

Typologie des cantons suisses : démarche et aspects méthodologiques

Autor(en): **Horber, Eugen / Joye, Dominique**

Objektyp: **Article**

Zeitschrift: **Annuaire suisse de science politique = Schweizerisches Jahrbuch für Politische Wissenschaft**

Band (Jahr): **19 (1979)**

PDF erstellt am: **27.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-172170>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

TYPOLOGIE DES CANTONS SUISSES

Démarche et aspects méthodologiques

par Eugen Horber

et

Dominique Joye

Assistants au Département de Science politique
de l'Université de Genève

Cet article essaie d'établir une typologie des cantons suisses. L'intérêt de cette démarche nous a semblé être à plusieurs niveaux. Dans un premier temps nous devions ainsi permettre un choix de cantons pour une recherche particulière, ce qui nous a amené, dans un deuxième temps, à explorer diverses méthodes de classification automatique et d'analyse dimensionnelle, et donc de proposer une solution pratique à de tels problèmes de recherche. Nous avons ainsi été amené à fonder une démarche exploratoire qui nous paraît plus féconde dans les situations où l'on a pas l'information nécessaire à des modèles statistiques forts. Enfin les résultats que nous avons ainsi obtenus montrent très bien les différences structurelles entre les cantons suisses, différences qui sont souvent révélatrices des anciens clivages.

Der vorliegende Aufsatz versucht die Probleme zu lösen, die sich bei der Auswahl von Kantonen im Rahmen eines Forschungsprojektes gestellt haben. Verschiedene typologische und dimensionenanalytische Verfahren werden auf ihre Eignung und optimalen Einsatz bei der konkreten Anwendung hin untersucht, wo nicht Verifizieren von Hypothesen im Mittelpunkt steht, sondern Entdeckung von Strukturen und Typenbenennung. Die Analyse lässt die klassischen „cleavages“ der Schweiz im Lichte der strukturellen Variablen erscheinen. Der Nutzen eines solchen Ansatzes beschränkt sich nicht auf den vorliegenden Fall; er kann mit Gewinn auf andere Probleme der Datenexploration angewendet werden, wo die üblichen statistischen Verfahren wenig fruchtbar sind.

I. Introduction et point de départ

Pourquoi faire une typologie des cantons suisses? Quel est l'intérêt d'une telle démarche? Nous pensons que cela deviendra plus clair si l'on expose d'abord ce qui nous a amené à cette étude. Une recherche est en cours au Département de science politique sur la culture politique des hauts fonctionnaires¹ menée par

1 P. Urio, G. Arigoni, I. Muller, *Pour une étude de la haute administration publique*, Département de science politique, Genève 1978.

Paolo Urio et son équipe. Cette recherche se fait par entretiens semi-directifs. Une première étape s'est faite sur l'ensemble des hauts fonctionnaires genevois et tessinois² et son extension est envisagée à l'ensemble de la Suisse. Comme il n'était évidemment pas possible d'interroger les quelques 2500 hauts fonctionnaires (un interview dure entre 2 et 3 heures et sa transcription revient à 200.—), il fallut opérer un choix qui pouvait se faire de deux manières: par échantillonnage direct des fonctionnaires ou bien en opérant un choix de cantons, ce qui permettait alors d'étudier les réseaux à l'intérieur d'un canton et de mieux distinguer les effets du service ou ceux du contexte. Dès lors, se posait le problème du choix des cantons.

De plus, il fallait également prévoir le refus éventuel d'une administration; il s'agissait de définir des groupes cohérents de cantons à l'intérieur desquels le chercheur pouvait effectuer son choix.

Pour opérer cette classification, nous avons fixé un certain nombre de critères. A la suite des travaux de R. Hofferbert³, notamment, il apparaissait clairement que nous devions tenir compte de facteurs politiques et socio-économiques. De plus, la recherche se faisant sur la haute administration, il nous fallait également des critères administratifs qui représentent d'une manière ou d'une autre la structure de l'administration. Il s'agira donc, dans un premier temps de choisir les variables pour représenter ces critères, de voir aussi comment elles s'organisent entre elles (c'est-à-dire de voir quels sont les liens à ce niveau entre le politique, l'administratif et le contexte socio-économique) avant de présenter les classifications que nous avons obtenues.

Toutefois, une telle étude ne va pas sans poser d'épineux problèmes méthodologiques: les typologies sont des choses récentes,^{3a} et leur utilisation en science politique ou en sociologie reste rare.⁴ De plus, il existe une grande variété de méthodes mais fort peu d'ouvrages comparant les diverses variantes;⁵ aussi, nous pensons que cet article pourra intéresser tout chercheur auquel se pose un problème de classification semblable. Mais il est évident que l'utilisation de ces méthodes implique une démarche analytique particulière, comme le dit Tukey:⁶ "Different ends require different means and different logical structure". Aussi,

2 Ceci grâce aux subsides nos. 1.542-0.77 et 1.859-0.78 du Fonds National Suisse de la recherche scientifique.

3 R. M. Hofferbert, *The Study of Public Policy*, Bobs Merrill, Indianapolis, 1974.

3a Il y eut une floraison d'articles dans les années 60. Plusieurs ouvrages ont parus au début des années 70, notamment: B. Everitt, *Cluster Analysis*, Heinemann, London, 1974; N. Jardine, R. Sibson, *Mathematical Taxonomy*, Wiley, New York, 1971; M. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York 1973; H. J. Cole, *Numerical Taxonomy*, Academic Press, New York, 1969; J. A. Hartigan, *Clustering Algorithms*, Wiley, New York, 1975.

4 Cf. par exemple, B. Russett, *Discovering voting groups in the UN*, Sage, 1972; P. Willets, *Cluster-Bloc Analysis and Statistical Inference*, APSR, 66, 1972.

