

GUTSATZ

**L'analyse des correspondances, système
de décision multidimensionnelle. Extrait
d'un mémoire de M. Gutsatz**

Les cahiers de l'analyse des données, tome 1, n° 1 (1976),
p. 47-59

http://www.numdam.org/item?id=CAD_1976__1_1_47_0

© Les cahiers de l'analyse des données, Dunod, 1976, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » 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/>

L'ANALYSE DES CORRESPONDANCES
SYSTEME DE DECISION MULTIDIMENSIONNELLE

Extrait d'un mémoire^(*) de M. Gutschatz⁽¹⁾

1. Où une application est tentée :

Nous avons tenté une application de l'analyse factorielle des correspondances au choix des systèmes d'aménagement de l'ouest parisien (**). Ce problème porte sur le choix d'une variante de tracé autoroutier (parmi 7) et sur l'étude des impacts dans divers domaines de ces différents tracés.

L'étude primitive proposait de "juger" les variantes selon 16 critères auxquels étaient associés autant d'indicateurs. Le tableau global était le suivant :

	CRITERES	INDICATEURS
NUISANCES EN FORETS	Effet de massif Micromorcellement Visibilité Nuisance phonique Emprise Détournement de fréquentation Accessibilité interne Sites affectés	Surface en massif de plus de 50 ha Nombre de parcelles isolées Surface où l'autoroute est visible Surface où l'autoroute est audible Surface d'emprise réelle Nombre de centres scolaires détournés de la forêt Nombre de chemins forestiers coupés Nombre de sites classés, d'étangs affectés
NUISANCES EN ZONE URBAINE	Atteinte aux équipements publics Expropriations Nuisance Compatibilité avec le PDUI	Nombre d'équipements publics à déplacer Nombre de personnes et d'entreprises à exproprier Nombre d'habitants dans la zone de nuisance Nombre de "caractères" du PDUI qui sont mis en cause, autres que les nuisances déjà citées.

(1) CEFI - Faculté de Sciences économiques - Université Aix-Marseille II

(*) Une analyse de l'ensemble de ce mémoire est publiée dans le présent cahier cf pp 95 ; le présent extrait constitue le §C du chapitre III du mémoire ; chapitre dont on a repris ici le titre.

(**) P. BERTIER, J. de MONTGOLFIER : On multicriteria analysis. An Application to a Forest Management Problem METRA, Vol. XIII, n°1, 1974.

TABLEAU 1

DEFINITION DES CRITERES

Critères	N°		Définition des indicateurs	Echelles	
				Unité	Intervalle
<i>Améliorer les transports</i>	1	Allongement de parcours	Somme pondérée des allongements de parcours	Kilomètre	0 à 5
<i>Structurer la banlieue</i>	2	Effet sur le développement urbain	Note qualitative (attribuée par des urbanistes)		0 à 6
<i>Minimiser le coût</i>	3	Coût financier	Coût	Millions de francs	0-2000
<i>Minimiser les nuisances en milieu urbain</i>	4	Effets sur les habitants de la zone de plus forte nuisance	Nombre d'habitants dans la zone de plus forte nuisance	1.000 habitants	0 - 10
	5	Effets sur l'environnement des communes traversées	Nombre d'habitants touchés dans leur environnement	1.000 habitants	0 - 100
	6	Atteinte aux sites	Nombre de sites touchés	1 site (viaduc dans le site de Bougival)	0 - 4
<i>Minimiser les nuisances sur les forêts atteinte au patrimoine, effet sur les promeneurs, perte de caractère des sites forestiers</i>	7	Surface détruite ou gravement endommagée	Nombre d'hectares de forêts dénaturées	Hectare	0 - 500
	8	Effets sur les promeneurs dans le reste de la forêt	Nombre d'hectares de forêts touchés (partie non dénaturée)	Hectare	0-1000
	9	Destruction du caractère des massifs	Nombre de massifs très touchés	1 massif (forêt de la Malmaison)	0 - 2

