

Université Ben Youssef Ben Khedda- Alger
Faculté des sciences économiques et de gestion

Thème

Modélisation spatiale du flux migratoire entre les wilayas
d'origine et wilayas de destination
(cas de l'Algérie)

Mémoire de fin d'études

Pour l'obtention du diplôme de magistère en sciences économiques

Option : Econométrie

Réalisation par : CHERFI Kahina épouse FEROUKHI

Sous la direction du : Professeur Mohamed Yassine FERFERA

Jury :

Mme.INAL Myriem, Professeur à l'Université d'Alger : Présidente

Mr. FERFERA Mohamed Yassine, Professeur à INPS /CREAD: Rapporteur

Mr.KHERBACHI Hamid, Professeur à l'Université de Bejaia : Membre

Mr.MASSAR Moncef, Maître de conférences à l'Université d'Alger : Membre

Mr.RAAD Ali, Maître de conférences à l'Université d'Alger : Membre

Année universitaire 2007/2008

Remerciements

Je tiens, en premier lieu, à remercier Monsieur le Professeur Mohamed Yassine FERFERA, pour son aide lors de l'élaboration de ce document, ses remarques judicieuses, sa confiance et sa grande disponibilité.

Je le remercie aussi de m'avoir procuré au sein du CREAD les conditions propices à l'accomplissement de ce modeste travail.

Mes remerciements vont également aux membres de mon jury, Madame la présidente, Professeur INAL Myriem, Monsieur le Professeur KHERBACHI Hamid, ainsi que Messieurs, MASSAR Moncef et RAAD Ali, les rapporteurs de mon mémoire qui ont bien voulu accepter cette tâche. Mon travail ne peut que bénéficier pleinement de leurs suggestions pertinentes.

Je remercie de même Monsieur Yacine BELARBI qui m'a fait entrevoir un monde fascinant, celui de l'économétrie spatiale. J'ai bénéficié de ses précieux conseils, sa rigueur et de son ardeur au travail. Je le remercie également pour le suivi régulier de mon travail, ainsi que sa disponibilité et sa grande gentillesse.

Je tiens enfin à remercier mes collègues du CREAD qui n'ont cessé de me renouveler leurs confiances et leurs encouragements.

Dédicace

Je dédis l'intégralité de ce mémoire à ma défunte tante maternelle Djamila, décédée en 2008, que dieu lui accorde sa miséricorde et l'accueille dans son vaste paradis.

Tables des matières	page
Introduction générale	01
CHAPITRE I : L'économie géographique et les modèles gravitaires	
Introduction	06
I-1 Le modèle de Von Thünen appliqué au cas d'une ville	06
I-2 Le modèle de Marshall dans un système local de production	08
I-3 Les modèles de concentration	08
I-3-1 La force de répulsion	09
I-3-1-1 Le modèle de Christaller et Lösch : Centralité et périphérie	09
I-3-2 la force d'attraction	10
I-3-2-1 Le modèle de Hotelling : la concurrence monopolistique localisée	10
I-4 Application des modèles gravitaires	11
I-4-1 Les lois de RAVENSTEIN (fin XIXe)	11
I-4-2 L'analogie gravitaire (PARETO, STEWART début XXe siècle)	12
I-4-3 La remise en cause de l'analogie gravitaire (fin XXe)	13
I-4-3-1 L'explication économique (ZIPF)	13
I-4-1-2 L'explication sociologique (S. STOUFFER)	13
I-4-1-3 L'explication psychologique (T. HAGERSTRAND)	13
I-5 Les trois hypothèses fondamentales pour modéliser les flux	14
Conclusion	15
CHAPITRE II : Introduction aux modèles d'économétrie spatiale	
Introduction	17
II-1 L'état de l'art de l'économétrie spatiales	18
II-2 Introduction aux modèles d'économétrie spatiale	19
II-2-1 Les effets spatiaux : définition et detection	19
II-2 -2 Autocorrélation et hétérogénéité spatiales	20
II-2 -2-1 Autocorrélation spatiale	20
II-2 -2-1-1 Autocorrélation spatiale et modèles économétriques	21
II-2 -2-1-2 Définition de l'autocorrélation spatiale et modèles économétriques	22
II-2-2-2 L'hétérogénéité spatiale et ses deux facettes	23
II-2-3 Les matrices de poids et les variables spatiales décalées	23
II-2 -3-1 Les matrices de contiguïté	24
II-2 -4 Le modèle autorégressif spatial : variable endogène décalée	25
II-2 -4 Modèle avec autocorrélation spatiale des erreurs	27
II-2 -4-1 Les erreurs suivent un processus autorégressif	27
II-2 -5 Estimation des modèles spatiaux	28
II-2 -5 La non-convergence des MCO et ses conséquences	28
II-3 Méthodes d'estimation Bayésienne	30
II-3-1 Introduction à l'approche Bayésienne	31
II-3-1-1 Principes de l'économétrie Bayésienne	31
II-3-1-2 Choix de distributions a priori	33
II-3-2 Définition des méthodes de Monte Carlo par Chaînes de Markov	34
II-3-2-1 Méthodes de simulation classiques	35
II-3-2-1-1 Simulation des méthodes MCMC	36
II-3-2-1-1-a Définition d'une chaîne de Markov	36
II-3-2-1-1-b Algorithme de Metropolis-Hastings	36
II-3-2-1-1-c Echantillonneur de Gibbs	37
Conclusion	38

