



**HAL**  
open science

# Une famille de modèles spatiaux et leur application à la matrice des migrants interdépartementaux français pour la période 1968-1975

Denise Pumain, Michel Poulain

## ► To cite this version:

Denise Pumain, Michel Poulain. Une famille de modèles spatiaux et leur application à la matrice des migrants interdépartementaux français pour la période 1968-1975. *Espace Populations Sociétés*, 1985, 3 (1), pp.33 - 42. 10.3406/espos.1985.997. halshs-01488362

**HAL Id: halshs-01488362**

**<https://shs.hal.science/halshs-01488362>**

Submitted on 13 Mar 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# Une famille de modèles spatiaux et leur application à la matrice des migrants interdépartementaux français pour la période 1968-1975

Denise Pumain, Michel Poulain

## Abstract

A family of spatial models applied to the 1968-75 migration matrix of the french "departements".

A family of models is proposed for the analysis of tables showing migration flows between geographical subdivisions. The aim is to use them as a filter, in order to remove, in observed migration flows, the part which is due to variations in size of populations and in geographical configurations (distance, compactness, contiguity) of every couple of zones being considered. The example of migrants having been exchanged between 94 french departments from 1968 to 1975 is given to show the efficiency of those models. The quality of ajustement is very good and the residuals can easily be interpreted in terms of regional variations in mobility behaviours.

---

## Citer ce document / Cite this document :

Pumain Denise, Poulain Michel. Une famille de modèles spatiaux et leur application à la matrice des migrants interdépartementaux français pour la période 1968-1975 . In: Espace, populations, sociétés, 1985-1. Migrations et urbanisation - Migrations and cities. pp. 33-42;

doi : 10.3406/espos.1985.997

[http://www.persee.fr/doc/espos\\_0755-7809\\_1985\\_num\\_3\\_1\\_997](http://www.persee.fr/doc/espos_0755-7809_1985_num_3_1_997)

---

Document généré le 21/03/2016

# UNE FAMILLE DE MODELES SPATIAUX ET LEUR APPLICATION A LA MATRICE DES MIGRANTS INTERDEPARTEMENTAUX FRANÇAIS POUR LA PERIODE 1968-1975

**Michel POULAIN**

(dépt Démographie,  
Université de Louvain-la-Neuve,  
Belgique)

**Denise PUMAIN**

(INED, Paris, France)

## 1. Elaboration d'une famille de modèles spatiaux

La famille de modèles spatiaux de la migration proposée brièvement ci-après et mise au point par l'un d'entre nous (Poulain, 1981) s'inscrit dans l'optique gravitationnelle qui remonte à Ravenstein (1885). A l'exemple de Thomlison (1961), les variables sociologiques et économiques n'y sont pas considérées bien qu'elles soient pertinentes pour expliquer le phénomène migratoire. Seuls sont considérés les facteurs spatiaux les plus significatifs mais également ceux qui sont susceptibles d'affiner au mieux la mesure du phénomène. Dans une seconde phase que nous n'abordons pas ici, les résidus par rapport à cette famille de modèles spatiaux devraient faire l'objet d'une tentative d'explication sur base des variables socio-économiques.

En ce qui nous concerne, trois types de variables seront retenus, ce sont, pour un découpage d'un pays en un nombre déterminé de zones et pour un couple de zones  $i$  et  $j$ :

- les effectifs de population peuplant ces zones  $P_i$  et  $P_j$ ;
- la distance  $d_{ij}$  entre ces zones;
- la configuration spatiale de ces zones (surface, longueur de frontière commune et contiguïté éventuelle).

Pourquoi introduire les effectifs de population dans un modèle spatial de la migration? Tout simplement parce que le découpage territorial adopté pour la mesure du phénomène fixe les effectifs de population des zones. Avec Courgeau (1975), nous introduirons la notion d'intensité de la migration de  $i$  vers  $j$  comme la probabilité qu'un individu tiré aléatoirement en début de période ( $t_0$ ) dans la zone  $i$  soit le même qu'un individu tiré aléatoirement en fin de période ( $t_1$ ) dans la zone  $j$ . On aura, par conséquent:

$$m_{ij} = \frac{M_{ij}}{P_i(t_0) P_j(t_1)} \quad (1)$$

A intensité de la migration  $m_{ij}$  égale, le nombre de migrants  $M_{ij}$  entre les zones  $i$  et  $j$  pour la période ( $t_0, t_1$ ) est donc proportionnel au produit des populations \*.