TABEAU 2
TABLEAU D'EVALUATION

Variantes	"Forêt - Viroflay"			'Bougival'		"Port - Marly"	
	Initial	Tranchée couverte dans Viroflay	Tunnel et tranchée couverte dans Viroflay	Tunnel	Viaduc	Tunnel	Viaduc
1. Qualité de service, Allongement de parcours (en km)	0	0	- 0,6	+ 3,9	+ 3,4	+ 4,9	+ 4,9
2. Urbanisme, Effet sur le développement urbain (note qualitative)	4	4	4	6	6	0 (1)	0 (1)
3. Coût financier (en millions de francs)	511	594	1.100	927	682	1.104	664
4. Nombre d'habitants dans la zone de plus forte nuisance (en milliers)	7,5	5,7	4,1	3,2	4,4	4,0	5,3
5. Nombre d'habitants touchés dans leur environnement (en milliers)	53	40	25	36	36	23	23
6. Nombre de sites touchés	2	2	2	2	3	2	4
7. Nombre d'hectares de forêt (2) dénaturés	400	400	370	60	70	20	20
8. Nombre d'hectares de forêt touchés (partie non dénaturée)	880	880	670	30	30	0	0
9. Nombre de massifs très touchés (Unité = Bois de Malmaison)	2	2	0	0	0	0	0

(1) La note serait 5 si la construction de cette autoroute A 86 n'était programmée qu'après 1985 (cf. S/D/A/U/ de Montesson).

(2) Il y a 3.200 hectares de forêts domaniale à l'intérieur de l'agglomération parisienne.

RESTRUCTURATION DE LA BANLIEUE	CRITERES	INDICATEURS
	Accessibilité aux emplois Accessibilité aux centres restructurateurs	Nombre d'emplois pouvant être atteints à moins de 30 mn. à partir des villes nouvelles Nombre d'actifs situés à moins de 30 mn. d'un centre restructurateur
CIRCULATION AUTOMOBILE	Temps de déplacement en banlieue Accessibilité de l'autoroute	Temps moyen de déplacement en banlieue Nombre d'habitants à moins de 2 km d'un échangeur
COÛT POUR L'ETAT	Coût	Coût

Il est évident que cet ensemble de critères est relativement riche et qu'il pourrait donner lieu à une discussion intéressante portant sur le passage critère → indicateur.

Or il se trouve que dans un premier temps les données dont nous disposions ne portaient que de loin sur cet ensemble de critères. En effet le tableau final que nous avons étudié est une simplification extrême du tableau précédent en ce sens que nous n'y disposons que de 9 indicateurs, dans l'ensemble fort éloignés des 16 précédents.

L'application de l'analyse des correspondances a porté alors sur deux étapes distinctes :

- . Codage des données
- . Analyse factorielle et étude des résultats.

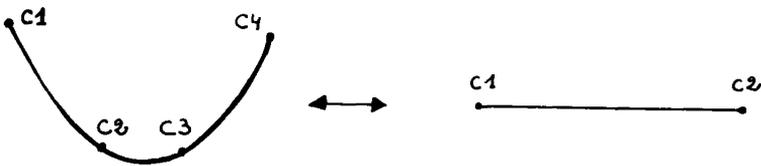
2. Le codage des données (Tableau 2 bis)

Indicateur	Nombre de classes	Codes	Classes
Allongement de parcours (I ₁)	2 classes	SERVB/SERV ^M	Nul/fort
Urbanisme (I ₂)	3 classes	URBAM/URBAB/URBATB	Mauvais/bon/très bon
Coût (I ₃)	2 classes	COUTFA/COUTFO	faible (de 511 à 682)/ fort (de 927 à 1104)
Nombre d'habitants dans la zone de plus forte nuance (I ₄)	4 classes	NUISTB/NUISB NUISM/NUISM	3 à 3,5/3,5 à 4,5/ 4,5 à 6/ - 6
Nombre d'habitants touchés dans leur environnement (I ₅)	3 classes	ENVT B/ENVT MOY/ ENVT M	23 à 25/36 à 40/53
Nombre de sites touchés (I ₆)	3 classes	SITES 2/SITES 3 SITES 4	2 / 3 / 4
Nombre d'hectares de forêt dénaturés (I ₇)	2 classes	DENAT B/DENAT M	< 100 / > 300
Nombre d'hectares de forêt touchés (I ₈)	2 classes	TOUCH B/TOUCH M	< 30 / > 600
Nombre de massifs très touchés (I ₉)	2 classes	MASSIF B/MASSIF M	0/2

Comme nous l'avons déjà précisé celui-ci consiste en un découpage de l'intervalle de variation de chaque indicateur en classes consécutives de probabilité à peu près égales. Vu le petit nombre de variantes en présence (7) nous avons en fait privilégié des classes "visibles" au détriment d'une équiprobabilité illusoire.