5 Citons toutefois l'article de R. M. Cormack, *A revue of classification*, J. of the Royal Statistical Society, A134, 1971, pp. 321–367; Everitt, *op. cit.*

6 In Cormack, *op. cit.*

dans un premier temps, il nous faudra préciser quelle est notre “logical structure”, ou, plus simplement, notre démarche.⁷

II. Logique de la démarche

Le but d'une recherche de ce type est d'arriver à dominer une masse d'informations, de manière à ce que le chercheur puisse lui donner une signification. Cet aspect de donner une signification nous semble essentiel; en effet, l'empirisme part généralement de la définition de l'objet comme étant donnée, et les données que l'on peut avoir sur lui comme l'épuisant progressivement. Or nous pensons au contraire qu'il ne peut être défini que par rapport à un certain nombre d'attributs (ces attributs le définissant mais ne l'épuisant pas). L'avantage des typologies c'est de rendre explicite un certain nombre de ces attributs. Mais, là encore, ce n'est pas suffisant dans la mesure où l'on obtient alors un regroupement, pour le moment arbitraire. Il importe alors d'interpréter ce regroupement, au besoin de créer les outils nécessaires à saisir cet objet construit, autrement dit à utiliser ce que nous avons établi. Cela implique toute une part de création conceptuelle.⁸ Dans notre cas, un élément nous permet de préciser notre interprétation: il s'agit des cantons suisses qui se sont créés à travers tout un processus historique connu (nous y reviendrons au moment de l'interprétation).

Ainsi, les analyses qui vont suivre vont porter d'abord sur les variables pour savoir ce que l'on a mesuré, et à partir de quoi l'on va définir nos regroupements. La deuxième étape porte sur les regroupements eux-mêmes. Et là aussi la discussion sera double: elle portera aussi bien sur la description des regroupements que sur ce qui les a produit. En effet, même si l'essentiel n'est pas la discussion statistique, il faut savoir ce qui s'est passé et les limitations de méthodes. Certes une configuration au hasard pourrait être suggestive, mais nous doutons que ce soit la plus suggestive; ce qui nous fera discuter les diverses méthodes, leur interprétation, et leur adéquation au problème posé. “Adéquation” en ce sens que la réduction des données se fasse avec un minimum de perte d'information (critères statistiques) et que la représentation soit significative aux yeux du chercheur.

Pour préciser, il nous faut introduire deux personnages (dont rien n'empêche qu'ils soient confondus en une seule et même personne): le chercheur qui définit les orientations et le méthodologue qui propose les représentations.

D'habitude, le chercheur travaille sur un modèle explicatif, où avec un certain nombre de variables il essaie de parvenir à une explication un exemple typique en est le modèle path). Or, nos typologies sont plutôt une description; l'inter-

7 Il faut relever ici que toute une “nouvelle école” de statistique s'est développée à Princeton, inspirée par John Tukey.

8 Ch. Roig, “Pour une théorie actualisée des typologies”, ASSP, 1976.

prétation elle fait intervenir une série de connaissances extérieures à nos variables pour décrire les types résultants. C'est à ce niveau que nous pensons qu'une description incluant des éléments extérieurs peut avoir une valeur explicative. Cette position se situe tout à fait dans la ligne des travaux de Goldmann sur les rapports explication-compréhension.⁹ Ainsi dans notre étude, nous interpréterons nos types en fonction d'autres éléments: évolution historique, système de mentalités, etc. . . Et nous pensons que cette inclusion de connaissances extérieures, inévitable dans les typologies, en fait le grand intérêt et est dans tous les cas génératrice d'hypothèses et d'idées nouvelles.

Le méthodologue a à sa disposition des méthodes que l'on peut classer sur un continuum qui va des méthodes exploratoires aux méthodes confirmatoires en passant par les méthodes dites robustes. Par méthode confirmatoire, on entend une méthode qui exige un modèle explicite (modèle "fort" du type analyse causale) et diverses propriétés statistiques des variables (théorèmes de Gauss-Markov: normalité, etc. . .). Une méthode robuste suppose aussi un modèle mais est peu sensible aux violations des hypothèses sur les données. A l'autre extrême, les méthodes exploratoires exigent un minimum d'hypothèses et pas de modèle.¹⁰

III. Les Variables

L'ensemble des variables disponibles comprend les principales publications du Bureau fédéral de statistique, ainsi que certaines données sur les administrations cantonales recueillies dans le cadre du projet sur les hauts fonctionnaires (P. Urio). Après un certain nombre d'analyses préliminaires (analyses factorielles, cf. plus loin), nous avons retenu les indicateurs suivants qui représentent nos trois critères de départ (politique, socio-économique, administratif):

variables socio-économiques

- JEUNE, VIEUX: % des jeunes, % des personnes âgées
- ALLMD: % des personnes parlant allemand pour mesurer la différenciation linguistique

9 L. Goldmann, „Epistémologie de la sociologie”, in: J. Piaget et al., *Logique et connaissance scientifique*, Paris, Gallimard, 1967.

10 “Once upon a time, statisticians only explored. Then they learned to confirm exactly – to confirm a few things exactly, each under very specific circumstances. As they emphasized exact confirmation, their techniques inevitably became less flexible. The connection of the most used techniques with past insights was weakened. Anything to which a confirmatory procedure was not explicitly attached was decried as “mere descriptive statistics”, no matter how much we had learned from it”. J. Tukey, *Exploratory Data Analysis*, Reading, Addison-Wesley, 1974.