Chapitre III : Structure migratoire et spatiale de l'Algérie	page
Introduction	40
III-1-1 Migration internationale	40
III 1-2 Migration nationale	40
III 1-3 Evolution de la population urbaine en Algérie de 1886 à 1998	41
III 1-3-1 Répartition de la population selon les trois grands ensembles de l'espace physique Algérien	41
III 1-4 La logique de la migration internationale de la population algérienne	43
III-1-5 La migration inter-wilayale de la population algérienne	44
III-1-5-1 Application du flux spatiale Origine-Destination	45
III 2 Structure spatiale de l'Algérie	48
III 2-1 Le découpage du territoire algérien de 1515 à nos jours	48
III 2-2 Le découpage du territoire algérien durant la période coloniale (1830-1962)	49
III 2-3 Les réorganisations territoriales de la période post-indépendance	50
III 2-3-1 Les réformes administratives de 1974	50
III 2-3-2 Les réformes administratives de 1984	51
Conclusion	52
Chapitre IV : Application des outils de l'économétrie spatiale à la modélisation des flux Origine_Destination	
Introduction	54
IV champs d'application du modèle économétrique inter wilaya O-D	55
IV-1 Origine-Destination : Notation et ordre	55
IV-2 Les modèles gravitaires avec indépendance des observations	56
IV-3 La dépendance spatiale dans les flux Origine-Destination	59
IV-4 Les spécifications du modèle spatial pour les flux Origine-Destination	62
IV-5 Estimation du modèle spatial des flux	63
IV-5-1 Estimation des paramètres et calcul des moments	65
IV-5-2 Log-déterminant pour une matrice de poids	65
IV-5-3 Log-déterminant pour le modèle du filtrage successif	65
IV-6 Application des outils de l'économétrie spatiale au modèle O-D	67
IV-6-1 Spécification du modèle économétrique spatiale O-D appliqué au flux migratoires	67
IV-6-2 Techniques d'estimations du modèle économétrique O-D	69
IV-6-3 Simulation du modèle O-D	69
IV-6-4 Interprétation des résultats du modèle O-D	70
IV-6-4-1 L'emploi	70
IV-6-4-2 Le taux de chômage	72
IV-6-4-3 Niveau d'instruction supérieur à la 3AS	73
IV-6-4-4 Population urbaine	74
IV-6-4-5 Population née dans sa wilaya de résidence	75
IV-6-4-6 La densité de population	76
IV-6-5 Interprétation globale des résultats du modèle O-D	77
Conclusion	78
Conclusion générale	80
Bibliographie	83
Annexes	86

Table des figures	page
Figure 1.1 : Localisation optimale de plusieurs cultures selon Von Thünen	07
Figure 1.2 : système local de production	08
Figure 1.7 Hypothèse 1	14
Figure 1.8 Hypothèse 2	14
Figure 1.9 Hypothèse 3	14
Figure 4.1 L'emploi agricole / 10 premiers flux d'entrées de wilayas	71
Figure 4.2 L'emploi BTP / 10 premiers flux d'entrées de wilayas	71
Figure 4.3 L'emploi agricole / 10 premiers flux de sorties de wilayas	72
Figure 4.4 Taux de chômage / 10 premiers flux d'entrées de wilayas	73
Figure 4.5 Niveau d'instruction supérieur à la 3AS/ 10 premiers flux de sortie de wilayas	73
Figure 4.6 Population urbaine/ 10 premiers flux de sortie de wilayas	74
Figure 4.7 Population urbaine/ 10 premiers flux d'entrées de wilayas	74
Figure 4.8 Population née dans sa wilaya de résidence/ 10 premiers flux d'entrées de wilayas	75
Figure 4.9 Taux de retour	75
Figure 4.10 Densité de population / 10 premiers flux de sortie de wilayas	76