Ainsi si l'on étudie de près les histogrammes des neuf indicateurs on est amené à la répartition suivante : indiquée dans le Tableau 2 bis.

On peut noter que la première exigence d'une telle démarche est la constitution des classes les plus fiables possibles : ceci ne peut être fait qu'à l'aide d'une connaissance parfaite des données. En effet on peut par exemple envisager de séparer en deux classes distinctes ce qui n'est dû en fait qu'à une approximation sur la détermination des données. Par ailleurs le nombre de classes correspondant à chaque indicateur a une importance en ce sens que l'on a toujours intérêt lors d'une première analyse à avoir de nombreuses classes : en effet certains effets non linéaires (d'intensité par exemple ; voir les effets Guttman (*)) peuvent être occultés sur des indicateurs à deux classes. On peut ainsi remplacer une répartition parabolique par une répartition linéaire :



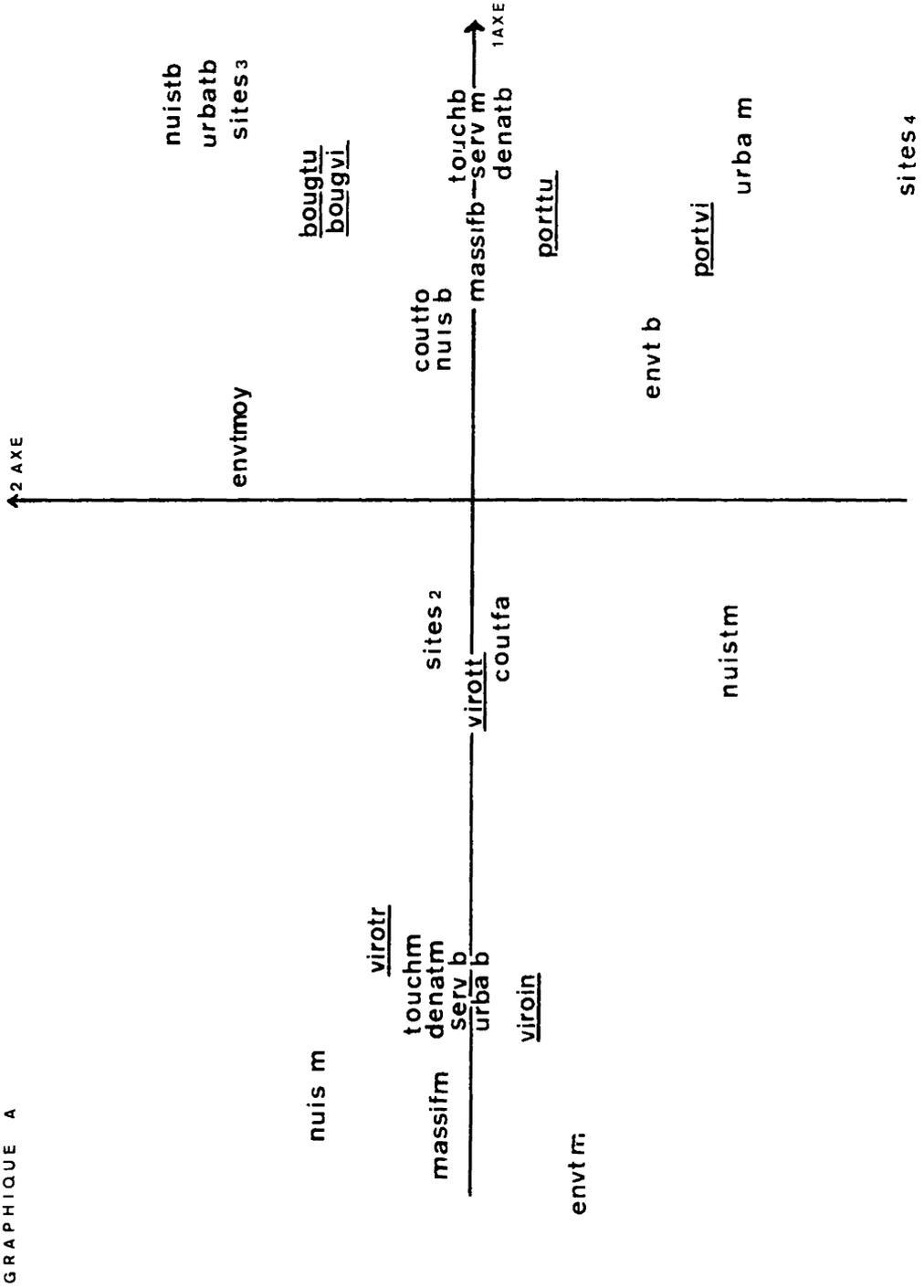
La stabilité de la méthode sera d'ailleurs jugée sur la variation des résultats en fonction de modifications du nombre de classes.