- URB: % des personnes vivant dans des communes de plus de 10 000 habitants mesurant le facteur urbanisation
- PRIM, SEC: % des personnes travaillant dans le secteur primaire ou secondaire mesurant la structure économique
- POPACT: population active (% des personnes ayant un emploi) permettant de préciser cette structure
- ETRA: le ratio population suisse sur population étrangère
- REVM: la richesse des cantons est mesurée par la moyenne des revenus¹¹
- GINI: mesurant la disparité des revenus par un indice d'inégalité.¹²

variables politiques et administratives

- PART: la participation électorale aux élections fédérales de 1975
- SICONS, SIRAD, SISOC, SIOTH: la proportion de sièges conservateurs, radicaux, socialistes et autres (regroupés) au parlement cantonal fournissant l'information sur la structure partisane du canton
- DEPSOC: la part des dépenses sociales (éducation, santé publique et dépenses sociales proprement dites) dans les dépenses sociales totales
- DEPAD: la part des dépenses pour l'administration
- DEPHAB: les dépenses par habitant comme indicateur de l'importance de l'activité étatique
- ADHAB: les dépenses pour l'administration par habitant comme indicateur de la taille relative de l'appareil administratif
- HABFON: le nombre d'habitant par fonctionnaire
- DEFFON: dépenses totales par fonctionnaire

En ce qui concerne ces cinq dernières variables "administratives" on constate que – suivant notre logique – on aurait dû les regrouper en deux nouvelles variables pour éviter les risques de multicollinéarité. Toutefois, ces variables ont été rarement utilisées dans des études de ce type. Aussi nous les avons gardées telles quelles dans un but exploratoire pour des procédures peu sensibles à la multicollinéarité. En revanche, nous les avons regroupées pour les analyses factorielles.

IV. Les méthodes

Quelle que soit la logique de notre démarche et nos variables, il nous faut maintenant préciser nos méthodes.

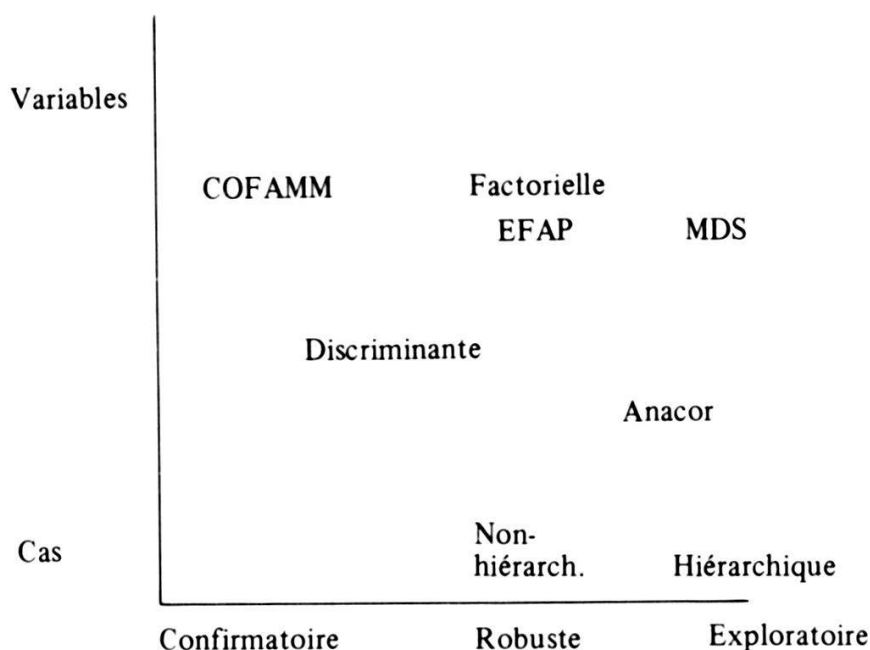
Nous avons déjà distingué entre méthodes confirmatoires et exploratoires. Il nous faut faire maintenant la distinction entre les méthodes qui opèrent des regroupements sur les cas ou sur les variables.¹³

11 Moyenne calculée sur les centres de classe.

12 Cet indice est un indice GINI modifié pour les données ordinales.

13 Cf. par exemple, l'article de R. N. Shepard, in A. K. Romney, *Multidimensional Scaling*, 2 vol., New York, Seminar Press, 1972.

Tableau 1: Classification des méthodes*



* MDS: Multidimensional Scaling, Cf. Coxon, A. P. M., Essex, ECPR; Kruskal, J. B., Multidimensional Scaling by Optimising Goodness of Fit to a Nonmetric Hypothesis, *Psychometrika*, 29, 1964.

Anacor: Analyse factorielle des correspondances, cf. Benzécri, J. P., (L'analyse des données, 2 vol., Paris, Dunod, 1973): c'est une analyse factorielle basée sur le chi-carré qui permet d'analyser les cas et les variables en même temps.

COFAMM, EFAP: Joreskog et ses collaborateurs (EFAP: Exploratory Factor Analysis Program, 1976; COFAMM: Confirmatory Factor Analysis with Model Modification, 1976, Chicago, National Education Resources) ont développé des méthodes d'analyse factorielle basées sur le maximum de ressemblance où l'une est exploratoire et l'autre confirmatoire, permettant de tester des modèles.

Anacor, COFAMM et EFAP ne sont pas pris en considération dans ce modèle.

1. Analyse factorielle

La première méthode, la plus connue, la plus ancienne aussi est l'analyse factorielle. Rappelons toutefois qu'elle peut être interprétée de deux manières:

- une manière "confirmatoire" où l'on suppose que le facteur est une variable explicative de nos variables manifestes;¹⁴
- une manière exploratoire où l'on cherche simplement à représenter nos variables (notre système d'attributs) en un espace de dimension moindre.

14 Cf. la discussion à ce propos et les limitations de ce paradigme in R. Boudon, *Analyse mathématique des faits sociaux*, Paris, 1967, p. 201 et s.

Quant aux hypothèses qu'elle demande, ce sont simplement, puisque le travail se fait sur une matrice de coefficients de corrélation de Pearson, les hypothèses classiques, notamment celle d'une relation linéaire entre les variables.

De plus, l'inversion de matrice intervenant, l'indépendance linéaire des variables est souhaitable.

