

TYOLOGIES DES VOYAGES À LONGUE DISTANCE

Christian CALZADA, Samuel BAYOMOG*



Le point fort des enquêtes structurelles ménages transports de l'Insee, réalisées en 1966-1967, 1973-1974, 1981-1982 et 1993-1994 (enquête transports et communications (ETC)) est leur caractère multimodal et leur capacité à fournir une vision d'ensemble cohérente des habitudes et pratiques de voyages des ménages, en particulier à longue distance.

Le dépouillement et l'exploitation des informations recueillies se heurtent aux problèmes classiques de données incomplètes (non-réponses partielles ou totales) et de synthèse intelligente du matériau collecté (construction de typologies). Les méthodes classiques d'analyses multidimensionnelles des données telles que les analyses canoniques et factorielles ont prouvé depuis longtemps leur efficacité à traiter des problèmes de corrélations entre variables au sein d'une population. Cependant ces approches sont intrinsèquement linéaires et ne permettent pas de représenter graphiquement de manière satisfaisante les corrélations entre plus de deux variables croisées. C'est pourquoi d'autres méthodes, dites neuronales, ont été développées. La présente note analyse les principes de ces méthodes et leur apport dans l'exploitation de l'enquête ménages transports.

Description des données utilisées

Sont analysés ci-après les voyages au sens de l'ETC 1993-1994, autrement dit l'ensemble des parcours effectués par une personne (dite B) à plus de 100 kilomètres du domicile. Un voyage est constitué au minimum d'un aller-retour, c'est-à-dire de deux déplacements.

Un voyage est décrit par les attributs suivants : le motif principal associé, le mode d'hébergement principal, sa durée (en nuitées), le mode de transport principal emprunté, la distance parcourue pendant le voyage, l'origine et la destination. En outre pour chaque individu voyageant on dispose des indicateurs socio-démographiques suivants : sexe, âge, condition familiale..., niveau d'études, PCS, strate de résidence ainsi que de variables dites *d'opportunité de voyage* comme l'existence de contacts privés ou professionnels, la prise en charge ou non des frais de déplacements, parc automobile du ménage, etc.

L'analyse porte sur un « échantillon utile » de 18 499 voyages, soit 80% de la totalité de l'échantillon des individus B qui décrivent leurs opportunités de contacts, sur la période allant de mai 1993 à avril 1994, dans un questionnaire rétrospectif de leurs déplacements à longue distance (à plus de 80 km à vol d'oiseau du lieu de domicile) et qui remplissent un questionnaire sur leurs voyages pendant les trois mois qui précèdent la visite de l'enquêteur.

	Nombre de voyages	Nombre d'individus
0 voyage	0	6 145
Entre 1 et 6 voyages	14 563	7 203
+ de 6 voyages	8 499	720
Total	23 062	14 068
Echantillon étudié	18 499	-

VOYAGES

Les méthodes classiques d'analyses des données

Plusieurs analyses multidimensionnelles ont été menées de front afin de dégager différentes typologies permettant de prendre en compte l'ensemble des types de variables descriptives des voyages : attributs du voyage, intensité de mobilité, caractéristiques socio-démographiques des individus mobiles, variables dites d'opportunité de voyage.

Typologie globale des individus effectuant des voyages à longue distance

Cette première approche vise à l'analyse des liaisons entre variables décrivant les voyages (variables dites « actives »), données caractéristiques des individus mobiles ainsi que variables dites « sociales » (variables dites « supplémentaires ») ; elle est menée au moyen d'une analyse des correspondances multiples (ACM) classique et appelle les commentaires généraux suivants :

- le nombre de voitures particulières (parc) varie dans le même sens que le nombre de voyages,
- deux types de pratique des voyages s'opposent : les personnes qui déclarent pratiquer régulièrement la conduite d'une voiture et ceux à pratique occasionnelle ou nulle,
- le nombre de voyages varie en sens inverse de la distance parcourue pour un voyage, autrement dit : plus on voyage, moins on va loin,
- une opposition entre deux grandes zones d'origine des voyages : l'Île-de-France et le reste de la France.