Ceci donne au total 23 classes pour 9 indicateurs et 7 variantes. Le tableau codé binaire des données est alors :

(*) Voir J.P. Benzecri, *op. cit.*, Tome II, pp. 45 et 192 sq.

TABLEAU 3

Classes Projets	I ₁		I ₂		I ₃		I ₄			I ₅			I ₆			I ₇		I ₈		I ₉	
	Bon	Mauv.	Bon	T.B.	Faib.	Fort	3 3,5	3,5 4,5	4,5 6	23 25	36 40	53	2	3	4	<100	>300	< 30	>600	0	2
Forêt Viroflay . Initial Code : VIR0IN	1	0	1	0	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1	0	1
Forêt Viroflay . Tranchée Code : VIRJYR	1	0	1	0	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0
Forêt Viroflay . Tunnel et tranchée Code : VIR0TT	1	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	1	0	1	0	1
Bougival . Tunnel Code : BOUGTU	0	1	0	0	1	0	1	0	0	0	1	0	1	0	0	1	0	1	0	1	0
Bougival . Viaduc Code : BOUGVI	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	1	0	1	0
Port-Marly . Tunnel Code : PORTTU	0	1	1	0	0	1	0	1	0	0	1	0	0	1	0	1	0	1	0	1	0
Port-Marly . Viaduc Code : PORTVI	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0



3. Les résultats de l'analyse factorielle

Les deux graphiques (A et B) ci-joints donnent les projections des 25 indicateurs et des 7 variantes dans les plans des 1er et 2ème, 3ème et 4ème facteurs.

En effet l'étude s'arrêtera au 5ème facteur, les parts de la variance totale expliquée étant respectivement :

Facteur	1	2	3	4	5	
% Variance	41	21	16	10	9	soit 97 %

. Le premier axe

Celui-ci oppose, au vu des tableaux de contribution (non présentés ici), et du graphique A :

		Code
<u>A droite</u>	Nombre d'hectares de forêt dénaturés faible (< 100)	DENAT B
	Allongement de parcours fort	SERV M
	Nombre d'hectares de forêt touchés faible (< 30)	TOUCHE B
	Effet sur le développement urbain très bon (note 6)	URBA TB
	Nombre de massifs très touchés nul	MASSIF B
	Effet sur le développement urbain mauvais (note 0)	URBA M
<u>A gauche</u>	Nombre de massifs très touchés fort (2)	MASSIF M
	Nombre d'hectares de forêt touchés fort (> 600)	TOUCHE M
	Nombre d'hectares de forêt dénaturés fort (> 300)	DENAT M
	Effet sur le développement urbain bon (note 4)	URBA B
	Allongement de parcours nul	SERV B
	Nombre d'habitants touchés dans leur environnement élevé (53.000)	ENVY M
	Nombre d'habitants dans la zone de plus forte nuisance élevé (> 6.000)	NUIS TM

On voit que cet axe oppose très nettement, si l'on regroupe les indicateurs correspondant à un même critère, les points de vue minimisation des nuisances en forêt et niveau de service faible aux points de vue maximum de nuisances en forêt et niveau de service élevé. On est d'ores et déjà confronté à un conflit inter-critériel : ou minimiser l'atteinte aux forêts et par là même allonger le parcours ou tirer au plus droit et traverser le plus de forêts.

Une remarque est à faire sur le critère restructuration urbaine. Nous nous sommes aperçus que figurent à la même extrémité du 1er axe les classes très bon et mauvais.