L'ensemble des modèles spatiaux de la migration fait intervenir la distance  $d_{ij}$  entre zone d'origine et zone de destination comme variable indépendante. De quelque façon

\* La population de  $i$  étant considérée en  $t_0$  et celle de  $j$  en  $t_1$ , ni le produit  $P_i P_j$ , ni le modèle proposé ne seront symétriques en  $i$  et  $j$ . Néanmoins, l'asymétrie sera d'autant plus négligeable que la période d'observation est courte et que les populations concernées évoluent sensiblement de la même façon.

qu'elle soit mesurée, elle peut être considérée à la fois comme une variable descriptive (elle décrit la configuration spatiale des zones et est, par conséquent, largement tributaire du découpage du territoire), et comme un résumé plus ou moins satisfaisant de variables explicatives inconnues (Courgeau, 1981).

Quel type de dépendance choisir ? Réalisant une étude critique et comparative des différentes fonctions appliquées aux mêmes données, Morill (1963, 1967) constate que le type Pareto surestime les migrations à très courte distance, alors que la forme exponentielle les sous-estime. Selon lui, la fonction Pareto-exponentielle se rapproche plus de la réalité mais son estimation est plus malaisée puisqu'elle fait intervenir deux paramètres. En fait, en l'absence de support théorique, aucune des fonctions proposées ne satisfait pleinement à l'ajustement des données empiriques. C'est pourquoi, et nous rejoignons en cela l'option de Stillwell (1977), nous retiendrons une solution composite :  $f(d_{ij})$  aura tantôt la formulation de Pareto, tantôt la forme exponentielle. Ce sont là les formes les plus simples - elles ne font intervenir qu'un seul paramètre - et les plus largement utilisées.

Le calcul des distances  $d_{ij}$  entre chaque partie de zone nécessite le recours à des hypothèses dont la plus simple suppose que tous les individus d'une zone sont concentrés en un seul lieu, qui peut être le chef-lieu de la zone, ou son centre de masse. Ce choix, s'il est acceptable pour des zones relativement éloignées l'une de l'autre, ne l'est plus dans le cas de zones plus proches et, a fortiori, de zones contiguës. Mais le calcul de la distance moyenne de migration entre deux zones n'a pas trouvé de solution simple. Pour ce faire, nous proposons le recours aux mesures de distances minimales et maximales entre zones. Sous certaines hypothèses de répartition de la population nettement moins contraignantes, la distance moyenne adoptée sera donnée par

$$D_{ij} = \frac{\int_{d_{\min}}^{d_{\max}} x f(x) dx}{d_{\min}} \quad (2)$$

où  $f(x)$  est soit la forme de Pareto ou la fonction exponentielle.

La prise en considération de la configuration spatiale des zones peut se faire également par l'introduction d'un indice de contiguïté. Conscient de l'importance numérique des migrations à courte distance on comprend que, lorsque deux zones sont contiguës, la majorité des migrants entre ces deux zones correspondent à des mouvements de courte distance par-delà la frontière commune. Dès lors, il nous semble important :

- de dissocier zones contiguës et non contiguës;
- de prendre en compte la longueur de frontière commune.

Pour ce, on introduira un indice de contiguïté sous forme d'une variable binaire à l'exemple de Weeden (1973) ou, mieux encore, on calculera celui-ci en tenant compte de la longueur de la frontière commune  $F_{ij}$  sous forme d'un indice synthétique considérant également la forme des zones par le biais de leur surface  $S_i$  et  $S_j$  :

$$C_{ij} = \frac{F_{ij}}{4 \sqrt{S_i S_j}} \quad (3)$$

Cet indice sera introduit sous forme additive dans le modèle comme il apparaît dans les équations ci-après (4 à 9).

