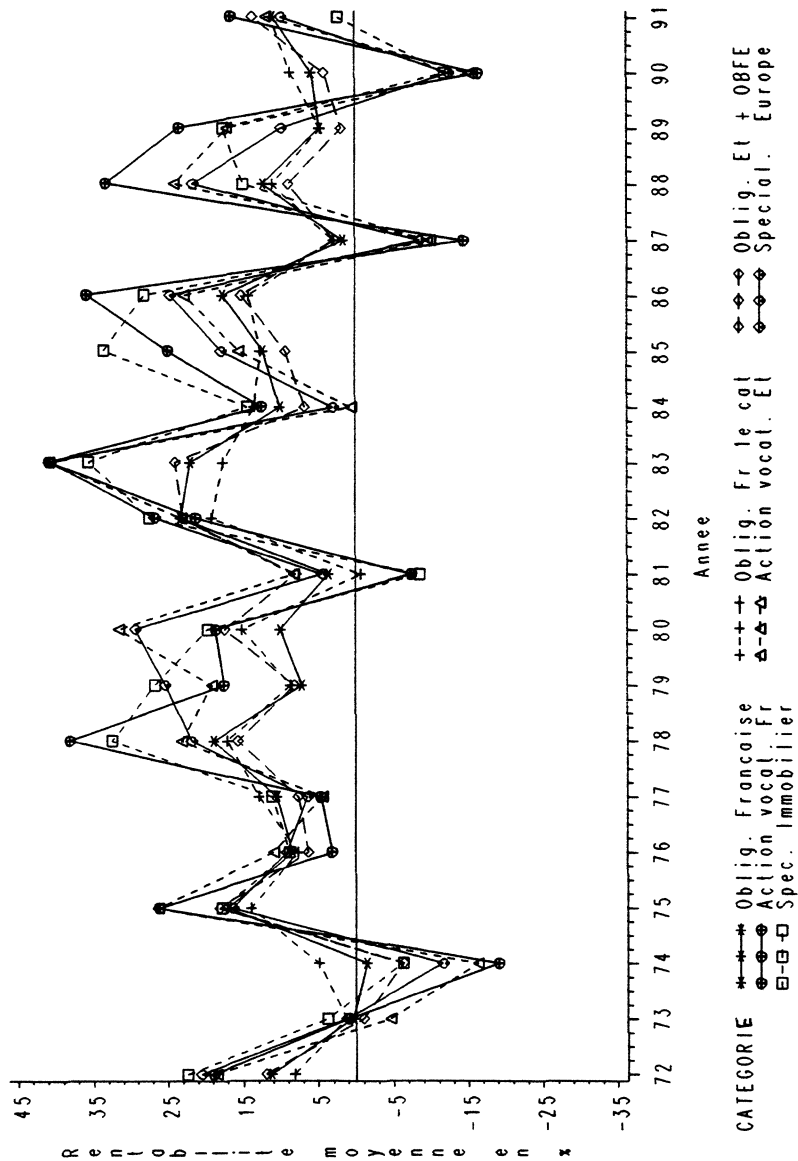
L'analyse des séries chronologiques de chaque sicav vue sous l'angle de l'analyse de la variance n'étant pas adaptée eu égard aux hypothèses requises, nous nous sommes tournés vers une étude des rangs calculés à partir des performances. Et, afin d'avoir une idée de la robustesse, au sens de la stabilité des résultats, nous avons été amené à étudier plusieurs historiques glissant sur les données.

1. La rentabilité moyenne d'une catégorie est la moyenne arithmétique calculée à partir des sicav constituant cette catégorie. Cette rentabilité représente la rentabilité d'un portefeuille fictif qui serait constitué de l'ensemble de ces sicav sans faire intervenir de pondération.

RENTABILITES ANNUELLES PAR CATEGORIES DE SICAV

Horizon 31/12/71 AU 31/12/91 soit 20 ANS

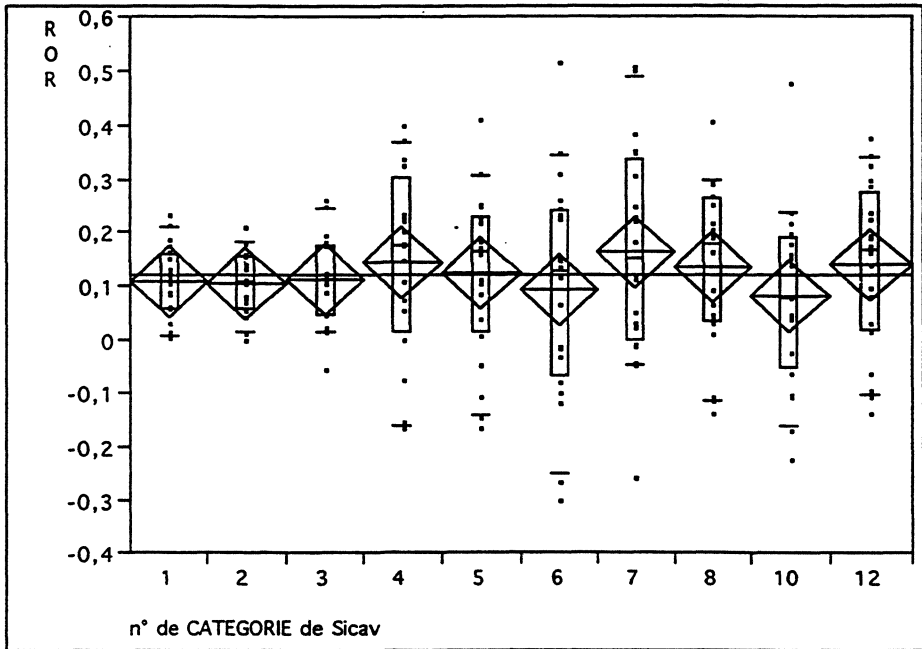
pour 7 CATEGORIES de SICAV



Graphique 1

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

Graphique 2. Analyse des ROR sur 20 ans



Signification des représentations visuelles des distributions

Ici on apprécie visuellement les différences entre les catégories, selon les différences de forme des losanges et des rectangles :

Losange

(Dans le cas particulier du graphique 2, les losanges sont proches d'un carré).

La diagonale horizontale d'un losange est proportionnelle à la **taille** de la catégorie, ici les effectifs sont les mêmes : 20 années.

La diagonale horizontale de chaque losange a pour ordonnée la **moyenne** de la catégorie correspondante ; celle-ci est à comparer à la moyenne générale notée en pointillé.

La diagonale verticale de chaque losange symbolise l'intervalle de confiance à 95 % de la moyenne (calculs issus de l'analyse qui suppose l'égalité des variances des groupes !)

Rectangle

Le rectangle symbolise les quartiles 25 % et 75 %.

La barre horizontale du rectangle symbolise la médiane (50 %).

Les barres à l'extérieur du rectangle symbolise les fractiles 10 % et 90 %.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

3.3. Caractéristiques des sous-périodes étudiées

En plus des périodes annuelles et des périodes quinquennales, cinq sous-périodes de 8 à 20 ans ont été choisies, comme le montre le tableau 2.

Tableau 2. Numéro d'étude et sous-période correspondante, nombre de sicav concernées

N° d'étude	Sous-périodes L1xx - L1yy : rentabilité annuelle du 31/12/xx au 31/12/yy + 1 L5xx - L5yy : rentabilité quinquennale du 31/12/xx au 31/12/yy + 5	Nombre d'années	Nombre de sicav	Nombre de catégories disponibles
1	L171 - L190	20	43	7
2	L181 - L190	10	101	10
3	L183 - L190	8	121	11
4	L171 - L180	10	49	8
5	L577 - L586	10 pts sur 14 ans	70	10
<p>Note L1xx : correspond à la rentabilité annuelle calculée à partir du 31/12/xx, ce qui correspond à la rentabilité de l'année xx + 1. L5xx : correspond à la rentabilité quinquennale calculée à partir du 31/12/xx, ce qui correspond à la rentabilité sur les 5 années xx + 1 à xx + 5.</p>				

Pour les 5 études, on peut repérer sur le tableau 2, « Numéro d'étude et sous-période correspondante », le nombre d'observations qui correspond au nombre d'années. Les tableaux à analyser sont des rentabilités annuelles et quinquennales pour l'étude n° 5. Les rentabilités quinquennales sont calculées selon les taux composés, c'est-à-dire en tenant compte du réinvestissement du dividende.

Ces 5 sous-périodes (soit 5 études) ont été choisies de manière à couvrir des horizons assez longs compatibles avec le nombre de sicav disponibles. Sur toute ces sous-périodes les sicav entièrement renseignées ont été sélectionnées. Le nombre est donc variable, et évolue entre 43 sicav pour l'horizon le plus long et 121 sicav pour le plus court et le plus récent, pour un stock potentiel de 239 sicav dans la base.

3.4. Études selon les rangs

Pour chaque sous-période, et pour chaque année, est calculé d'abord le rang de chaque sicav à partir des tableaux de rentabilités des sicav individuelles, puis un rang moyen sur la sous-période.

Ces classements sont ensuite regroupés par catégorie, afin de calculer la moyenne des rangs des sicav composant chaque catégorie.