Liste des tableaux	page
Tableau 1.1 : Modèle de Newton vs modèle de Stewart	12
Tableau 2.1 : Les définitions de la contiguïté pour une configuration spatiale régulière	24
Tableau 2.2 : Exemples de lois conjuguées naturelles pour des lois appartenant à la famille exponentielle	32
Tableau 3.1: Evolution de la population urbaine en Algérie de 1886 à 1998	41
Tableau 3.2 : Structure géographique de l'Algérie.	42
Tableau 3.3 : Flux migratoires (1987-1998)	45
Tableau 3.4 : Les trois principaux flux migratoires	46
Tableau 3.5 : Évolution des circonscriptions administratives en Algérie	49

Liste des cartes	page
Carte 3.1 : Flux migratoires O-D de 1987 à 1998	47
Carte 3.2 : Evolution des wilayas (1974 -1984)	51

Introduction générale

Introduction

L'histoire de la pensée économique montre que la science économique standard¹ s'est développée pendant longtemps en ignorant l'espace. Dans cette perspective aspatiale, les phénomènes économiques sont appréhendés comme se déroulant tous en un seul point, dans une sorte de « pays des merveilles sans dimension », selon l'expression de W.Isard (1956).

Néanmoins, si on peut faire abstraction plus au moins facilement de l'espace quand on raisonne sur des agrégats macroéconomiques, la démarche apparaît moins satisfaisante au niveau microéconomique où les explications sont fondées sur le comportement d'unités de division individuelle, principalement les ménages et les firmes, et sur les interactions entre ces unités qui sont par définition localisées en des lieux différents.

Le mouvement de la population est à l'origine de la transformation de l'espace, ce mouvement appelé communément migration, peut être dû à des logiques économiques, démographiques ou politiques. La décision d'émigrer revient à comparer le revenu présent et futur des deux pôles de l'alternative : rester sur place, partir ailleurs. Car l'une des conséquences de la migration interne est la modification entre autres, de la composition par âge, par genre et du taux d'analphabétisme, des régions de départ et d'arrivée.

Pour les régions de départ cela conduit entre autres, à une érosion démographique (vieillesse de la population).

Pour les régions d'accueil, cela stimule la croissance démographique et la consommation, mais la migration peut aussi contribuer à la création de problèmes économiques et sociaux.

Pour capter les migrations internes de la population algérienne (échanges migratoires inter-wilaya), la mesure des flux migratoires s'impose, mais cette mesure pose un problème méthodologique, et cela tient au caractère spécifique de la migration qui, à la différence des autres phénomènes démographiques, fait intervenir deux populations, la population de départ, et celle de l'arrivée. Ainsi, il a été notamment demandé à toutes les personnes recensées en juin 1998 où elles résidaient à la date du précédent recensement de 1987 (on ne prendra en compte que les changements de wilaya de résidence).

Cependant notre travail ne réduit pas l'analyse de la migration au simple calcul d'indices de migration et d'immigration, mais utilise des équations et modèles mathématiques visant à montrer dans quelle mesure les flux migratoires dépendent d'un nombre plus ou moins grand de variables. Ces modèles dynamiques de flux migratoire sont des processus évolutifs : partant d'une répartition initiale donnée de population, on va étudier son évolution dans le temps sous certaines conditions².

Néanmoins, l'utilisation de ces modèles porte atteinte à l'hypothèse de Gauss-Markov utilisée dans la modélisation de la régression. En ce qui concerne la dépendance spatiale entre les observations, rappelons que Gauss-Markov suppose que les variables explicatives sont

¹ La tradition néoclassique.

² Courgeau, 1970

indépendantes. Or la dépendance spatiale viole cette hypothèse. De même, l'hétérogénéité spatiale viole l'hypothèse de Gauss-Markov selon laquelle une seule relation linéaire existe à travers l'échantillon de données. Les formes fonctionnelles et les paramètres varient selon leurs localisations et ne sont donc pas homogènes.

D'autres procédures d'estimation sont donc nécessaires pour mener à bien ce type de modèle, qui prennent en compte les caractéristiques particulières de chaque localisation de l'échantillon.

L'objet de ce mémoire vise donc à apporter des éléments de réponses sur l'impact des migrations sur la répartition spatiale de la population, avoir une meilleure compréhension sur le lien qu'entretient l'espace (localité géographique) sur les flux migratoires et fournir une méthode spécifique de régression spatiale qui puisse être appliquée à des modèles à interaction spatiale, qui se réfère à l'absence d'indépendance entre des observations géographiques (l'autocorrélation), et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace. Car omettre à tort une autocorrélation spatiale, produit des estimateurs non biaisés mais inefficients et l'inférence statistique basée sur les méthodes des moindres carrés ordinaires (MCO) est biaisée.