Somme toute, à partir des deux modèles simples (P = Pareto et E = exponentiel), en combinant les deux alternatives retenues, on obtient les 4 modèles suivants :

Modèles simples	$(P) \hat{M}_{ij} = k P_i P_j d_{ij}^a \quad (4)$	(4)
	$(E) \hat{M}_{ij} = k P_i P_j e^{bd_{ij}} \quad (5)$	(5)
Avec distances minimales et maximales	$(Ia\_P) \hat{M}_{ij} = k P_i P_j d_{ij}^a \quad (6)$	(6)
	$(Ib\_E) \hat{M}_{ij} = k P_i P_j e^{bd_{ij}} \quad (7)$	(7)
Avec indice de contiguïté	$(IIa\_P) \hat{M}_{ij} = k P_i P_j d_{ij}^a (1 + \alpha c_{ij}) \quad (8)$	(8)
	$(IIb\_E) \hat{M}_{ij} = k P_i P_j e^{bd_{ij}} (1 + \alpha c_{ij}) \quad (9)$	(9)

L'estimation finalement proposée sera la moyenne arithmétique de ces quatre derniers modèles : il s'agira d'une estimation optimale au sens de plusieurs modèles, évitant le problème du choix d'une fonction de dépendance par rapport à la distance, et intégrant deux approches indépendantes pour résoudre le problème des zones proches ou contiguës.

## 2. Application à la matrice des migrants interdépartementaux (France, 1968-1975)

Sous les hypothèses que toutes les zones soient peuplées d'individus à comportement moyen identique, que populations d'origine et de destination soient également responsables des courants migratoires  $M_{ij}$ , qu'aucune préférence n'intervienne pour expliquer la destination des migrants et que seule la distance joue un effet négatif, les courants  $\hat{M}_{ij}$  estimés par la famille de modèles comparés aux courants observés fourniront des résidus qui pourront être considérés comme indicateurs de différences dans le comportement moyen des individus par rapport à la norme du modèle.

La fiabilité des courants migratoires étant généralement proportionnelle à la racine carrée de leur importance (Poulain, 1978), l'estimation de ce modèle non linéaire se fera par moindres carrés pondérés à l'aide d'un processus d'itérations successives *sans recours* à la linéarisation par passage au logarithme. La qualité de l'ajustement réalisé par le modèle, ajustement qui correspond en fait à une minimisation du chi carré, se fera par le calcul d'un coefficient de corrélation non linéaire pondéré :

$$R^2 = \frac{\sum \frac{1}{\hat{M}_{ij}} (M_{ij} - \hat{M}_{ij})^2}{\sum \frac{1}{\bar{M}_{ij}} (M_{ij} - \bar{M}_{ij})^2} \quad (10)$$

L'application de ces divers modèles aux migrants interdépartementaux de la période 1968-1975 en France (Tobelem, 1983) a donné les ajustements suivants :

	Modèles simples		Avec distances minimales et maximales		Avec la contiguïté		Modèle global
	(P)	(E)	(P)	(E)	(P)	(E)	
R <sup>2</sup>	0,937	0,916	0,951	0,916	0,961	0,954	0,958

Brièvement, nous constaterons le haut niveau d'explication de ces modèles (plus de 95 % de l'inertie expliquée par trois des modèles) et les performances supérieures des modèles utilisant la formule de Pareto (P) sur ceux de forme exponentielle (E), et de ceux intégrant l'indice de contiguïté sur ceux qui emploient les distances maximales et minimales. Le modèle global réduit de 29 et 53 % la variance résiduelle des modèles simples.

La valeur estimée des autres paramètres ne retiendra guère notre attention, car l'on sait qu'ils ne sont strictement comparables que pour une même population et un même découpage du territoire. Signalons, toutefois, que l'exposant  $a$  de la fonction de Pareto dans le modèle simple vaut  $-0,99$ , tandis que le paramètre  $b$  de l'équation exponentielle simple est de  $-0,00443$ .