Cette apparente contradiction provient du fait que les projets ayant obtenu une bonne appréciation pour l'indicateur "restructuration urbaine" (note 4) traversent tous trois des forêts, alors que les quatre autres projets ne traversent pas les forêts, et correspondant au côté droit du 1er axe, obtiennent soit la note 0 soit la note 6 (Mauvais ou Très Mauvais). Implicitement la présence de ce critère sur le 1er axe ne fait que recouper l'interprétation que nous avons donnée du 1er facteur. On ne sera donc pas amené à utiliser ce critère de façon discriminante sur cet axe.

. Le deuxième axe

On y trouve opposés :

<u>En haut</u>	Nombre d'habitants touchés dans leur environnement moyen (36.000 à 40.000)	ENVT MOY
	Structuration urbaine bonne	URBA B
	Nombre d'habitants dans la zone de plus forte nuisance faible (3.200)	NUIS TB
	Nombre de sites touchés moyen (3)	SITES 3
<u>En bas</u>	Nombre de sites touchés élevés (4)	SITES 4
	Structuration urbaine mauvaise (note 0)	URBA M
	Nombre d'habitants dans la zone de plus forte nuisance élevé (> 6.000)	NUIS M
	Nombre d'habitants touchés dans leur environnement faible (23.000 à 25.000)	ENVT B

Si l'on tente de regrouper ces indicateurs en critères, on observe que le critère "nuisances en milieu urbain" se retrouve des deux côtés de l'axe ; ce au travers les indicateurs :

- . Nombre d'habitants dans la zone de plus forte instance,
- . Nombre d'habitants touchés dans leur environnement,
- . Nombre de sites touchés.

Le critère "structuration urbaine" apparaît aussi au niveau de cet axe. On retrouve là les deux critères "urbains", le haut de l'axe étant associé à de bonnes performances sur ces critères, le bas de l'axe correspond à des performances médiocres. On peut alors appeler cet axe, axe urbain, et il paraît évident que nous n'avons pas ici le conflit inter-critériel relevé sur le premier axe : au contraire les deux critères interdépendants se trouvent groupés et c'est leur échelle de valeur interne que l'on retrouve sur ce deuxième axe.

Il nous faut tout de même remarquer une fausse note dans ce consensus : l'indicateur "nombre d'habitants touchés dans leur environnement" agit au sens exactement opposé aux autres indicateurs : ceci doit nous amener à mettre en doute ses capacités à décrire le phénomène nuisances en milieu urbain. Où chercher la véritable signification de cet indicateur sinon sur le premier axe ? En effet nous nous apercevons que les trois classes de cet indicateur se projettent sur le premier axe dans leur ordre naturel : mauvais/moyen/bon de la gauche vers la droite. Or de la signification du premier facteur, il ressort immédiatement que cet indicateur en fait est lié au critère "nuisances sur les forêts" plus qu'au critère "nuisances en milieu urbain". Nous avons affaire ici à un indicateur à double appartenance et dont la véritable place apparaît à l'étude du graphique. De plus si l'on revient à la définition de cet indicateur on note qu'il correspond à un effet de coupure : on compte le nombre d'habitants des communes coupées en deux par l'autoroute. Or le même type de "comptage" a été effectué pour le chiffrage des indicateurs "nombre d'hectares de forêt dénaturés" et "nombre d'hectares de forêt touchés" : il est évident alors que ces indicateurs structurellement identiques se comportent de manière identique.

. Le troisième axe (voir graphique B)

Son interprétation ne pose aucun problème : c'est un axe de coût opposant les variantes à coût faible aux variantes à coût fort. En effet ces points contribuent à 35% de la formation de l'axe (contre 3% en moyenne aux autres). On peut noter d'ailleurs qu'aux "coûts" faibles est associé le maximum de nuisances urbaines, alors qu'aux "coûts" forts est associé le minimum de nuisances sur la forêt. Ceci ne peut que nous confirmer dans la fragilité de l'estimation des coûts dits directs eu égard à la difficulté de calculer les coûts indirects qui sont ici de loin les plus importants. La structure particulière de ce facteur ne peut que nous inciter à ne pas perdre ce problème de vue.