La sicav qui a le meilleur ROR sur l'année considérée prend le rang 1. Le rang maximum correspond à l'effectif des sicav de la sous-période.

Ces tableaux montrent l'hétérogénéité au cours du temps. Si on retient les sicav qui ont la meilleure performance (rang = 1), leurs performances l'année précédente et/ou l'année suivante sont très souvent dégradées.

Ainsi dans l'étude n° 1, sur la période 1971-1991, la sicav *9536 Épargne Revenu*, appartenant à la catégorie 1 (Obligations françaises) est au rang 1 en 1987, encadrée par les rangs 39 en 1986 et 40 en 1988. Ce qui correspond aux rangs des moins bonnes performances (le maximum étant 43).

Durant cette même période, *9789 Épargne Institution*, appartenant à la catégorie 2 (Obligations françaises de 1^{re} catégorie) est au rang 1 en 1984 encadrée par les rangs 15 en 1983 et 43 en 1985. Ce dernier chiffre correspond également à la moins bonne performance.

Citons encore *9511 Sliva France* appartenant à la catégorie 4 (Actions à vocation française), au rang 1 en 1978, encadrée par les rangs 30 en 1977 et 14 en 1979.

Le résumé des rangs moyens par catégorie (voir le tableau 3) montre à l'évidence que les performances fluctuent dans le temps, puisque ces rangs sont tous très élevés.

Ces résultats corroborent ceux obtenus par Albouy en 1982 sur un échantillon de 50 sicav, et qui concluait à *l'instabilité* des classements des sicav. Il est clair que **les performances d'une année ne sont pas un gage pour l'avenir.**

Remarque : tout lecteur intéressé peut demander au secrétariat de l'IOF (Faculté de Droit, BP 6739, 45067 Orléans cedex 2 - Fax : 33 - 38 41 73 80) le document de même titre 1-94-1-MIF qui contient 38 pages d'annexe chiffrée : les rangs, les AFC en couleur, les dendrogrammes et les croisements entre catégories d'origine et groupes d'affectation.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

Tableau 3. Rang moyen des catégories de sicav par sous-période

Sous-période	Rang maximum	N° de catégorie	Rang moyen <i>critère de classement</i>
L171 - L190	43	12-S Immobilier et foncier 4-A Vocation française 8-S Europe 5-A Vocation étrangère 1-O Françaises 2-O Françaises de 1 ^{re} catégorie 3-O Étrangères ou franç. & étrang.	18 19 20 21 25 26 26
L181 - L190	101	4-A Vocation française 12-S Immobilier et foncier 8-S Europe 5-A Vocation étrangère 1-O Françaises 2-O Françaises de 1 ^{re} catégorie 3-O Étrangères ou franç. & étrang. 7-S Japon 6-S Amérique du Nord 10-S Or	40 45 49 52 54 57 58 58 69 70
L183 - L190	121	4-A Vocation française 9-S Régions françaises 12-S Immobilier et foncier 1-O Françaises 2-O Françaises de 1 ^{re} catégorie 8-S Europe 5-A Vocation étrangère 3-O Étrangères ou franç. & étrang. 7-S Japon 10-S Or 6-S Amérique du Nord	44 51 58 60 62 64 65 70 77 85 95
L171 - L180	49	12-S Immobilier et foncier 8-S Europe 5-A Vocation étrangère 4-A Vocation française 6-S Amérique du Nord 1-O Françaises 3-O Étrangères ou franç. & étrang. 2-O Françaises de 1 ^{re} catégorie	20 22 24 26 28 30 30 30
L577 - L586	70	4-A Vocation française 7-S Japon 12-S Immobilier et foncier 8-S Europe 5-A Vocation étrangère 1-O Françaises 2-O Françaises de 1 ^{re} catégorie 3-O Étrangères ou franç. & étrang. 6-S Amérique du Nord 10-S Or	22 23 26 29 33 43 45 48 49 61
<p>Note Les catégories sont classées selon le rang moyen. Les rangs des sicav obligataires sont soulignés. A : Actions, O : Obligations, S : Spécialisées.</p>			

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

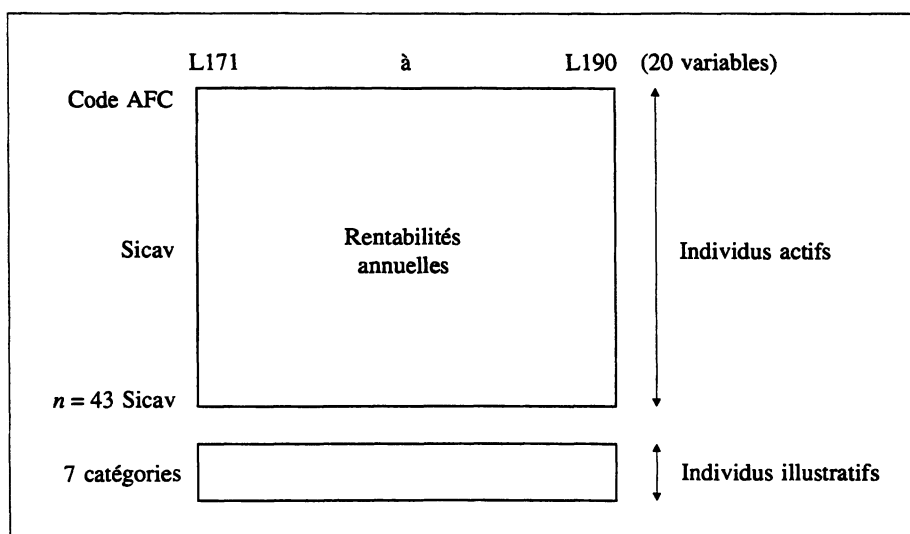
Au vu de ces approches variées, graphiques, analyse de variance, études des rangs, nous avons abouti au constat qu'il devait exister dans les données une organisation, mais que les techniques statistiques classiques ne permettaient pas de la mettre en évidence.

4. Les rentabilités des sicav vues sous l'angle de l'analyse de données

4.1. Constitution des tableaux de données, observations * variables

Dans les tables soumises aux analyses, les *observations* sont les sicav repérées par leur code AFC¹), et les variables actives sont les rentabilités sur l'horizon choisi.

Tableau 4. Table des données pour l'étude n° 1



Le nom des variables (limité à 4 caractères par les contraintes logicielles) permet de repérer le type de rentabilité retenu, L1 pour les rentabilités annuelles et L5 pour les rentabilités quinquennales. Les 2 derniers caractères du nom font référence au 31/12 de l'année où débutent les calculs de rentabilité. Ainsi L171 correspond à la rentabilité de l'année 72.

1. Attention à l'homonymie entre code AFC qui est la norme de codification des sicav, et l'Analyse Factorielle des Correspondances qui a le même acronyme.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

A chaque table sont ajoutées en ligne des observations illustratives calculées pour chaque catégorie sous forme de centre de gravité. Ces observations permettent de trouver un représentant typique fictif de chaque catégorie et d'aider ainsi à l'interprétation.¹

4.2. Les analyses factorielles

Les données étant des variables continues et la notion de corrélation entre deux séries ayant un sens cognitif important, il nous a paru nécessaire d'utiliser d'abord une ACP afin d'avoir une vue globale sur les liaisons entre variables, puis une AFC pour ses facilités d'interprétation conjointes des deux ensembles, I pour les individus et J pour les variables.

Le tableau 5 donne le résumé des valeurs propres pour l'ensemble des 5 ACP et 5 AFC qui ont été réalisées. Pour toutes les analyses on note la très forte importance des variances extraites dès le 1^{er} axe. Le cumul sur les 3 premiers axes est toujours supérieur à 70 % et atteint même 94,43 % pour les AFC5.

Il y a une très forte organisation dans les données, c'est-à-dire une structure forte qui doit apparaître sur les graphiques factoriels.

Tableau 5. Valeurs propres (en %)

Étude	Nombre de variables	% 1 ^{er} axe	% 2 ^e axe	% 3 ^e axe	Cumul sur 3 axes
ACP1	20	36,33	22,51	11,88	70,73
AFC1	20	43,97	21,52	12,36	77,85
ACP2	10	48,68	16,06	11,86	76,60
AFC2	10	58,00	17,73	6,71	82,43
ACP3	8	53,22	18,52	13,78	85,52
AFC3	8	64,08	16,20	6,90	87,18
ACP4	10	37,08	24,76	18,68	80,53
AFC4	10	44,38	25,35	17,55	87,28
ACP5	10	45,70	39,65	7,59	92,95
AFC5	10	65,85	23,81	4,76	94,43

Pour les interprétations, nous nous limiterons à une ou deux études pour éviter le côté fastidieux des répétitions.

1. Pour les AFC qui nécessitent un tableau de mesures positives, les rentabilités ont été majorées de 1 (ROR → 1 + ROR), pour assurer cette contrainte de positivité.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

4.2.1. Analyse des variables sur le premier plan factoriel

Sur le graphique 1 des rentabilités annuelles des catégories, nous avons déjà remarqué que les catégories de sicav ont des profils semblables avec des chutes et des remontées en coïncidence. Ainsi les années de chute avec des rentabilités très négatives : 1974, 81, 87 et 90 (sauf pour les sicav obligataires), sont suivies de remontées spectaculaires en 1975, 82, 88 et 91 avec des rentabilités positives.