C'est ce que nous avons démontré dans ce présent travail, car l'application des flux migratoires entre les 48 wilayas, a démontré l'existence des flux origine-based et destination-based, invalidant de ce fait même, l'utilisation des estimations des moindres carrés. On a démontré également, qu'en Algérie il n'y avait pas de migration d'élite et que la migration se faisait au profit des wilayas les plus urbaines et à fort taux d'emploi dans le BTP et ne se faisait pas au profit de wilayas à fort taux d'emploi agricole et à fort taux de chômage.

Ce document est structuré en quatre chapitres, le premier chapitre intitulé « L'économie géographique et les modèles gravitaires » car les premiers modèles d'analyse des flux migratoires se réfèrent à cette théorie et prennent en compte la distance entre les points de départ et le point d'arrivée. Le pouvoir de répulsion du point de départ et celui d'attraction de la zone d'arrivée, mais la prise en compte de la distance dans le modèle n'a pas suffi à saisir avec précision le phénomène de mobilité spatiale, il a fallu attendre l'élaboration d'autres modèles plus complexes en économétrie spatiale : c'est l'objet du chapitre 2 intitulé « Introduction aux modèles d'économétrie spatiale » qui est une introduction aux modèles d'économétrie spatiale, c'est par cette approche spatiale que l'on a estimé les modèles spatiaux tout en tenant compte à la fois des problèmes d'hétérogénéité, des observations extrêmes, et en modélisant les interactions spatiales à partir de fondements économiques.

Dans le chapitre 3 intitulé « Structure migratoire et spatiale de l'Algérie », on a retracé la structure migratoire et spatiale de l'Algérie, car le maillage actuel du territoire algérien était marqué par une volonté d'un développement harmonieux du territoire, volonté qui s'est exprimée notamment par le biais des politiques de promotion administrative et industrielle ;

élaborées dans le but d'avoir une action en matière d'aménagement plus ciblée, et davantage en cohérence avec les préoccupations et attentes des collectivités locales.

Cependant, notre étude sur les mobilités spatiales dans le territoire national faite avec ArcGis (logiciel de cartographie), a démontré que le but fixé par les politiques d'aménagement et d'harmonisation du territoire n'est pas encore atteint, car la redistribution de la population sous l'effet des migrations se fait au profit d'un nombre limité de wilaya accentuant de ce fait, les déséquilibres existants.

Enfin dans le chapitre 4 intitulé «Application des outils de l'économétrie spatiale à la modélisation des flux Origine-Destination » nous avons appliqué les modèles économétriques spatiaux aux flux Origine-Destination, ce qui nous a conduit à de nouveaux résultats en développant des techniques de simulation sous le logiciel Matlab, pour rendre faisable l'estimation bayésienne et l'estimation du maximum de vraisemblance pour des données de flux O-D. Ceux-ci exploitent la structure naturelle de la famille des modèles gravitaires, tout en appliquant la régression spatiale, qui nous a permis d'avoir une inférence et une estimation des flux O-D.

CHAPITRE I

L'économie géographique et les modèles gravitaires

Introduction

David Ricardo (1823) avait démontré que la rente foncière était établie en fonction de la fertilité des sols, c'est-à-dire du potentiel agricole. Or, le modèle de VON THUNEN (1826) a par la suite clairement illustré que la rente foncière était aussi liée à la distance du marché. Distance directement associée à des coûts de transport à assumer par les producteurs et par le fait même, par les consommateurs.

Plus les activités économiques s'éloignent du centre (marché), plus les coûts de transport sont élevés et moins le propriétaire du sol ne peut exiger une rente importante pour l'utilisation de son bien. Cette règle influence la localisation des cultures agricoles ou encore le type des activités qui utilisent le sol.

S'appuyant sur les premières réflexions des précurseurs des économistes (mercantilistes) sur la contrainte de la distance dans le fonctionnement de l'économie, VON THUNEN s'est intéressé à la mesure précise de cette composante spatiale. Il en a formulé une véritable loi économique qui explique, par des cercles concentriques autour des lieux de marché, l'effet de la distance sur la localisation de divers types d'exploitation agricoles et aussi, inévitablement, sur la rente foncière.