Maintenant, un problème supplémentaire se pose à l'utilisateur: un système intégré d'analyse statistique comme SPSS propose plusieurs analyses différentes. Que choisir? Un premier choix s'impose entre le modèle en composantes principales (PA1) et les analyses factorielles proprement dites (PA2, RAO, ALPHA).

Sans entrer dans les détails, ces analyses factorielles nous semblent préférables dans la mesure où elles autorisent les variables à avoir une partie en "erreur", sans lien avec la structure. Ce qui veut dire, en d'autres termes, qu'elles essaient de maximiser la structure en ne prenant en compte que la partie commune des variables, tandis que l'analyse en composantes principales essaie de minimiser simplement la perte d'information. Comme nous voulons avant tout définir une structure, ce sont bel et bien les analyses factorielles proprement dites qui sont préférables.

Nous n'avons pas retenu non plus la méthode ALPHA car son but est de maximiser un coefficient de ressemblance entre l'échantillon et une population théorique de variables. Or, il est clair que ce n'est pas ce que nous cherchons.

Entre RAO et PA2, le choix a été plus difficile. En effet, RAO peut apparaître comme plus adéquate du point de vue statistique (elle s'apparente aux méthodes du maximum de vraisemblance) et elle permet de faire un test sur les facteurs utiles. De plus, elle est peu sensible à la taille de l'échantillon et peut donc sembler ad hoc pour un petit jeu de données. Toutefois, nos données sont assez fortement corrélées et sa méthode de travail est telle qu'elle est plus sensible à ce genre de problème. De plus, elle exige une hypothèse statistique plus forte: la multinormalité des variables. Suivant le cas, on utilisera donc l'une ou l'autre.¹⁵

Nous avons mentionné plus haut que l'utilité de l'analyse factorielle était de permettre la réduction du nombre des variables et d'éviter la multicollinéarité. Aussi, par exemple nous en avons effectué une sur nos cinq variables administratives.

Les deux facteurs résultants expliquent 77 % de la variance totale dans la matrice de données. Le premier est un facteur typiquement "fonctionnaire" et le second "dépenses". A partir de ce résultat, on a remplacé dans les analyses ces cinq variables par ces deux facteurs: FAC1 et FAC2.

15 Pour les problèmes techniques de l'analyse factorielle, on se référera à: H. H. Harman, *Modern Factor Analysis*, Chicago, Chic. U. P., 1967. R. J. Rummel, *Applied Factor Analysis*, Evanston, Northwestern UP, 1969. J. P. Van de Geer, *Introduction to Multivariate Analysis for the Social Sciences*, Freeman, San Francisco, 1971. N. H. Nie, *SPSS*, 2nd ed., New York, McGraw Hill, 1975.

Tableau 2: Analyse factorielle (RAO) des variables administratives (factor loadings)

	Facteur I	Facteur II
DEBHAB	0.24	0.76
ADHAB	- 0.20	0.98
DEPAD	- 0.35	0.54
HABFON	0.83	- 0.15
DEFFON	0.99	0.06

Jusqu'à maintenant, l'analyse factorielle nous a servi à choisir et à préciser nos variables de départ. Nous allons aussi l'utiliser pour clarifier les relations entre nos variables. Pour cela, nous avons procédé à une analyse de type PA2 de nos variables. Avec trois axes, nous obtenons:

- une première dimension (la plus importante) regroupe surtout les variables de contexte: les oppositions primaire-secondaire, GINI-revenu, ainsi que la variable dépense (FAC2)
- le deuxième facteur reprend notre critère politique: sièges conservateurs contre sièges socialistes et dépenses sociales
- le troisième n'est pas défini par une opposition mais révèle plutôt un type de canton alémanique à structure économique secondaire

Tableau 3: Analyse factorielle (PA2) de toutes les variables (factor loadings)

	Factor 1	Factor 2	Factor 3
ETRA	- .66671	- .20688	.05363
PART	- .36458	- .13700	.13505
JEUNE	- .72753	- .63994	.23174
VIEUX	.22010	.75537	- .19362
SICONS	- .59063	- .59767	- .27364
SIDRAD	.47457	- .21872	- .41651
SIDSOC	.42292	.43207	.48638
SIOTH	.20705	.63339	.31538
POPACT	.74857	.49703	.23292
ALLMD	- .13941	- .10110	.79809
URB	.64345	.39132	- .00283
DEPSOC	.38732	.64336	.11051
PRI	- .88564	- .12875	- .08430
SEC	.03157	- .06209	.82136
REV	.69343	.04423	.02755
GINI	- .89651	- .23582	- .14198
FAC1	.05312	- .62687	.32164
FAC2	.77008	.19715	- .17344

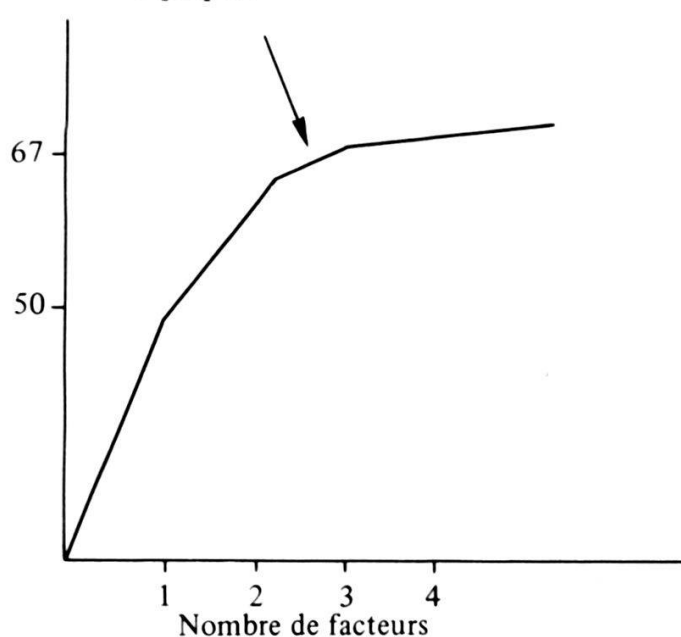
Notre variable fonctionnaires (FAC1) se corrèle surtout sur les facteurs 2 et 3 ce qui rejoint ce que nous disions au début sur la non-indépendance de nos critères.