Les rentabilités des sicav obligataires sont beaucoup moins volatiles. Elles sont selon les années au-dessus des autres catégories comme en 1974, 87 et 90 ou en dessous comme en 1978, 79, 83, 85, 86, 88 et 89, ou encore intermédiaires entre les moins rentables et les plus rentables.

Sur la figure 1, « Plan factoriel F1 * F2 de l'ACP1 – projection des variables », on retrouve ces mêmes oppositions mais cette fois-ci en tenant compte de l'ensemble des sicav prises dans leurs résultats individuels et en projetant sur ce plan factoriel les centres de gravité des catégories, pour faciliter les interprétations.

Sur F1, par identification des variables principales (années) on repère successivement L189 (cor = 0,851), L186 (cor = 0,757) et L173 (cor = 0,650) qui s'opposent à L182 (cor = 0,824), L187 (cor = 0,708) et L171 (cor = 0,584).

Sur le tableau 6, « Moyennes et écarts types des rentabilités annuelles toutes catégories confondues », on note que L173, L186 et L189 correspondent aux années de *moins bonnes performances* (F1 < 0). Ces années sont de plus à rentabilités négatives et elles s'opposent (F1 > 0) aux années de très *fortes* rentabilités moyennes : L171, L182, L187 pour les mieux représentées sur l'axe 1, puis L174, L177, L179 et L185.

Mais c'est en terme de corrélations que les interprétations sont plus éloquentes. Sur ce plan factoriel sont matérialisées les corrélations les plus élevées en valeur absolue pour montrer les analogies dans les dynamiques des années, ainsi que les liaisons fortes entre années consécutives, c'est-à-dire les corrélations importantes proches de la diagonale principale. Ces dernières mettent particulièrement en évidence les *inversions de rangs* d'une année sur l'autre (corrélation négative).

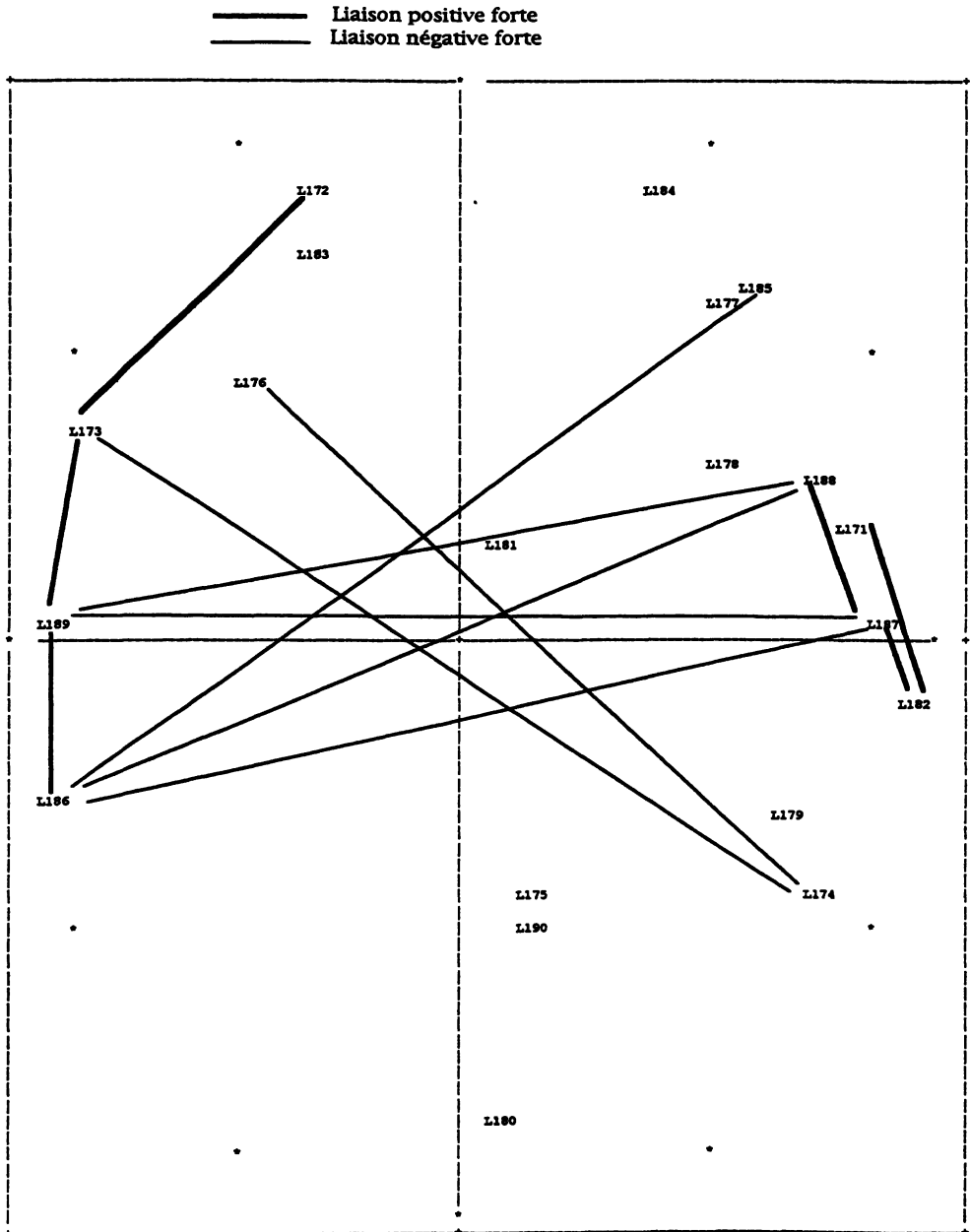
Les corrélations positives sont représentées par un trait épais et les corrélations négatives par un trait plus fin.

La dynamique des rentabilités en 90 (L189) est très corrélée à celle de 87 (L186), dans ses interrelations avec l'année antérieure et l'année postérieure. Les performances des sicav se sont inversées de l'année $n - 1$ à l'année n , puis de nouveau de n à $n + 1$. C'est dire que dans l'ensemble les sicav performantes sont devenues les moins performantes et *vice versa*.

En utilisant le logiciel JMP qui combine des analyses statistiques à des représentations visuelles, il est plus facile de voir ce qui caractérise ces années.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

Figure 1. Plan factoriel F1 * F2 de l'ACP 1 - projection des variables



SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

En L189, L186 et L173 les rentabilités des sicav obligataires sont toutes supérieures à celles des autres sicav, cf. figure 2 : « Visualisation des rentabilités pour l'année 1990 (L189) ».

En L179, L182 et L188 le phénomène est inverse, les sicav obligataires sont systématiquement moins performantes que les autres sicav, cf. figure 3 : « Visualisation des rentabilités pour l'année 1983 (L182) ».

Sur F2 l'opposition vient de L180 (cor = 0,680), avec L172 (cor = 0,629) et L184 (cor = 0,611). Les rentabilités de 1981 (L180) qui déterminent l'axe 2 montrent un comportement de l'ensemble des sicav très hétérogène.

Tableau 6. Moyenne et écart-type des rentabilités annuelles toutes catégories confondues

Variable	N	Moyenne	Écart type
L171	43	0,169	0,046
L172	43	- 0,018	0,044
L173	43	- 0,112	0,082
L174	43	0,223	0,062
L175	43	0,090	0,050
L176	43	0,067	0,052
L177	43	0,236	0,076
L178	43	0,167	0,098
L179	43	0,240	0,110
L180	43	0,039	0,087
L181	43	0,232	0,052
L182	43	0,348	0,096
L183	43	0,054	0,076
L184	43	0,170	0,090
L185	43	0,224	0,080
L186	43	- 0,069	0,070
L187	43	0,198	0,080
L188	43	0,135	0,095
L189	43	- 0,091	0,101
L190	43	0,109	0,049