Malgré tout si l'on souhaite disposer d'une vue synthétique de l'ensemble des informations disponibles, la projection des 94 modalités des variables étudiées, comme c'est le cas de l'analyse développée ci-dessus, rend les graphiques vite illisibles.

Apport des méthodes dites neuronales

Les méthodes neuronales présentent quelques avantages :

- contrairement aux techniques classiques, les algorithmes utilisés (Kohonen en particulier) ne nécessitent pas de calculs matriciels importants. Parmi les algorithmes neuronaux, celui d'auto-organisation dit de Kohonen est particulièrement célèbre [voir références bibliographiques], et permet en particulier d'obtenir une carte plane lisible de données multidimensionnelles nombreuses,
- elles constituent tout d'abord un outil d'analyse exploratoire des données (EDA), complémentaire à d'autres outils, ne nécessitant pas la construction d'hypothèses contraignantes *a priori* sur la structure du modèle à tester,
- un autre avantage est de permettre de « révéler » des relations comportementales dites du second ordre. La nature exploratoire de ces techniques permet en cela d'identifier des variables que nous qualifierons de complexes (styles et cycles de vie), dont on sait qu'elles peuvent conditionner des comportements spécifiques de transport et sur lesquelles il s'avère impossible de questionner directement mais dont on peut appréhender les contours à l'aide des variables socio-démographiques usuelles,
- enfin elles permettent la prise en compte de phénomènes de nature intrinsèquement non linéaires, plus proches des notions de cycles de vie, de chaînes d'activité, etc.

VOYAGES

Typologie neuronale des individus effectuant des voyages à longue distance

Dans cette étape on a procédé à l'analyse simultanée des correspondances entre toutes les variables entrant dans l'étude. Vue l'hétérogénéité des variables analysées, cette démarche ne semble pas opératoire mais ce travail a pour but de rassembler dans une même étape la description des types de voyages et la typologie des individus qui les effectuent. On étudie les relations entre toutes les variables au moyen de l'algorithme de Kohonen (KACM), analyse suivie d'un regroupement par classification hiérarchique.

Sur la carte de Kohonen [carte 1], cinq classes de voyages apparaissent, il s'agit :

- des **voyages d'une journée**, qui sont surtout des voyages domicile travail et sont entrepris par des personnes travaillant en Île-de-France, en provenance du bassin parisien. Les personnes qui effectuent ce type de voyage sont des ouvriers, de niveau d'étude CAP ou BEP, des couples avec enfant, et des personnes réalisant au moins six voyages dans l'année.
- des **voyages à moyenne distance**, effectués par des individus appartenant à des ménages possédant au moins une voiture, faisant au moins deux voyages en voiture et résidant soit dans l'Ouest de la France dans des communes de plus de 100 000 habitants soit dans la région parisienne. Dans ce groupe, on trouve également des couples non mariés, des personnes âgées de 26 à 40 ans, d'autres qui conduisent régulièrement et exercent une profession intermédiaire, des hommes ou des membres de ménages avec au moins trois enfants ainsi que des femmes de niveau d'études baccalauréat et des employés. Les voyages s'effectuent principalement dans la région du sud-ouest ou à destination du bassin parisien.
- des **voyages à une distance comprise entre 500 et 1000 kilomètres**, souvent en hébergement chez des amis, qui durent de une à trois nuitées. Ils sont également le fait d'individus disposant d'une résidence secondaire, de personnes âgées de 51 à 60 ans et de couples mariés ou de ménages sans enfant.
- des **voyages de plus de 1000 kilomètres** qui sont en général effectués par des retraités en avion et durent au moins quatre nuitées, à l'étranger en location. Ces voyages ont en majorité pour origine l'Île-de-France (unité urbaine de Paris). Le motif de ce type de voyages est d'ordre privé et sont le fait de personnes déclarant avoir plusieurs contacts privés.
- des **voyages pour motif professionnel** qui correspondent à des voyages de personnes bénéficiant d'une prise en charge (complète ou partielle) et qui déclarent avoir des contacts privés et professionnels. Ces voyages sont également effectués par des personnes de 40 à 50 ans, artisans, commerçants, chefs d'entreprise ou cadres supérieurs.