L'intérêt de ce genre d'étude est surtout de voir le système de relation qui unit nos variables ce qui nous facilitera l'interprétation des analyses typologiques proprement dites.

Toutefois avant d'en arriver là, une remarque s'impose: comment déterminer le nombre de facteurs adéquat? Différents critères sont possibles (Critère de Kaiser, test de X^2 , etc. . .). En fait, un critère nous semble particulièrement simple et utile: celui de Scree: ¹⁶ il consiste à dessiner un petit graphique comme le tableau 3, avec d'un côté le nombre de facteurs et de l'autre les pourcentages de variance expliquée; l'optimum étant atteint au coude de la courbe.

Tableau 4: Nombre optimal de facteurs

% de variance expliquée



Un critère qui intervient aussi, suivant d'ailleurs la logique de notre démarche, est l'interprétabilité des facteurs: en effet, il est clair que si l'on a autant de facteurs que de variables la perte d'information sera nulle. Mais, en contrepartie, l'on aura rien gagné en réduction d'information et donc la structure restera illisible.

Nous aurons le même problème pour les typologies. Le seul critère statistique applicable est le critère de Scree ou des méthodes de simulation numérique. ¹⁷

16 Cf. *Rummel*, 1969, *op. cit.*

17 Cf. *Handley*, *Optimum for Typology*, papier non publié, Genève.

Toutefois, la question de la dénomination des types est encore plus importante. Un critère supplémentaire sera donc celui de l'interprétabilité en termes des critères du chercheur.

Pour notre part, dans cet article, nous nous sommes fixés un chiffre de quatre types car il nous permettait d'avoir généralement une bonne représentation, pas trop de cantons isolés formant un cas unique, et surtout nous l'avons gardé quelle que soit la méthode, pour permettre la comparaison. Mais, au lieu d'aborder déjà cette question, examinons plutôt les méthodes typologiques dont nous disposons.

2. *Analyses typologiques*

Il en existe une grande quantité et de nombreuses variantes que l'on peut regrouper suivant divers critères;¹⁸ le premier étant: hiérarchiques-non-hiérarchiques.

a) Hiérarchiques

Un exemple en est la méthode de Johnson qui consiste, à partir d'une matrice de distance de construire des types en regroupant chaque fois le plus proche voisin.¹⁹ Cette méthode est intéressante dans la mesure où elle ne fait pas l'hypothèse d'une métrique mais uniquement du respect de certaines relations d'ordre (inégalité ultramétrique).²⁰

Un autre avantage en est sa simplicité qui la rend facilement accessible ainsi que les représentations ou Dendrogram que l'on peut en tirer.

En contre-partie, cette méthode a tendance à entraîner des regroupements forcés, des enchaînements. Ceci ne serait pas gênant si l'on suppose qu'il existe une hiérarchie sous-jacente à nos regroupements. Pourtant, vu l'ensemble de dimensions qui sous-tend notre problème (ce que nous avons montré dans l'analyse factorielle) nous ne pensons pas que cette méthode corresponde entièrement à notre problème.

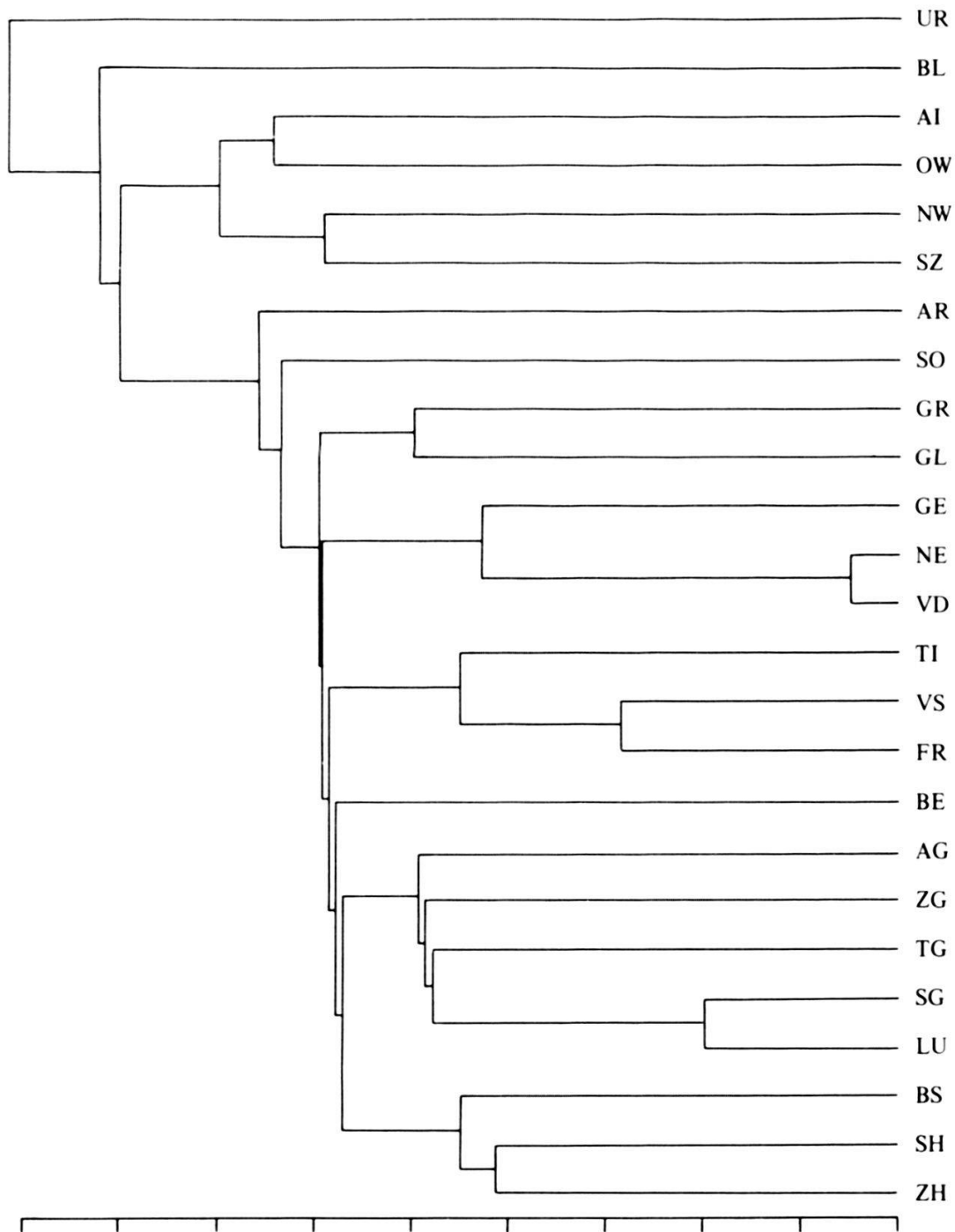
Elle reste de toute manière intéressante comme configuration de départ pour d'autres analyses que nous allons invoquer maintenant.