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

Figure 2. Visualisation des rentabilités en 1990 (L189) selon les catégories

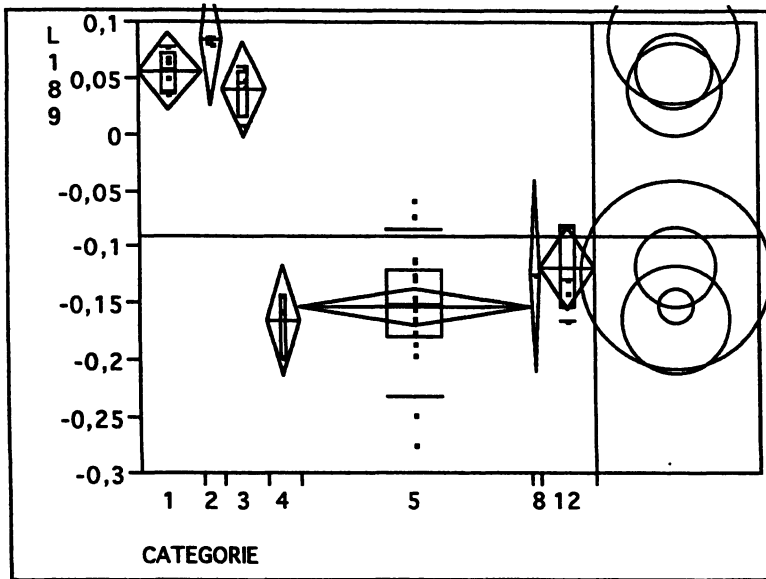
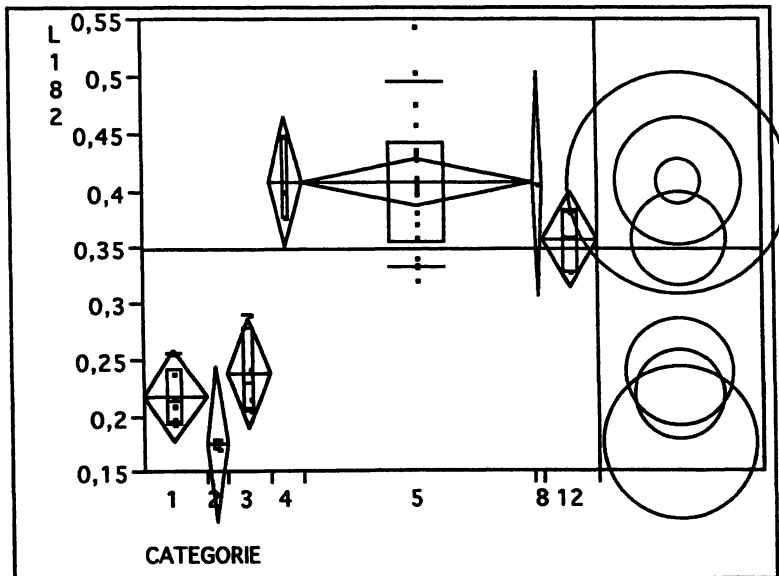


Figure 3. Visualisation des rentabilités en 1983 (L182) selon les catégories



SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

4.2.2. Analyse des observations sur le premier plan factoriel

Si nous projetons sur ce plan factoriel (1,2) les sicav repérées par leur catégorie, ainsi que les centres de gravité, on fait apparaître une remarquable discrimination (cf. figure 3).

L'axe 1 oppose les sicav obligataires ($F1 < 0$) toutes regroupées avec, là encore, une distinction entre les obligations françaises (1), les obligations françaises 1^{re} catégorie (2) et les obligations étrangères ou françaises et étrangères (3), beaucoup plus faible mais néanmoins visible.

A l'opposé, on repère la catégorie 5, « Actions à vocation étrangère », majoritairement regroupées dans le cadran $F1 > 0, F2 < 0$.

Sur l'axe 2 la catégorie des sicav immobilières (12) se sépare nettement des autres catégories et se rapproche de la catégorie 4, « Actions à vocation française ».

Dans les analyses ACP1 ou AFC1, les résultats sont très analogues et les corrélations des centres de gravités sont remarquablement liées aux axes¹ (voir tableau 9).

Sur $F1$ on trouve *C1 (cor = 0,976), *C2 (cor = 0,920), *C3 (cor = 0,855), *C5 (cor = 0,692)².

Sur $F2$ c'est principalement *C12 (cor = 0,822) qui est bien corrélé.

Sur $F3$ apparaît l'unique sicav de la catégorie 8 – 9677 *Euro-Croissance* – à spécialisation géographique tournée vers l'Europe qui est mal représentée sur les 2 premiers axes.

Tableau 9. Tableau des coordonnées factorielles des centres de gravité des catégories avec les contributions relatives du facteur à un élément (COR) et les contributions relatives de l'élément au facteur (CTR = 0)

ISUP	QLT	POID	INR	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR
*C1	989	23	19	-3 954	976	0	88	0	0	443	12	0	51	0	0
*C2	944	23	32	-5 034	920	0	511	9	0	304	3	0	-546	11	0
*C3	961	23	18	-3 624	855	0	-1 149	86	0	401	10	0	381	9	0
*C4	896	23	25	2 755	359	0	2 336	258	0	2 283	247	0	814	31	0
*C5	995	23	5	1 721	692	0	-1 107	286	0	-218	11	0	-143	5	0
*C8	494	23	5	614	85	0	530	63	0	-1 237	343	0	110	3	0
*C12	894	23	22	310	5	0	3 971	822	0	-1 135	67	0	-27	0	0

Dans toute cette homogénéité on note un anachronisme pour la sicav 9740 *Via Investissement* qui a une très forte inertie et n'apparaît que sur le 4^e axe. Sur le plan $F1 * F2$, *Via Investissement* se trouve proche des C4.

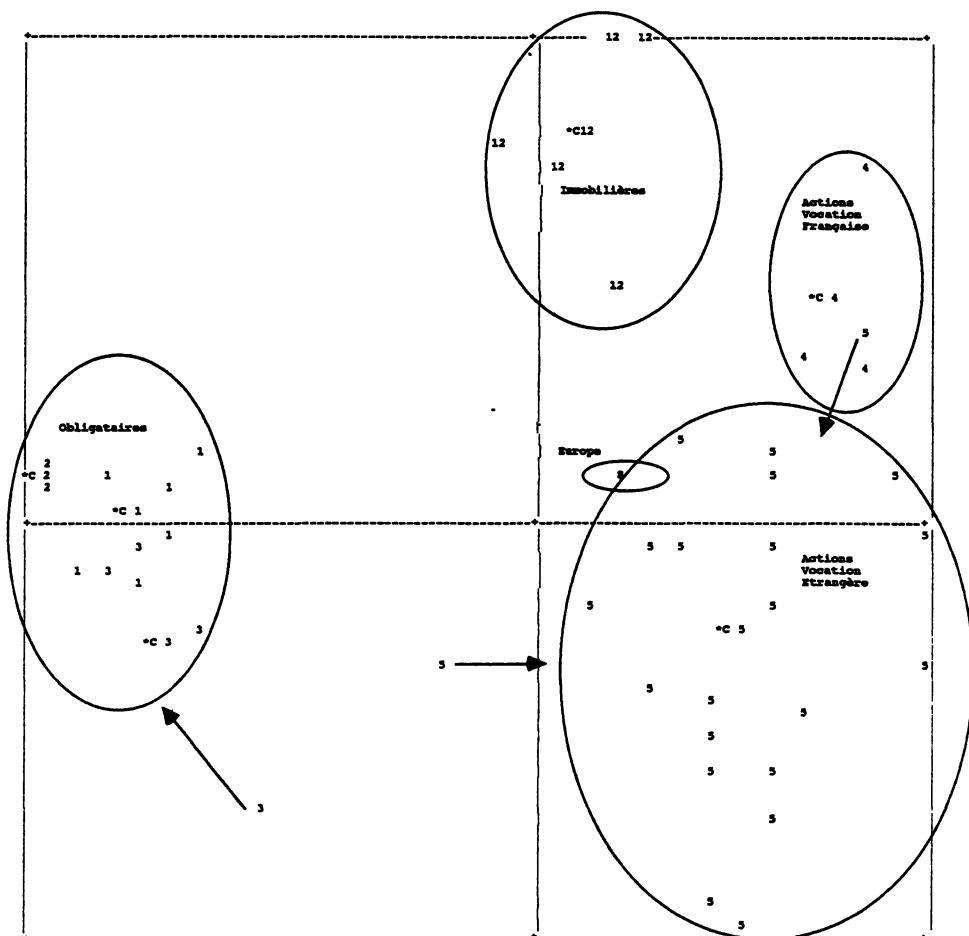
1. On rappelle que ces centres de gravité ne participent pas à la détermination des axes factoriels, ils n'interviennent qu'en éléments illustratifs.

2. Cor représente le carré du cosinus de l'angle entre la direction du point dans l'espace d'origine et l'axe factoriel. Cor est donc compris entre 0 et 1. On appelle aussi cet indicateur : contribution relative du facteur à l'élément.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

La partition en catégories définie *a priori* est bien une réalité *a posteriori* du comportement des rentabilités tout au long de ces 20 années.

Figure 4. Plan factoriel F1 * F2 de l'ACP1 – projection des catégories



Pour les autres ACP et AFC des autres études, on retrouve systématiquement les mêmes discriminations et les mêmes caractéristiques.