. Les 4ème et 5ème axes

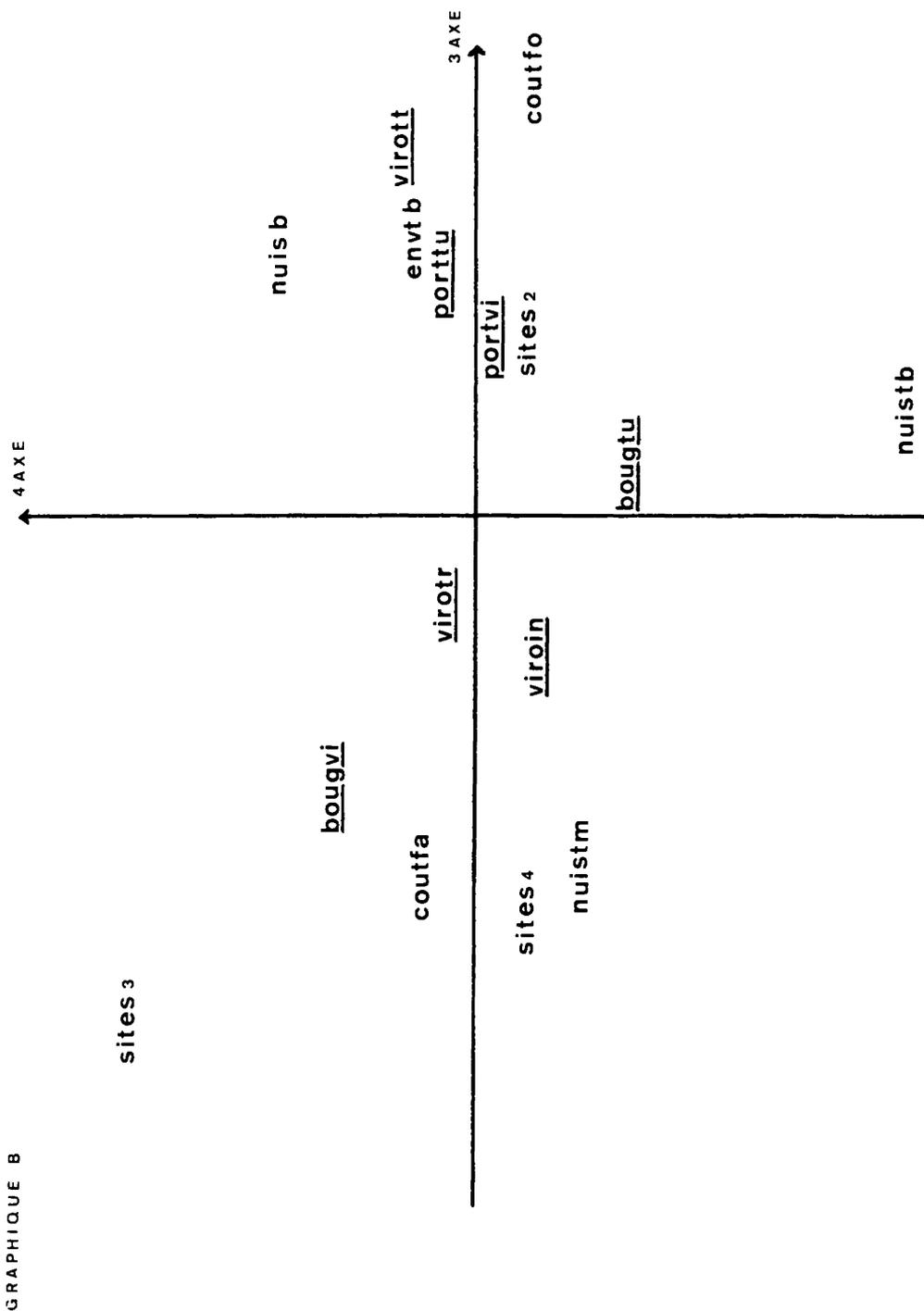
Ces facteurs ne portent, au vu des tableaux de contribution, que sur des distinctions entre variantes particulières :

. Ainsi le 4ème facteur oppose les variantes 4 et 5 (variantes Bougival) qui obtiennent les mêmes scores partout sauf sur les indicateurs "nombre de sites touchés" et "nombre d'habitants dans la zone de plus forte nuisances".

. De même le 5ème facteur oppose les variante 1 et 2 (variantes Viroflay) qui présentent les mêmes particularités vis-à-vis des seuls indicateurs "nombre d'habitants touchés dans leur environnement" et "nombre d'habitants dans la zone de plus forte nuisance".