18 La plupart sont disponibles dans une série de programmes: CLUSTAN 1c du Computer Center, University College, London. On a également utilisé le système SYAT de M. Bergonier implémenté et adapté sur l'Univac 1108 par l'équipe du département de science politique, université de Genève.

19 Cf. S. C. Johnson, "Hierarchical Clustering Schemes", *Psychometrika*, Vol. 32, 3, 1967.

20 Cf. Coxon, *op. cit.*

Tableau 5: Résultat d'une analyse hiérarchique: dendrogramme (méthode: single linkage)



b) Analyses non-hiérarchiques

Une distinction doit être faite entre les méthodes où le classement est exclusif et les méthodes qui autorisent un cas à appartenir à plusieurs types.

Ce sont les méthodes à classement exclusif qui nous ont le plus intéressés: en effet, notre problème était, rappelons-le, de classer les différents cantons.

Parmi ces méthodes, on doit faire une nouvelle distinction entre les méthodes cherchant à optimiser un critère de classification et celles fondées sur l'analyse de concentrations de points. Les méthodes qui cherchent à optimiser un coefficient classificateur sont souvent basées sur l'analyse de la variance (ou des mesures dérivées). On veut en effet que la variance inter-groupes soit la plus grande possible. Cela oblige à certaines hypothèses: une métrique de mesure pour le critère et l'hypothèse que le comportement du coefficient soit le même dans tous les groupes: autrement dit que la distribution ne diffère pas d'un groupe à l'autre (c'est d'ailleurs la même hypothèse qui est à la base de l'analyse discriminante; nous y reviendrons).

La méthode DNDRITE²¹ par exemple applique une méthode de maximisation d'un critère dérivé de la théorie des graphes.

RELOC²² part d'une configuration initiale donnée; il revoit la répartition des cas entre les divers types pour maximiser un critère de similarité.

Ce genre de méthode a une autre limitation: dès que le nombre d'individus dépasse une certaine limite, le temps d'ordinateur nécessaire à l'analyse de toutes les configurations devient considérable.²³ Une solution à ce problème est de prendre un échantillon d'individus. C'est ce que fait SYAT notamment. Il²⁴ fonctionne, en fait, en quatre phases. A partir de plusieurs échantillons aléatoires, on construit une typologie à TM types (nombre maximal spécifié) par analyse de toutes les distances (phase 1). Un échantillon plus important est réduit à TM types (phase 2). Ensuite, TM types sont réduits à tm (nombre minimal de types) et l'optimum est déterminé (phase 3). Tous les individus sont repris et situés, puis la solution est optimisée (phase 4).

SYAT présente l'avantage d'être fiable même pour un nombre d'individus très élevé en un temps de calcul raisonnable.

Enfin, une autre série de méthodes travaille à partir de l'analyse des concentrations, des distributions de points.²⁵ Mais la plupart de ces méthodes échappent rarement à un risque: la sphéricité, c'est-à-dire la supposition que le regroupement a une forme plus ou moins sphérique.²⁶

21 Cf. D. Wishart, *CLUSTAN User Manual*, London, University College, 1975.

22 idem.

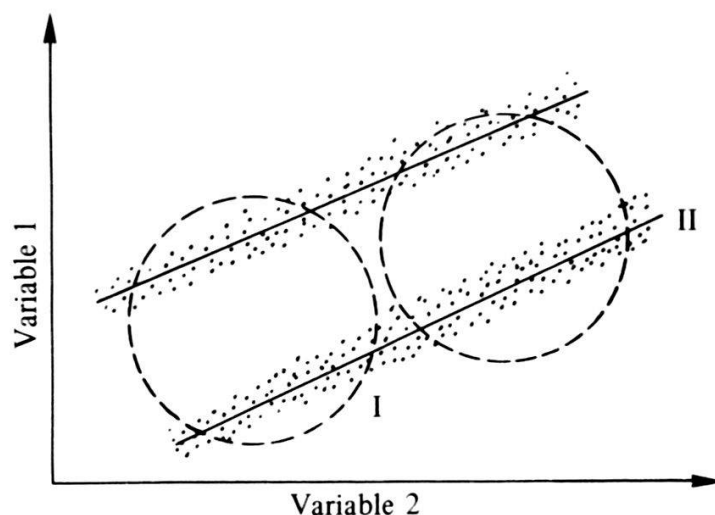
23 6 minutes pour notre problème.