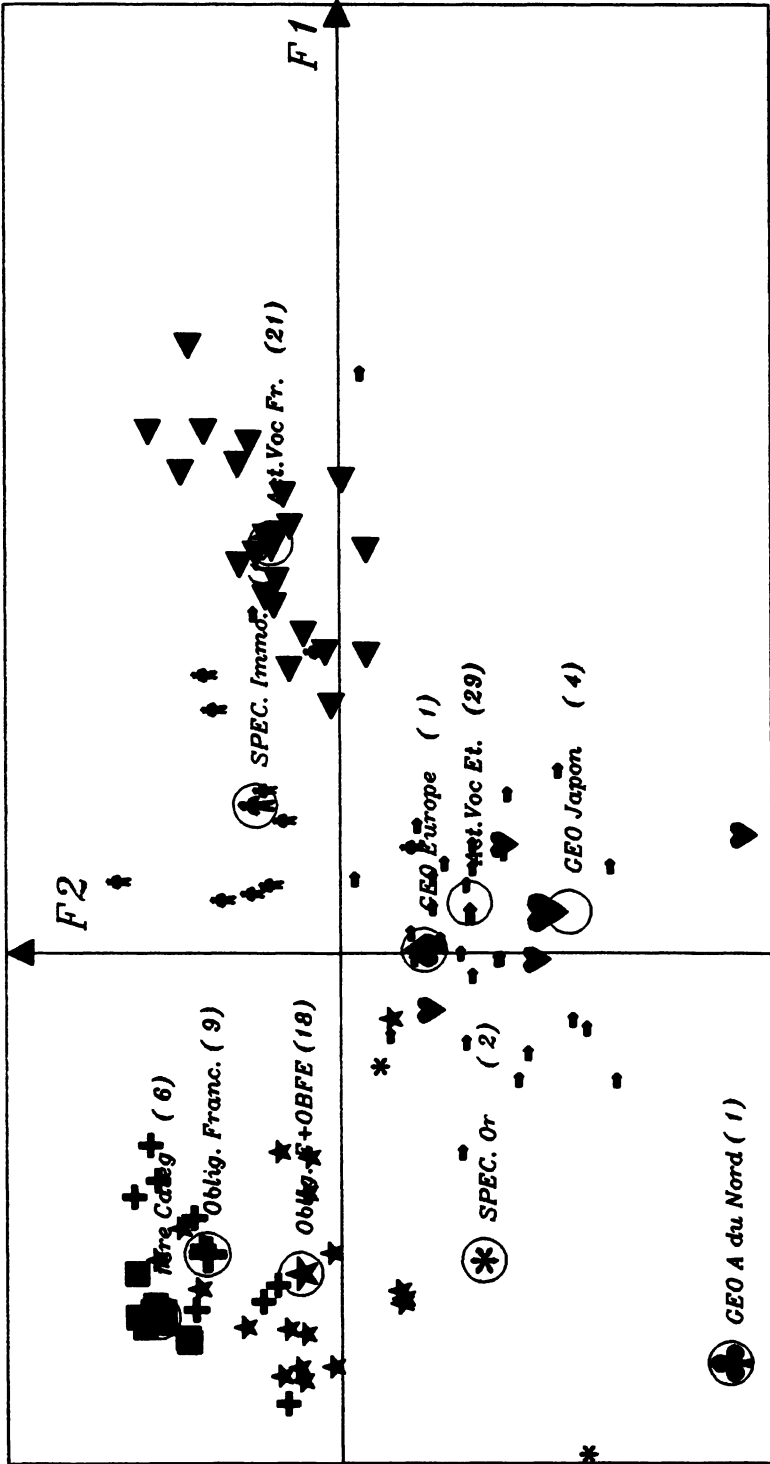
Mentionnons pour l'étude 2 (cf. graphique 2), qui porte sur 10 années et 101 sicav, une particularité qui vient de la catégorie 7 (sicav spécialisées sur le Japon). L'axe 2 montre l'écrasante domination en 83 de cette catégorie avec une rentabilité moyenne

SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse en Composantes Principales

101 Sicav – Horizon 31/12/81 – 31/12/91 soit 10 années

champ



Note: Les centres de gravité des catégories sont entourés d'un cercle suivi du nom et de l'effectif de la catégorie ()

Graphique 2

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

de 52,76 %. Les records sont atteints par 9547 *Gestion UniJapon* avec une rentabilité de 66,94 %, suivi de 9557 *Aurore* avec une rentabilité de 53,94 %, puis 9622 *Laffitte Tokyo* avec une rentabilité de 53,4 %.

Malheureusement, cette année record est suivie par deux années de très mauvaises performances pour ces mêmes sicav (en moyenne - 0,69 %). *Gestion UniJapon* descend à - 6,92 % en 84, puis - 1,85 % en 85, *Aurore* dans une moindre mesure descend à 1,16 %, puis 5,30 % et *Laffitte Tokyo* ne fait guère mieux avec 9,13 % et 1,13 %.

Dans le cadran $F1 < 0 * F2 < 0$, la seule sicav spécialisée sur l'Amérique du Nord - 9556 *Gestion Amérique* - apparaît isolée par rapport aux nuages des points. Elle détermine l'axe 4. Son originalité est due à sa chute de rentabilité en 86 (L185). C'est sur cette année la seule sicav à rentabilité négative.

Ici comme pour toutes les analyses et plans factoriels $F1 * F2$, la sicav 9740 *Via Investissement* a un comportement plus proche des sicav de la catégorie C4, sicav actions à vocation française, que de sa catégorie C5, sicav actions à vocation étrangère.

4.2.3. Plans factoriels $F1 * F2$ pour les 5 études sur les rentabilités

Afin de mettre en évidence visuelle cette homogénéité des catégories dans toutes les coupes temporelles, nous avons représenté sur les figures 1-5 portées en annexe, l'ensemble des premiers plans $F1 * F2$ d'analyses factorielles des correspondances.

Chaque sicav est projetée sur le plan factoriel $F1 * F2$, et est reliée au centre de gravité de sa catégorie. On fait ainsi apparaître des constellations représentatives des catégories.

Quelle que soit la longueur des horizons retenus, les catégories sont toujours bien discriminées sur le 1^{er} plan. Les sicav obligataires sont toujours bien regroupées dans un même cadran et s'opposent aux sicav actions françaises bien séparées des sicav actions étrangères. Les sicav spécialisées sont également discriminées sur le premier plan.

Quels que soient les horizons, la seule sicav spécialisée sur l'Amérique du Nord apparaît systématiquement à l'écart du nuage des autres sicav.

Voyons ce que peut apporter une classification ascendante hiérarchique sur les mêmes ensembles de données et comparons la classification a priori avec une classification a posteriori basée sur les résultats (rentabilités).

4.3. Classification a posteriori

Dans le couplage classique AFC et classification, la construction des classes s'effectue souvent à partir des coordonnées factorielles obtenues par une analyse préalable. Ici nous avons effectué les classifications sur les mêmes tableaux de données définis pour les études n^{os} 1 à 5.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

La construction de l'arbre d'agrégation hiérarchique ou dendrogramme est réalisée en utilisant le critère d'agrégation de Ward. Ce critère correspond à une réduction minimale de variance lors de l'agrégation, il est d'autre part homogène au critère d'inertie utilisé pour la détermination des axes factoriels. L'arbre est construit avec l'algorithme rapide de recherche des voisins réductibles¹.

« *Les zones denses de points homogènes s'agrègent au niveau bas de l'arbre et se séparent en classes naturelles quand les branches de raccordement s'allongent.* »²

Pour définir une partition on coupe l'arbre selon divers critères qui sont bien souvent liés à la facilité de l'interprétation des classes retenues et à leur homogénéité. En pratique, il est rare qu'une seule partition s'impose.

L'arbre hiérarchique de l'étude n° 1, confirme les résultats obtenus sur les analyses factorielles, voir figure 5 : dendrogramme³.

Au plus bas de l'arbre se trouve regroupées les 3 catégories obligataires puis les sicav composant la catégorie 4 des sicav actions à vocation française.

On retrouve *Via Investissement* et *UAP Actions Sélectionnées* classées avec la catégorie 4 et non la 5. Quatre des sicav immobilières sont également bien regroupées et très séparées des autres sicav puisqu'elles ne se regroupent que dans une partition en 3 classes.

La sicav immobilière *IMSI Gestion Immo* se trouve agrégée avec l'unique sicav européenne, *Euro-Croissance*, proche de *Victoire*, et bien centrée au milieu de l'ensemble des sicav actions à vocation étrangère.

Cette sicav immobilière est la seule à avoir un classement « technique », bien éloigné de sa catégorie d'origine. Son profil des rentabilités est plus proche des sicav actions à vocation étrangère que de celui des immobilières.

La partition en 4 classes fait aussi ressortir les 4 classes obtenues sur le plan factoriel F1 * F2. On repère successivement, les sicav obligataires les plus rapprochées et les plus distantes des autres sicav, puis le groupe des sicav actions étrangères, qui, bien que plus nombreuses en effectif, ont des profils sur les 20 années très similaires. Les 2 dernières catégories, sicav immobilières et sicav actions à vocation française, d'effectif faible, se différencient cependant des deux autres classes.

Si on retient une partition en 4 classes et en croisant la catégorie d'origine et le numéro d'affectation (cf. tableau 10) on remarque que :

- Les 3 catégories des obligataires ne se différencient pas, toutes ces sicav se retrouvent dans une même classe.
- Les 3 sicav actions à vocation française forment bien une classe à part.

1. Les programmes utilisés sont ceux de la bibliothèque ADDAD en liaison avec le progiciel SAS via la procédure ADDADSAS de M.-O. Lebeaux (CNRS) et P.-O. Flavigny (CNRS), installée au CIRCE.