4. En conclusion

Il nous reste à résoudre, si possible, le problème du choix de la "meilleure" variante. Or nul besoin ici de simuler 36 politiques possibles et de retrouver quasiment toutes les variantes comme "meilleur-choix". Tout se lit ici directement sur les graphiques.



Nous proposons de "simuler" la démarche de la façon suivante (en ne perdant pas de vue que les axes sont "indépendants").

1ère étape :

On peut rejeter sans hésitation le bas du 2ème axe dans la mesure où il correspond à un maximum de nuisances en milieu urbain (solution inadmissible) : ceci ne nous laisse que trois variantes, les deux variantes BOUGIVAL à droite du 1er axe, la variante VIROFLAY TRANCHEE à gauche.

2ème étape :

En examinant le premier facteur, nous nous apercevons que le choix est à faire entre bon niveau de service et destruction des forêts d'une part, mauvais niveau de service et protection des forêts d'autre part.

Or on peut penser par ailleurs que le niveau de service lié à l'allongement de parcours, sera donc lié au coût. On peut donc s'attendre ici à ce que la variante VIROTR soit nettement moins chère que les variantes BOUGIVAL.

Or en nous référant au troisième facteur (puis aux données) nous voyons qu'il n'en est rien. En fait une des variantes BOUGIVAL (BOUG VI) a un coût comparable au coût de la variante VIROTR (682 M contre 594 M) alors que la variante BOUGTV est nettement plus onéreuse (927 M). Cette remarque nous met sur la voie d'un processus de négociation entre les diverses parties concernées :

. Concrétisons celles-ci par les actants "EQUIPEMENT" qui cherche à minimiser les coûts et "SERVICE FORESTIER" qui cherche à minimiser les nuisances sur les forêts.

. On peut alors imaginer le dialogue suivant :

Service Forestier

Je ne veux absolument pas de la catastrophique variante VIROTR. Je préférerais la variante BOUGTU qui épargne le plus les forêts concernées.

Equipement

Il est hors de question d'envisager une variante aussi onéreuse. Je reste sur VIROTR qui coûte moitié moins cher.

Analyste

Puis je vous suggère de voir les mérites de la variante BOUGVI qui coûte à peine plus cher que VIROTR et qui a des performances (bien que légèrement inférieures) comparables à celles de BOUGTU en ce qui concerne les nuisances sur les forêts.

Equipement et (en choeur) Pourquoi pas ?

Service Forestier (à part) Et merci à l'analyse des données.

Il nous paraît intéressant de revenir sur une comparaison entre la méthode que nous proposons et celle communément usitée. L'attitude implicite de l'analyste multicritère traditionnel relève de la mystique dite "fonction d'utilité". En effet vouloir chercher à rendre comparables les échelles d'évaluation des différents indicateurs suppose obligatoirement que ces points de vue différents sont tous réductibles sur une seule dimension. Ce qui n'est rien d'autre que la recherche de la fonction d'utilité du ou des décideurs ayant présidé à l'élaboration des données.

Or il est absurde de tenter ainsi de réduire le multidimensionnel à l'unidimensionnel : il peut paraître pour le moins paradoxal, s'agissant d'un choix multidimensionnel, qu'on ne puisse en fait que l'approximer par une représentation sur une dimension unique. On retrouve là, dans le domaine du qualitatif, les travers des analyses coûts-avantages.

En effet si l'on regarde de près nos graphiques, nous nous apercevons que si l'on peut considérer comme réductibles à une seule dimension (correspondant au 1er axe) les critères "niveau de service", "nuisance sur les forêts" et dans une moindre mesure "coût" (*), il est évident que les critères "restructuration urbaine" et "nuisance en milieu urbain" se déplacent sur une toute autre dimension. Il nous faudrait donc en tout état de cause utiliser au moins deux, sinon trois, dimensions pour décrire le phénomène. Il paraît alors impossible de réduire systématiquement à l'unidimensionnel ce qui n'a aucune raison de l'être.

En résumé nous ne pouvons qu'être confortés dans nos vues, dans la mesure où nous respectons ici la structure du phénomène étudié et où nous remplaçons un instrument de marchandages (taux de substitution, poids) par un instrument de négociations.

(*) Dans la mesure où les points COUTFA et COUTFO se retrouvent dans des positions relatives identiques sur le 1er et le 3ème axe. Mais ceci n'est qu'une approximation.