Le tableau 1 présente quelques valeurs extrêmes pour les flux observés (1-1) et résiduels (1-2). Les plus forts courants de migration se sont produits entre les départements de la région parisienne, très proches et très peuplés, ainsi qu'entre des départements de

**Tableau I. - Valeurs les plus fortes des courants de migrants observés et des erreurs d'estimation par le modèle global**

**1—1: Les courants les plus forts**

REGION PARISIENNE			PROVINCE		
Départements Origine-destination	Courant observé	Résidu pondéré	Départements Origine-destination	Courant observé	Résidu pondéré
75-92	117 905	+ 34	62-59	63 420	+ 55
75-94	109 010	+ 129	59-62	40 760	— 38
75-93	103 270	+ 68	54-75	27 000	— 16
75-91	72 860	+ 151	69-38	26 955	+ 93
92-78	69 335	+ 231	69-01	20 020	+ 145
75-78	60 470	+ 188	38-69	18 485	+ 29
92-75	60 285	— 109	54-57	18 220	+ 67
94-91	52 490	+ 249	42-69	17 245	+ 9
75-95	50 400	+ 191	57-54	15 645	+ 44
92-91	44 400	+ 140	13-83	15 070	+ 24

**1—1: Les plus forts résidus pondérés ( $M_{ij} - \hat{M}_{ij} / \sqrt{\hat{M}_{ij}}$ )**

POSITIFS (sous-estimation par le modèle)			NEGATIFS (surestimation par le modèle)		
Départements Origine-destination	Résidu pondéré	Courant observé	Départements Origine-destination	Résidu pondéré	Courant observé
75-06	374	17 910	92-75	109	60 285
30-34	341	11 820	93-75	109	41 100
83-29	304	4 950	93-92	105	22 150
07-26	268	11 350	92-93	100	24 820
57-13	256	11 545	75-59	100	6 830
94-91	249	52 490	94-75	96	36 720
84-30	238	6 245	80-59	96	5 475
30-84	221	245	59-80	84	7 395
92-06	206	7 600	92-59	84	2 950
75-12	205	2 815	93-59	82	2 530

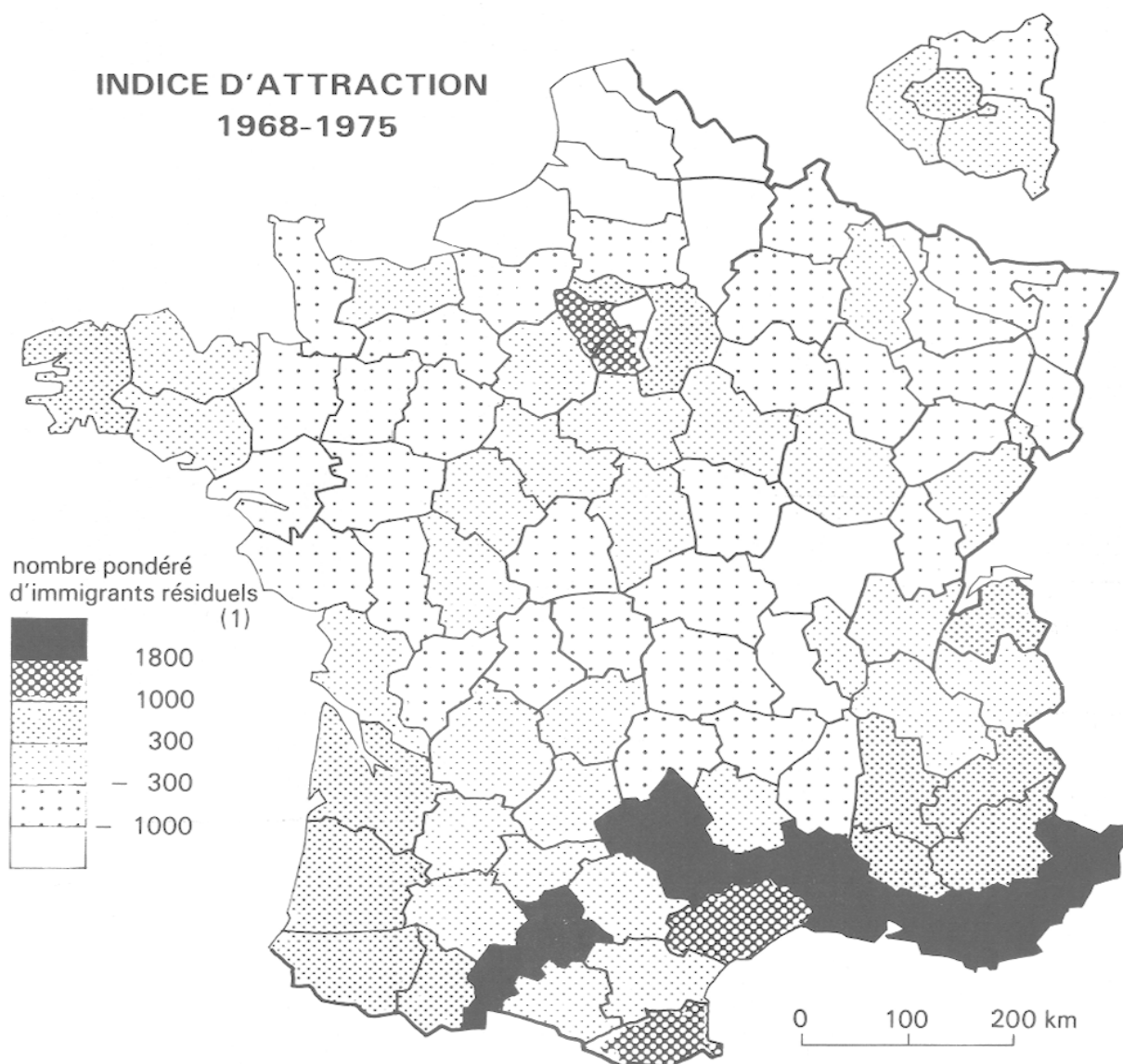
province contigus, aux populations nombreuses (Nord et Pas-de-Calais, départements voisins du Rhône où se situe l'agglomération lyonnaise, départements lorrains, région de Marseille) ou bien encore entre deux foyers de population distants mais très peuplés (de la Meurthe-et-Moselle vers la Seine).