2. A. Morineau, « SPAD.N, logiciel pour l'analyse statistique de données », *Revue de MODULAD INRIA*, mars 1989, n° 6, p. 27-59.

3. Les autres arbres hiérarchiques ne sont pas reproduits ici, cf. note 1, page 50.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

- Les 22 sicav actions à vocation française et étrangère sont majoritairement bien regroupées dans la classe 2 ; seules 2 d'entre elles, à savoir *Via Investissement* et *UAP Actions Sélectionnées*, déjà mentionnées, sont regroupées avec les actions à vocation française.
- Les sicav immobilières forment la classe 3.

Ainsi l'adéquation entre les 2 classements donne 90,7 % des sicav bien classées (39/43).

Tableau 8. Croisement entre la catégorie d'origine et le numéro d'affectation issu de la classification

Étude n° 1	Catégorie	Partition en 4 classes				Total
	Fréquence	1	2	3	4	
	1. Oblig. françaises	6	0	0	0	6
	2. Oblig. 1 ^{re} catégorie	2	0	0	0	2
	3. Oblig. fr. et étr.	4	0	0	0	4
	4. Act. voc. française	0	0	0	3	3
	5. Act. voc. fr./étr.	0	20	0	2	22
	8. Géo. Europe	0	1	0	0	1
	12. Spéc. immobilière	0	1	4	0	5
	Total	12	22	4	5	43

Si on reprend ces mêmes tableaux de croisement pour l'ensemble des autres études, on trouve des pourcentages de bien classés très importants, compris entre 73 % et 90,7 %.

Tableau 9. Pourcentages de bien classés selon la partition retenue

N° Étude	Partition en 4 classes	Partition en 5 classes
1	90,7 %	*
2	87,1 %	85,1 %
3	85,1 %	87,6 %
4	73,5 %	73,5 %
5	82,9 %	84,3 %

* La partition en 5 classes conduit à séparer en deux classes d'effectifs égaux la classe C5.

Dès les partitions en 4 classes, on peut s'apercevoir que la catégorie C4, actions à vocation française, et la catégorie C12, sicav immobilières, bien que proches sur le 1^{er} plan factoriel, sont bien 2 classes **différenciées**. Seule l'étude n° 5 sur les rentabilités quinquennales ne peut les différencier, même dans une partition en 5 classes.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

En comparant les arbres hiérarchiques sur les 5 études, le fait marquant concerne la catégorie 4 « actions à vocation française ».

Dans la sous-période la plus éloignée dans le temps (étude n° 4, 1971 à 1981), les sicav de cette catégorie sont peu nombreuses ($n = 4$) et ne forment pas une catégorie homogène. Elles se rapprochent selon les cas vers les sicav immobilières comme pour *9511 SlivaFrance* ou vers les sicav à vocation étrangère, pour *9544 Essor*, *9509 Optimavalor* et *9543 Intersélection France*.

Dans les sous-périodes plus rapprochées, leur nombre augmente jusqu'à $n = 21$ pour la sous-période du 31/12/81 au 31/12/91, et $n = 24$ pour la sous-période du 31/12/83 au 31/12/91. On remarque que sur les sous-périodes plus récentes cette catégorie s'homogénéise et vient alors à se différencier nettement des autres catégories.

Pour la période la plus récente, on trouve au bas de l'arbre toutes les sicav de la catégorie 4 **très bien regroupées**, seule *9627 Laffitte France* dans un éloignement peu important se rapproche cependant de la catégorie 12.

L'homogénéité des catégories est remarquable.

Au vu de ces résultats, il est incontestable que les catégories de sicav sont homogènes. Il y a là, la découverte d'une **cohérence des gestionnaires de sicav**.

Si cette homogénéité est indéniable, il y a cependant des anomalies. On va donc s'intéresser par la suite à identifier les sicav dont le comportement diffère de celui de leur catégorie d'origine.

Mais auparavant, pour positionner chaque catégorie, nous allons rechercher les sicav les plus proches des centres de gravité.

4.3.1. Recherche des parangons

Dans son article, Alain Morineau propose d'identifier dans chaque classe un « parangon ». Un parangon est l'individu le plus proche du centre de gravité de la classe et qui est donc le représentant typique d'une classe.

Dans chaque sous-période nous avons donc recherché la sicav la plus représentative de chaque catégorie (seules les catégories d'effectif supérieur à 2 sont retenues).

Pour toutes les sous-périodes et toutes les catégories, le parangon élu dans une catégorie fait bien partie de ladite catégorie, voir l'annexe *Parangons* qui donne les tableaux analogues pour les autres études.

Pour les catégories obligataires qui sont très peu différenciées, 3 erreurs de classement sont à noter mais les parangons restent toujours classés dans les obligataires.

Parmi ces parangons, les plus **représentatifs** et les plus **stables** sont *9505 Slivarente* pour les sicav obligations françaises, *9789 Épargne Institutions* pour les obligataires de 1^{re} catégorie et *9569 Sélection Croissance* pour les sicav actions étrangères.

4.3.2. *Les sicav atypiques de leur catégorie*

A l'inverse des parangons qui constituent les noyaux des catégories, on a repéré les sicav atypiques de leur catégorie et qui sont repérables par les anomalies de classement.

On retrouve comme déjà signalé sur les plans factoriels, la sicav 9740 *Via Investissement* et la sicav 9684 *UAP Actions Sélectionnées* de la catégorie 5 (actions à vocation étrangère) ainsi que la sicav 9545 *Convertimmo* de la catégorie 12 (immobilières et foncières).

Au bilan le nombre très faible d'anomalies constatées entre les catégories définies *a priori* et celles obtenues *a posteriori* à partir des résultats effectifs, montre bien l'identité des modes de gestion des sicav.

CONCLUSION

Nous avons développé une démarche inductive, c'est-à-dire une analyse investigative que tout individu déploie naturellement, lorsqu'il ne sait pas et qu'il veut découvrir, dans le but d'apprendre.

L'analyse des données permet cette démarche exploratoire, lorsque sur les données nous n'avons au départ aucun *a priori*. Mais le revers de la médaille est bien connu :

« Qui procède par induction n'offre ni démonstration ni déduction nécessaire... cependant il fait voir quelque chose. »¹

Pour éprouver la stabilité des structures révélées nous avons effectué plusieurs études en découpant en sous-périodes « l'horizon » historique, afin d'appréhender la robustesse des formes révélées.

Nous avons utilisé plusieurs méthodes d'analyses de données en les associant à des procédures de visualisation, pour faciliter les interprétations. C'est une démarche qui va dans le sens de la tradition de l'École française d'analyse exploratoire des données où la représentation graphique est présente et les hypothèses et les modèles sont absents.

Le processus d'analyse utilisé ici visait à fournir une aide à la compréhension de la réalité en recherchant les homogénéités et les particularités dans les rentabilités des sicav.

1. Jean-Paul Benzecri, « D'Aristote à l'analyse multidimensionnelle : logique et ontologie dans l'induction statistique », *Cahiers de l'analyse de données*, 1988, n° 3, p. 325-338.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

Lorsque l'on étudie l'évolution des rentabilités au cours du temps, ce n'est pas le **niveau** instantané des rentabilités qui apporte de l'information, mais l'**écart** entre les rentabilités à une même date.

C'est bien ce que prend en compte l'analyse factorielle des correspondances, puisque cette technique opère sur les écarts de profils aux mêmes dates.

Cette méthodologie d'analyse a permis de montrer la grande stabilité dans l'homogénéité des catégories définies a priori, et par là, de porter un jugement sur la qualité des gestionnaires de sicav.

Nous avons par cette méthodologie révélé une forte régularité des gestionnaires qui ne peut être mise en évidence par les autres techniques statistiques classiques.

Cette régularité dans les « images de synthèse » nous permet de conclure que le microcosme des gérants de sicav ont de par leur formation, leurs techniques, leur mode de raisonnement et les contraintes, **des gestions très voisines**.

La partition en catégories définie a priori est bien une réalité a posteriori du comportement des rentabilités tout au long de ces 20 années.

Nous avons fait apparaître que la catégorie 4 des sicav actions à vocation française n'était pas, dans le passé, une catégorie différenciée. Au cours du temps, cette catégorie est devenue une catégorie à part entière, homogène et différenciée.

De manière complémentaire, nous avons pu repérer les sicav au comportement atypique.

Pour aller plus loin et pour interpréter les analyses, il faudrait projeter des variables illustratives qui pourraient expliquer la spatialisation obtenue sur les plans factoriels, en repérant les facteurs responsables, sans doute à chercher dans le comportement de la profession.

On pourrait envisager d'introduire comme variables explicatives, des variables de techniques de gestion utilisées, obtenues par enquête auprès des gérants de ces sicav. Plus classiquement, des variables macro-économiques, comme le taux de croissance de la production industrielle, le taux de variation de l'inflation, des séries de taux d'intérêts, etc. pourraient aider aux interprétations.