Identification des types de voyages

Chaque voyage est décrit par un certain nombre de variables dites « principales » : motif, mode d'hébergement, durée, moyen de transport, distance parcourue. On réalise ci-après une classification des modalités de ces variables à l'aide de l'algorithme KACM. La carte n°2 permet d'identifier trois grands groupes de voyages :

- **les voyages à très longue distance** à plus de 1000 kilomètres du domicile, en mode d'hébergement locatif, par avion et de durée supérieure ou égale à quatre nuitées. Ces voyages sont essentiellement d'ordre privé.
- **les voyages à moyenne distance** sur une distance comprise entre 300 et 1000 kilomètres de la zone de résidence chez des parents ou amis, en général en voiture et pour une durée d'une à trois nuitées. Ils sont pour une grande partie d'ordre privé mais peuvent avoir un motif professionnel.

VOYAGES

- **les voyages à faible distance** d'une durée qui n'excède pas la journée, généralement en train. Le motif « principal » de ces voyages est professionnel et souvent de type domicile-travail.

La prise en compte de l'intensité de mobilité

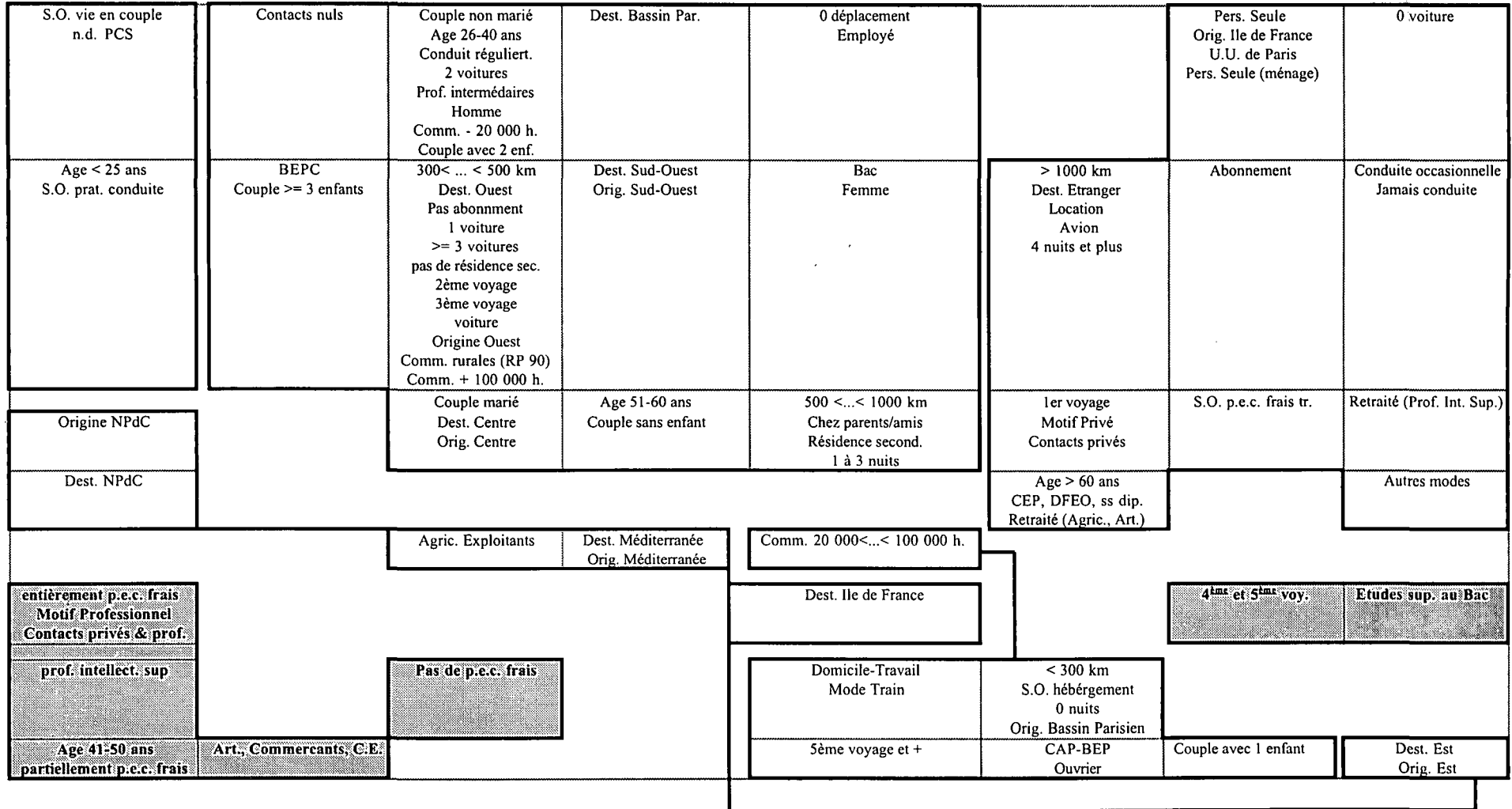
Dans cette partie, on incorpore aux variables décrivant chaque voyage une variable résumant l'intensité de la mobilité (nombre de voyages annuels effectués par l'individu concerné). A cet effet, chaque voyage entrant dans l'analyse a été associé à un numéro d'ordre qui indique que l'individu ayant décrit ce voyage a effectué un nombre de voyages au moins égal à ce numéro. On applique ensuite l'algorithme KACM sur l'ensemble des variables décrivant un voyage augmenté de la variable numéro d'ordre du voyage et on obtient la carte n°3.