Les plus forts courants résiduels (tableau 1 et figure 1) relèvent d'explications de nature différente mais déjà connues. Les modèles ont ainsi sous-estimé des flux de migrants entre départements très éloignés, qu'il s'agisse de ceux que la Seine ou les Hauts-de-Seine ont envoyés vers des départements méridionaux (Alpes Maritimes, Aveyron...) ou des flux de militaires échangés par les départements où sont localisées des bases navales (Finistère et Var). Mais la sous-estimation concerne également les flux. Parmi les cas où les modèles ont surestimé les flux de migrants entre départements, figurent les courants de la périphérie de la région Ile-de-France vers la Seine, ou des départements de cette région, les moins densément peuplés, vers les départements les plus densément peuplés. On a donc bien, à l'intérieur de cette région Ile-de-France, une nette dissymétrie des échanges tendant à refouler

Illustration non autorisée à la diffusion

Figure 1 : Echanges migratoires résiduels 1968-1975

la population à la périphérie. Un autre cas de surestimation par les modèles concerne les flux à destination des départements du Nord ou de la Lorraine, qui apparaissent ainsi comme répulsifs (la surestimation ne porte que sur l'un des flux). Enfin, d'autres cas de surestimation correspondent à des phénomènes de « frontière » observés depuis longtemps, qui minorent dans les deux sens les courants échangés entre départements alsaciens et lorrains, ou entre départements lorrains et ceux de la région du Nord. Si l'on effectue, pour chaque département, la somme de ces courants résiduels pour les flux de migrants entrés d'une part et sortis d'autre part, on obtient deux indices, respectivement d'attraction et d'émission. Ces indices mesurent le sens et l'intensité globale des écarts entre les effectifs de migrants reçus ou envoyés par chaque département, en éliminant l'effet des masses de population et de sa position par rapport à elles. Les cartes de la figure 2 permettent une analyse fine des zones attractives ou répulsives, en indiquant s'il s'agit d'un excès ou d'un défaut d'entrées et de sorties. Ainsi, la France méridionale et la région parisienne constituent à la fois des zones d'attraction et d'émission fortes, tandis que les départements du Nord, de l'Alsace, des marges armoricaines ou de la bordure nord du Massif Central apparaissent comme peu mobiles.