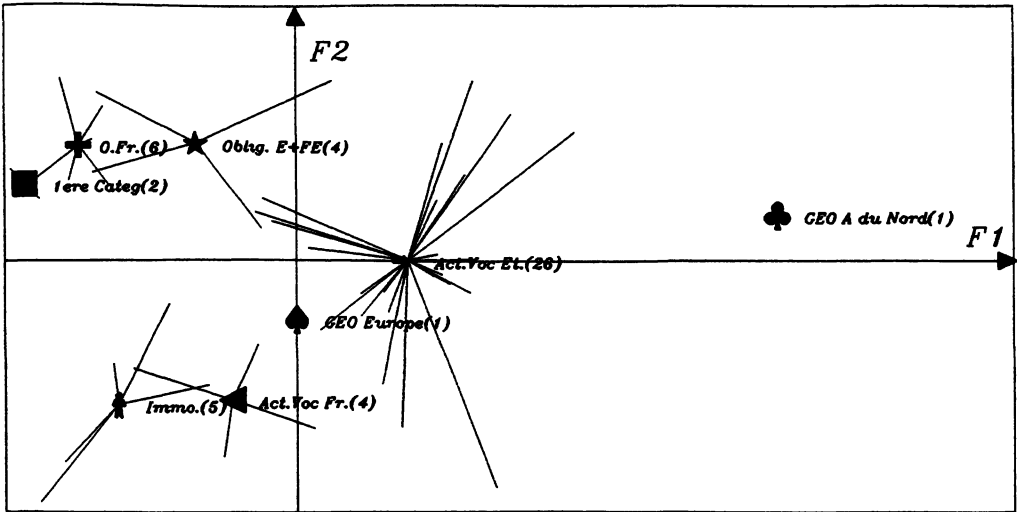
Pour conclure, nous avons analysé le passé, mais analyser le passé est indispensable pour anticiper et prévoir l'avenir. En empruntant la formule à Pascal Mazodier¹, l'**après-vision** conduira peut-être un jour à la **prévision**.

1. « La prévision ou l'après-vision », débat animé par Pascal Mazodier, 3^e journée de méthodologie statistique, ENSAE, 15-16 décembre 1993.

ANNEXES

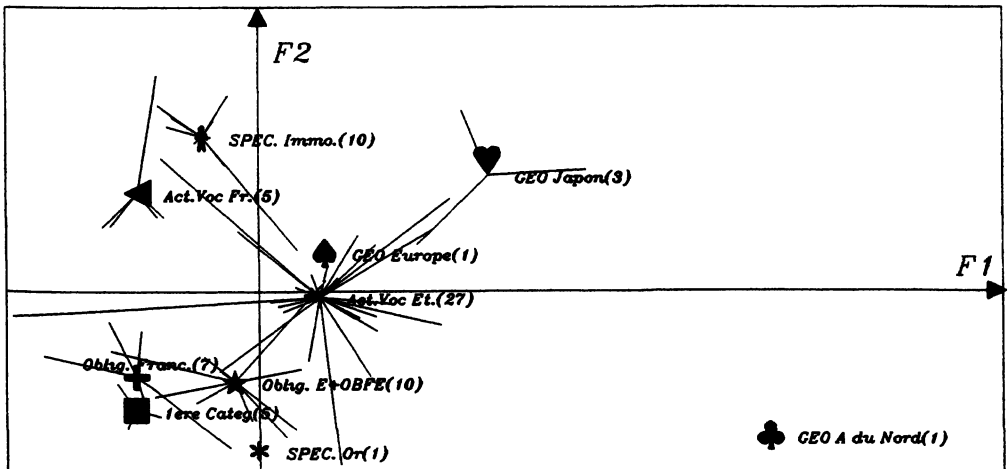
SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse Factorielle des Correspondances
 champ 49 Sicav - Horizon 31/12/71 - 31/12/81 soit 10 années



SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse Factorielle des Correspondances
 pour des rentabilites composees sur 5 ans
 champ 70 Sicav - Horizon 31/12/77 - 31/12/81 soit 14 années

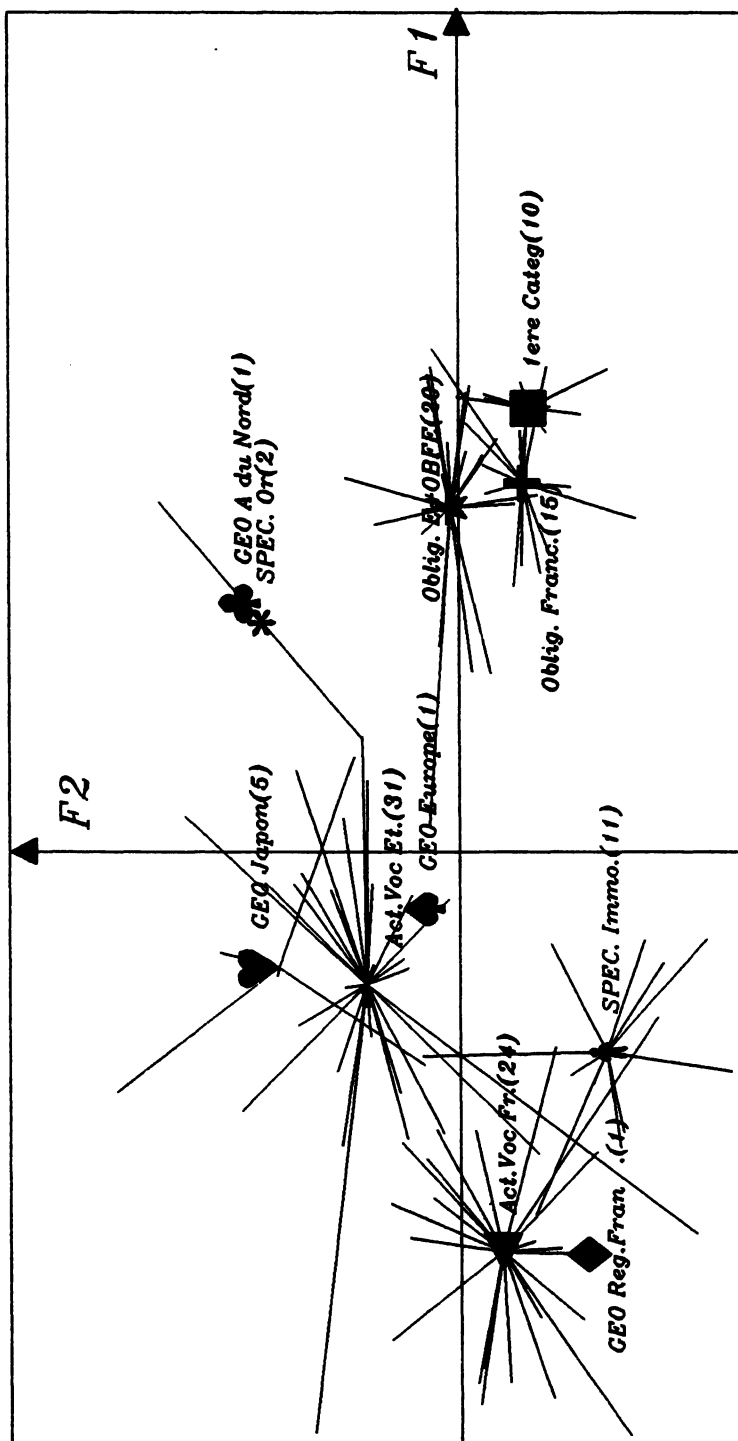


SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse Factorielle des Correspondances

121 Sicav - Horizon 31/12/83 - 31/12/91 soit 8 années

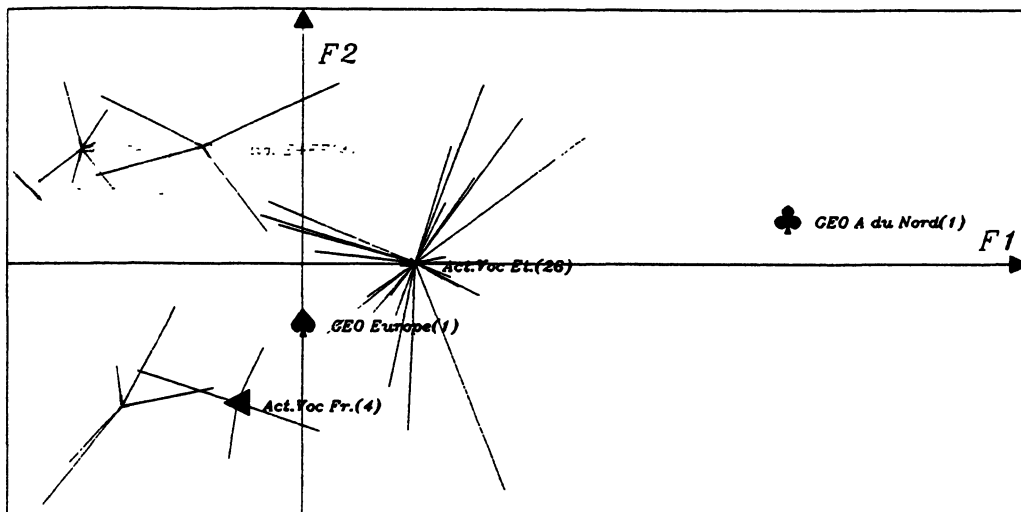
champ



Note: Par categorie, les SICAV sont reliees au Centre de Gravite
 Nom de la categorie (effectif de la categorie)