On retrouve les trois grands groupes de voyages identifiés dans la partie précédente, la plus grande intensité de mobilité (5^{ème} voyage) étant associée aux distances les moins longues (< 300 km).

Références bibliographiques

- *M. Cottrell, P. Letrémy : Classification et analyse des correspondances au moyen de l'algorithme de Kohonen : application à l'étude de données socio-économiques.* Prépublication du SAMOS n°42 (1995), Université Paris 1.
- *S. Ibou, M. Cottrell : Multiple Correspondence analysis of crosstabulations matrix using the Kohonen algorithm.* Prépublication du SAMOS n°49 (1995), Université Paris 1.
- *F. Blayo, M. Verleysen : Les réseaux de neurones artificiels, Que sais-je ?, n° 3042, PUF.*
- *D. Shmuelli, I. Salomon, D. Sheffer : Neural network analysis of travel behaviour : evaluating tools for prediction, Transportation Research - C, Vol. 4. No. 3, pp. 151-166, 1996.*
- *C. Calzada, S. Midenet : Des neurones dans les réseaux, Notes de Synthèse de l'OEST, N°97, décembre 1995.*
- *N. Gouider : Les Français voyagent de plus en plus, Notes de Synthèse du SES, N°113, septembre-octobre 1997.*
- *N. Gouider, A. Kerihuel : Les voyages touristiques de courte durée en 1994, Notes de Synthèse du SES, N°114, novembre-décembre 1997.*

Toutes les analyses ont été menées sous le logiciel SAS© version 6.11 (Orlando).



Carte n° 1 : Représentation des modalités sur une carte de Kohonen
Algorithme KMCA sur 22 variables (400 itérations)

VOYAGES

VOYAGES

Carte n° 2 : Représentation des modalités par l'algorithme KMCA dans l'analyse des correspondances des voyages (120 itérations)

Motif Professionnel	Mode 'Autres'	Train	< 300 km Domicile Travail 0 nuitées
Location Avion		300 < ... < 500 km	
> 1000 km			500 < ... < 1000 km
4 nuitées ou +	Chez parents/amis	Motif Privé Voiture	1 à 3 nuitées

Carte n° 3 : Représentation des variables de voyages et du nombre de voyages

5 ^{ème} voyage et + Domicile-travail	< 300 km 0 nuitées		3 ^{ème} voyage
Train		Motif Professionnel	
	300 < ... < 500 km 2 ^{ème} voyage		500 < ... < 1000 km Chez parents/amis 1 ^{er} voyage Motif Privé Voiture 1 à 3 nuitées
		Autres modes	
4 ^{ème} et 5 ^{ème} voyage			Avion > 1000 km Location 4 nuits ou +