(1) Somme des écarts entre flux d'entrants observés et flux calculés par ajustement au modèle global.



Illustration non autorisée à la diffusion

**Figure 2: Attraction et émission différentielles**

La *différence* entre ces deux indices donne une mesure proche de celle des taux de migration nette, tandis que leur somme est un indice de mobilité, proche de la mesure de migration totale<sup>1</sup>. Les cartes de la figure 3 montrent que, pour la période 1968-75, et à l'échelon départemental, les zones les plus attractives sont très souvent aussi des zones de forte mobilité, alors que des zones de forte mobilité, comme certains départements de la région parisienne, apparaissent avec des bilans négatifs.

Cependant, plus que dans ces mesures globales, l'intérêt essentiel du modèle nous semble d'être un instrument d'analyse fine des flux entre zones (fig. 1), ou des courants d'entrée et de sortie (fig. 2). Il permet de comparer ces flux en faisant abstraction de la taille et de la configuration des zones sur lesquelles ils sont observés.

---

<sup>1</sup> L'intérêt de ces deux indices est d'être pondérés par la racine carrée du flux estimé, qui rend comparables les écarts dus aux flux entre départements, qu'ils soient proches ou très éloignés.

## REFERENCES

- COURGEAU, D., «L'intensité des changements de catégorie de communes», *Population*, 30, 1975, 81-102.
- COURGEAU, D., *Analyse quantitative des migrations humaines*, 1980, Paris, New York, Barcelone, Milan, Masson, 230 p.
- MORILL, R.L., "The Distribution of Migration Distances", *Papers and Proceedings of the Regional Science Association*, 11, 1963, 75-84.
- MORILL, R.L. et PITTS, F.R., "Marriage, Migration and the Mean Information Field: a Study in Uniqueness and Generality". *Annals of the Association of American Geographers*, 1967, 401-422.
- POULAIN, M., «Du registre de population aux statistiques de migration interne en Belgique: Critique des sources et Corrections des données». *Population et Famille*, 45, 1978, 1-45.
- POULAIN, M., «Contribution à l'analyse spatiale d'une matrice de migration interne». Recherches démographiques, Université Catholique de Louvain, Cabay, 1981, 225 p.
- RAVENSTEIN, E.C., "The Laws of Migration". *Journal of the Royal Statistical Society*, 48, 1885, 167-235 et 52, 1889, 241-289.
- STILLWELL, J.C.H., Some Historical Tests of Spatial Interaction Models for Inter-Area Migration. *Part. I: Inter-Country Migration. Working Paper 185*, School of Geography, University of Leeds, 1977, 70 p. *Part. II: Inter-Region Migration. Working Paper 199*, School of Geography, University of Leeds, 1977, 37 p.

Illustration non autorisée à la diffusion

Illustration non autorisée à la diffusion

**Figure 3: Indices de migration nette et migration totale**

TOBELEM, Ch., Application de modèles d'interaction spatiale à l'étude des migrations interdépartementales en France de 1968 à 1975. *Université de Paris I*, 1983, mémoire de maîtrise sous la direction de D. PUMAIN.

THOMLISON, R., "A Model for Migration Analysis". *Journal of the American Statistical Association*, 56, 1961, 675-689.

WEEDEN, R., "Interregional Migration Models and their Application to Great Britain". *National Institute of Economic and Social Research, Regional Papers* 11, 1973, 43-104.

**ABSTRACT: A FAMILY OF SPATIAL MODELS APPLIED TO THE 1968-75 MIGRATION MATRIX OF THE FRENCH "DEPARTEMENTS"**

A family of models is proposed for the analysis of tables showing migration flows between geographical subdivisions. The aim is to use them as a filter, in order to remove, in observed migration flows, the part which is due to variations in size of populations and

in geographical configurations (distance, compactness, contiguity) of every couple of zones being considered. The example of migrants having been exchanged between 94 french departments from 1968 to 1975 is given to show the efficiency of those models. The quality of adjustment is very good and the residuals can easily be interpreted in terms of regional variations in mobility behaviours.

**KEY-WORDS:** Models, France.

**MOTS-CLES:** Modèles, France.