24 Cf. H. Bergonier, *Méthodes d'analyse de données pour la détection et la description de structures*, Genève, Département de science politique, Etudes et recherches, no. 3, 1974.

25 Cf. méthode MODE in Wishart, *op. cit.*

26 Nous travaillons d'ailleurs sur des techniques permettant de mieux connaître les qualités et défauts de nos données, ce qui facilite le choix de la méthode à utiliser.

Tableau 6: Exemple de sphéricité



Quant aux méthodes qui permettent à un individu d'appartenir à plusieurs types, comme la méthode KDEND de Jardine-Sibson,²⁷ que nous avons utilisée pour vérifier la force de nos "clusters" et voir s'il y avait des cantons à la limite d'un type difficile à classer.²⁸

c) Analyse discriminante

L'analyse discriminante nous a permis de voir si les types définis par les différentes méthodes pouvaient se retrouver, même avec une méthode qui exige des hypothèses plus précises.

C'est une analyse qui essaie de prévoir l'appartenance à un groupe préalablement défini.²⁹ Par exemple, si l'on a deux groupes, on essaiera de construire une fonction linéaire de nos variables qui permette de prédire l'appartenance à un des groupes (pour N groupes, il faudra N-1 fonctions). Cette analyse fait partie des méthodes d'analyse de variance-covariance (comme l'analyse factorielle) et a donc les mêmes hypothèses, avec en plus celle que les distributions des variables soient semblables dans les différents groupes. Il en existe deux variantes principales: une méthode directe et une méthode qui sélectionne les variables les unes après les autres pour les faire entrer dans l'analyse (ou non). Pour notre part, étant donné la multicollinéarité dont souffrent nos variables, nous avons préféré la méthode pas à pas.³⁰

Nous avons procédé à une telle analyse de chacune des typologies: à chaque méthode nous avons obtenu une capacité de prédiction de 100 % ce qui veut dire que nos données supportent parfaitement les hypothèses les plus "fortes".

27 Cf. Wishart, *op. cit.*, Jardine-Sibson, *op. cit.*

28 Par exemple le cas du canton de Bâle-Campagne que l'on verra dans l'interprétation.

29 Cf. Nie et al., *op. cit.*, Van de Geer, *op. cit.*, M. G. Kendall, *A Course in Multivariate Analysis*, London, Griffin, 1972.

30 Lambda de Wilks en SPSS, Cf. Nie, *op. cit.*

V. Les résultats

Tableau 7: Résultats des analyses typologiques retenues

Méthode	Type 1				Type 2				Type 3				Type 4			
	ZH AR NE	BE SG	GL AG	ZG TG	SO TI	SH VD	BS BL	GE	LU GL SO AR SG	GR AG TG	FR TI	VD VS NE GE	FR NW BL	SZ FR SZ	FR NW BL	GR SZ NW BL AI
SYAT																
RELOCATE	ZH VD	BE VS	FR NE	BS GE	SH TI											
UNDRITE	ZH AR	BE SG	LU TG	ZG AG	BS SH											
KDEND*)	ZH ZG SG VS	BE FR GR NE	LU SO AG NE	SZ BS TG GE	OW SH TI VD											
Hierarchie (WARD)	ZH	BS	SH	VD	NE	GE										
Hierarchie (complète link.)	ZH SG NE	LU AG	ZG TG	FR TI	BS VD	SH VS										

* Les cantons en italiques appartiennent à plusieurs types

Comme notre tableau de résultats le fait ressortir, il y a des différences entre les résultats des diverses méthodes qui proviennent des poids différents donnés à certaines variables, ainsi que l'analyse discriminante nous l'a montré. Toutefois, les différences restent faibles. Aussi, nous allons d'abord interpréter ces analyses avant d'essayer de généraliser, de retrouver les constantes des diverses méthodes. Voyons d'abord les résultats de SYAT:

Tableau 8: Définition des types part SYAT

Ecarts à la moyenne

1 + correspond à un écart positif de 0.05 Ecart-type

1 - négatif

Les variables sont listées selon les valeurs de R^2 décroissantes explication moyenne générale 478 0/00

La liste est limitée aux variables représentant au total 95 0/00 de l'explication totale soit 19 variables sur 21 (90.0 0/0)

Explication moyenne 513 0/00 pour ces variables

Expl. cum. 0/00	R^2 en 0/00	Titre des variables ou des éventualités	Numéros des types			
			1	2	3	4
081	817	Primaire	+++++	---	-	+
160	792	Dép. administr. par habitant		+++++	-	-
235	756	GINI	++	----		++
310	754	Etrangers	++++	--	-	++
373	626	Jeunes	++	---	-	++
433	609	SiLeges conservateurs	--	-	-	++
491	574	Dépenses par habitant	-	++++	-	+
545	548	Urbanisation	---	+++		-
598	531	Dépenses administration	++	++		--
650	525	Population active	-	++	+	--
697	469	Sièges socialistes	---	++	+	-
740	427	Participation	----	-	+	-
779	393	Dépenses sociales	--	++	+	--
817	379	Habitants par fonctionant	++++	-	-	+
854	377	Secondaire	--	-	+	-
890	361	Sièges radicaux	----	+		
918	275	Dépenses tot par fonctionnant	++		-	+
945	274	Sièges autres	--	+	+	-
971	260	Vieux	+		+	--

Le tableau 8 présente l'essentiel de ces résultats. Il permet de définir le profil de chaque type en termes de variables utilisées.

Le quatrième type (Appenzell-Rhodes intérieur et Obwald) est caractérisé par l'importance du secteur primaire, des administrations de taille réduite.³¹ Par

31 Rappelons que nos indicateurs sont toujours relativisés par rapport au contexte.

rapport aux autres, ces deux cantons ont encore gardé une structure relativement plus traditionnelle.

Le deuxième type présente des caractéristiques opposées, c'est-à-dire des cantons urbanisés, avec une forte population d'étrangers, une distribution des revenus plus égalitaire et une administration plus importante, et en outre un secteur tertiaire très important.