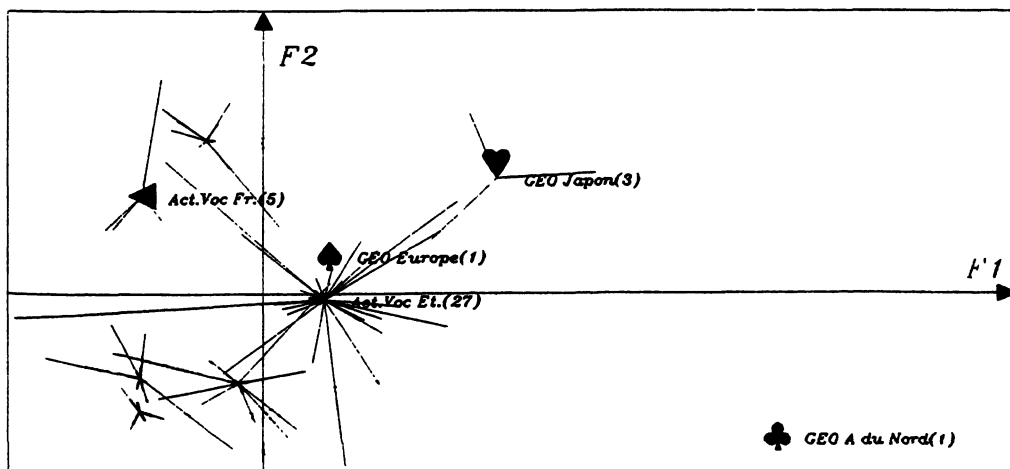
SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse Factorielle des Correspondances
champ 49 Sicav - Horizon 31/12/71 - 31/12/81 soit 10 années



SICAV par CATEGORIE et CENTRES de GRAVITES

Analyse Factorielle des Correspondances
pour des rentabilités composées sur 5 ans
champ 70 Sicav - Horizon 31/12/77 - 31/12/91 soit 14 années



RÉFÉRENCES

- ALBOUY M. (1982) *Stabilité du classement des sicav diversifiées de 1974 à 1980*, juin 1982, CERAG.
- ALBOUY M. et MENGE A. (1980) *Objectifs et performances de sicav*, juin 1982, CERAG.
- ALUDAAT K. et QUAQAZEH A. (1988) "Corrélation entre séries chronologiques : méthode des séries décalées", *Cahiers de l'analyse de données*, 1988, n° 4, 495-502.
- BENZECRI J.-P. (1973) "La place de l'a priori", *Encyclopedia Universalis*, vol. 17, Organum, 11-24.
- BENZECRI J.-P. (1976-1977) "Histoire et préhistoire de l'analyse des données", *Cahiers de l'analyse de données*, 1976, n° 1, 9-32 ; n° 2, 101-120 ; n° 3, 221-241 ; n° 4, 434-446 ; 1977, n° 1, 9-40.
- BENZECRI J.-P. (1988) "D'Aristote à l'analyse multidimensionnelle : logique et ontologie dans l'induction statistique", *Cahiers de l'analyse de données*, 1988, n° 3, 325-338.
- BENZECRI J.-P. (1987) "L'analyse des séries chronologiques, Méthodes des séries décalées", *Cahiers de l'analyse de données*, 1987, n° 4, 289-309.
- BOUROCHE J.-M. et SAPORTA G. (1983) *L'Analyse de données*, Que Sais-je ? n° 1854, PUF.
- BRETON T. (1991) *La dimension invisible, le défi du temps et de l'information*, Éditions Odile Jacob.
- CLAIRAMBAULT J. et CELEUX G. (1981) "Analyse discriminante appliquée à l'étude du rythme cardiaque", *Revue de MODULAD*, n° 8, décembre 1981, INRIA.
- CLEVELAND W.S. et MCGILL R. (1987) "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data", *J. R. Statist. Social A*, 1987, part. 3, 192-229.
- DROESBEKE J.-J., FICHET B. et TASSI Ph. (1989) *Séries chronologiques*, ASU.
- DROESBEKE J.-J. et TASSI Ph. (1990) *Histoire de la statistique*, Que Sais-je ? n° 2527, PUF.
- DE DROUAS F., DE VITRY B. et BOULIER J.-F. (1988) *Lois stables et cours boursiers*, décembre 1988, Journées AFFI.
- FÉNELON P. (1981) *Qu'est-ce l'analyse de données ?*, Lefonén, Paris.
- FONTAINE P. (1990) "Peut-on prédire l'évolution des marchés d'actions à partir des cours et des dividendes passés ? – tests de marche au hasard et de co-intégration", *JSSP* 1990-1, 16-36.
- GALLAIS-HAMONNO G. (1992) *Sicav et fonds communs de placement, les OPCVM en France*, Que Sais-je ? n° 2654, PUF.
- GALLAIS-HAMONNO G. et LE GUEN M. (1992) "Les résultats des sicav sur longue période", *Banque*, décembre 1992, 1139-1146.
- GOURIÉROUX Chr. (1990) "Quelques développements récents en séries temporelles", *JSSP* 1990-1, 7-15.
- GOURIÉROUX Chr. et MONTFORT A. (1990) *Séries temporelles et modèles mathématiques*, *Economica*.
- JAMBU M. et LEBEAUX M.-O. (1978) *La classification automatique pour l'analyse de données*, tome 1 : *Méthodes et algorithmes*, tome 2 : *Logiciels*, Dunod.
- JMP SAS Institute : *Logiciel de statistiques et visualisation*.

SÉRIES CHRONOLOGIQUES ET ANALYSE DE DONNÉES

- JOURNAL DE LA SOCIÉTÉ DE STATISTIQUE DE PARIS (1992) "Marchés financiers et gestion de portefeuilles : une mise en perspective des nouveaux outils", 1992-4.
- KAUFMAN H. et GROBOILLOT J.-C. (1974) *Les techniques de la prévision à court terme*, Dunod.
- LEBART L. (1989) "Stratégie du traitement des données d'enquêtes", *Revue de MODULAD*, INRIA, juin 1989, 21-29.
- LA LOIRE J.-C. (1972) *Méthode du traitement des chroniques*, CIRO.
- LEBART L., MORINEAU A. et FÉNELON J.-P. (1979) *Traitement des données statistiques*, Dunod.
- LE GUEN M. (1983) *Séries chronologiques et analyse de données, des images de synthèse pour répondre à la question "ex-post, la classification COB des sicav est-elle pertinente ?"*, Rapport de recherche 1993.
- LE GUEN M., SAINT-MARC M. (1984) *Recherche sur les méthodes économétriques de la datation des cycles : application aux cycles monétaires des XIX^e et XX^e siècles*, colloque Structures économiques et Économétrie, Aix-en-Provence, 1984.
- LOUVET P. et TARAMASCO O. (1992) *Analyse descriptive de la dépendance sérielle sur les rentabilités journalières à la Bourse de Paris*, décembre 1992, Journées AFFI.
- LUCAS M. (1993) "La communication graphique", *Courrier du CNRS*, février 1993, n° 80, 102.
- MARCOTORCHINO F. (1991) *La classification automatique aujourd'hui*, Publications Scientifiques et Techniques d'IBM France, novembre 1991, 35-91.
- MOLES A.A. (1990) *Les sciences de l'imprécis*, Seuil, 1990.
- MORINEAU A. (1989) "SPAD.N, logiciel pour l'analyse statistique de données", *Revue de MODULAD*, INRIA, mars 1989, n° 6, 27-59.
- ORLÉAN A. (1991) "Les désordres boursiers", *La Recherche*, mai 1991, 668-672.
- PHILIPS L., BLOMME R. et VAN DEN BERGHE C. (1981) *L'analyse chronologique*, Cabay-Économica.
- PONCET P. (1990) "Performances des OPCVM et efficience des marchés boursiers", *AFFI, Lettre d'information* n° 39, mars 1990.
- QUAQAZEH A. (1987) "L'analyse des séries chronologiques décalées. Principes d'interprétation sur des cas modèles", *Cahiers de l'analyse de données*, 1987, n° 4, 407-418.
- ROSENWALD G. (1988) *Performances des sicav françaises : 1982-1988*, Les Entretiens de la Finance, AFFI, décembre 1988.
- SAS System for Elementary Statistical Analysis (1987) SAS Institute.
- TARAMASCO O. et GIRERD-POTIN I. (1992) *Les rentabilités à la Bourse de Paris sont-elles chaotiques ?*, Les Entretiens de la Finance, AFFI, décembre 1992.
- VOLLE M. (1988) "Les Techniques de l'Observation Statistique", *Cahiers français*, mars-avril, 5-10.
- WHELLWRIGHT S.C. et MAKRIDAKIS S. (1974) *Choix et valeurs des méthodes de prévision*, Dunod.
- WONG LAI KUEN et S.Y.L. CHEUNG (1989) "La Bourse de Hong Kong : Analyse des cours en février-mars 1987 par la méthode des séries décalées", *Cahiers de l'analyse de données*, 1989, n° 2, 133-142.