Le troisième type est en sorte une version "développée" du premier: tout en présentant des aspects ruraux, ces cantons ont connu un développement économique plus poussé que les cantons du type 1 tout en conservant une structure moins différenciée de l'administration. Il est intéressant de noter ici que ces cantons sont des cantons du Sonderbund, si l'on fait abstraction des Grisons.

Le premier type est en sorte un type moyen dans le sens où les critères de discrimination et l'optimum des 4 types forcent un regroupement moins "typé" des cantons non encore classés en un groupe qui se rapproche de la moyenne de la Suisse. Cette remarque est valable mutatis mutandis pour les autres méthodes.

La méthode RELOCATE produit les résultats suivants:

Le type 1 est un regroupement de Zurich, Zoug, Bâle-ville et Schaffhouse d'une part (les cantons alémaniques les plus urbanisés avec tout ce que cela comporte: structure de la population, de l'administration, etc. . .) et d'autre part, les cantons latins: Fribourg, Tessin, Vaud, Neuchâtel, Genève et Valais.

Il est intéressant de noter que ces cantons se séparent des autres par l'urbanisation, l'indice d'inégalité et surtout les dépenses par fonctionnaire.

Le type 3 ne comprend que le seul canton d'Uri (on en reparlera)

Le type 4 est constitué par: Schwyz, Obwald, Nidwald, Appenzell Rhodes intérieur et l'inclassable Bâle-Campagne.

Quant au deuxième type, il regroupe les cantons alémaniques à prépondérance du secteur secondaire comme l'Argovie ou la Thurgovie. Il est, en quelque sorte, notre type "moyen".

Pour DNDRITE les différences sont surtout dans le fait qu'un type est composé des seuls cantons romands (discriminés par DEFFON et ADHAB). Uri a rejoint Nidwald et Bâle-Campagne. Les autres petits cantons restent groupés avec, en plus, les Grisons.³² Il reste toujours un type "moyen" de cantons alémaniques.

D'une manière générale, si l'on essaie de synthétiser ces résultats, on obtient:

- un regroupement de petits cantons de la Suisse dite "primitive";
- un regroupement de cantons alémaniques du "mittelland" où le secteur secondaire est relativement important;
- un regroupement des cantons ville comme Genève ou Bâle-Ville

32 La taille d'un canton est un phénomène complexe. Cette notion n'implique pas seulement la taille extérieure (surface, population), mais surtout une interdépendance entre taille de la population, développement économique, structure socio-économique, structure de l'administration etc. . .

– un regroupement des cantons latins (Fribourg et Valais ayant toutefois gardé une structure un peu particulière qui les différencie parfois).

– Deux cas spéciaux:

Bâle-Campagne qui se trouve tantôt avec les petits cantons, tantôt avec Bâle-Ville. Bien que présentant des aspects de canton moyen, le rayonnement de plusieurs centres économiques à la frontière marque son profil.³³

Uri qui se trouve souvent isolé dans les analyses. Cette situation s'explique par les variables "administratives" qui diffèrent des autres "petits" cantons à cause de besoins très spécifiques: les routes et tunnels du Gotthard. Un contrôle sur ces facteurs fait rentrer Uri dans le groupe des petits cantons.

Si l'on regarde maintenant quelles sont les variables qui définissent ces regroupements, on s'aperçoit que celles qui agissent le plus fortement sont les dépenses par fonctionnaire, la structure de la population, les variables d'urbanisation et la distribution des revenus (beaucoup plus, d'ailleurs, que le revenu lui-même).

A notre avis, une explication possible nous est donnée si l'on considère la typologie conceptuelle suivante (qui reprend nos résultats empiriques):

– des cantons "périphériques" où la structure est encore dominée, marquée par le poids du secteur primaire; la population active est plus réduite ce qui explique les grandes proportions de "jeunes" et de "vieux". Enfin la structure particulière du canton détermine les besoins administratifs, ce qui explique le niveau des dépenses par fonctionnaire;

– un autre type de cantons également "périphérique" est celui où une partie importante de la population travaille hors du canton: dès lors, ce statut n'apparaît pas dans la composition de la population mais bien au niveau des variables "administratives";

– quant aux autres cantons, il y a une distinction à faire entre les cantons marqués par le secteur secondaire, où les dépenses administratives restent à un niveau "moyen" et les cantons marqués par le tertiaire, urbanisés, avec un haut niveau de revenus et de fortes dépenses.

VI. Conclusion

On avait dit que le chercheur souhaitait une représentation adéquate des cantons suisses. En l'occurrence, il avait besoin de groupes de cantons à l'intérieur desquels il puisse les considérer comme "équivalents" par rapport aux critères spécifiés.

Cependant, plusieurs représentations ont été produites. Laquelle choisir? A part les critères méthodologiques dont nous avons parlé, c'est au chercheur de trancher selon l'importance qu'il accorde à tel ou tel critère. L'intérêt de la

33 Une méthode comme KDEND a d'ailleurs bien montré les difficultés de classement de Bâle-Campagne puisqu'elle classait ce canton dans trois types différents.

méthode se trouve donc dans cette interaction entre le chercheur et le méthodologue.

Mais l'intérêt réside aussi dans la classification obtenue: non qu'elle soit particulièrement nouvelle, elle reprend en effet largement ce qu'on a l'habitude de penser à propos de la Suisse, mais aussi par les liens qu'elle suggère entre la structure socio-économique, la structure politique et administrative. En effet, une telle juxtaposition de variables est génératrice d'hypothèses et c'est un des avantages de la méthode.

En d'autres termes, cette méthode reprend finalement les propositions que Mills fait pour stimuler "l'imagination sociologique".³⁴

34 Cf. C. W. Mills, *L'imagination sociologique*, Paris, Maspéro, 1967.