I-1 Le modèle de Von Thünen appliqué au cas d'une ville

Le temps constituait de longue date un des points centraux de l'analyse économique, il faudra attendre Von Thünen, au début du XIX^{ème} siècle, pour voir naître la première véritable théorie économique il propose en 1826 une réflexion moderne sur l'organisation de l'espace rural qui est généralisé aujourd'hui à l'organisation urbaine.

Il établit, sur un espace aux propriétés particulières, que la localisation des productions agricoles obéit à une logique de maximisation de la rente foncière, nette des coûts de transport des marchandises vers le consommateur. Car la rente foncière est un revenu touché par un producteur, attribuable à sa seule localisation, cette logique trouve sa traduction dans l'espace rural sous la forme de cercles concentriques admettant pour centre la ville, point d'écoulement unique des denrées agricoles. Formellement, le système de localisation optimal de Von Thünen peut être résumé ainsi :

Les producteurs utilisent les mêmes techniques de production et la fertilité de la terre étant partout la même, le coût de production est identique en tous point de l'espace. Si l'on admet que l'écoulement de la production se fait en totalité sur le marché urbain qui concentre tous les consommateurs, seul varie le coût de transport T , qui est fonction de la distance d la ville des lieux de production. On suppose que T varie linéairement avec la distance de telle sorte que $T=ad$. Quant à la rente foncière, elle représente la différence entre le prix de vente P , avec la distance. La rente foncière $R=P-C-T=b-ad$ (avec $b=P-C=cst$) décroît donc linéairement avec la distance.

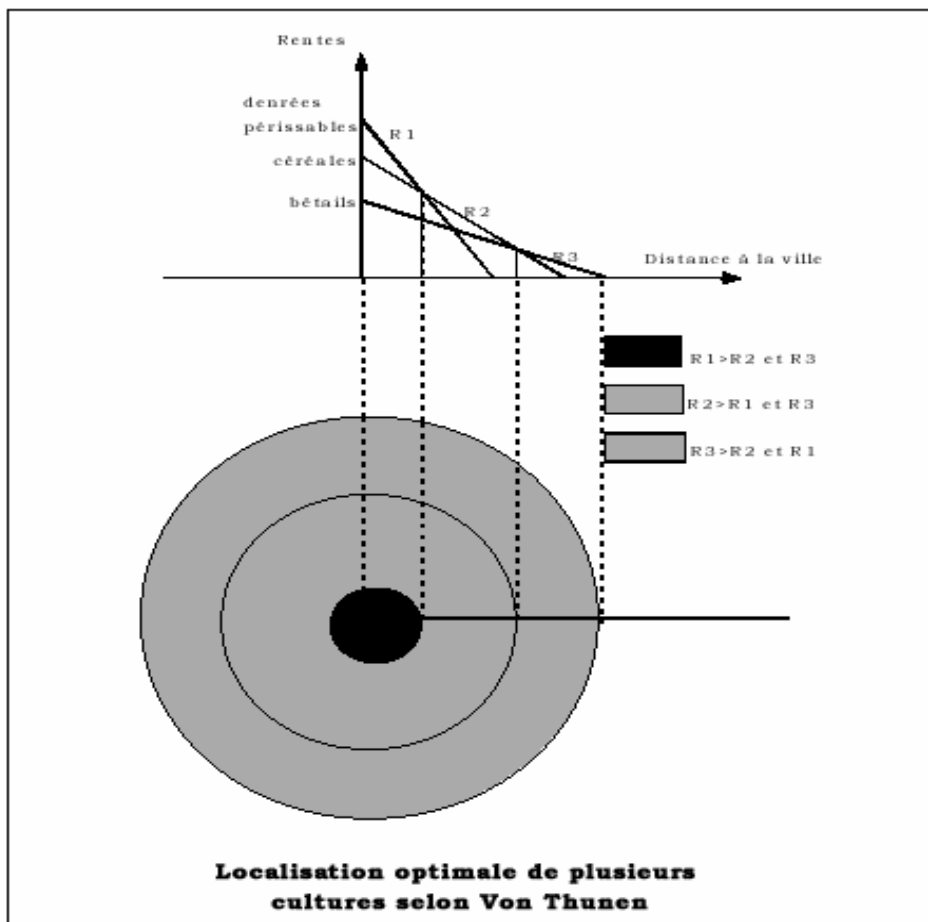
Au-delà d'une certaine distance critique, les frais de transport annulent la rente foncière, la culture envisagée sera donc abandonnée. La rente est évidemment la plus élevée possible pour les terres immédiatement voisines de la ville et s'annule lorsque $P=C-ad$. La culture s'inscrit

à l'intérieur d'un cercle dont le centre est occupé par la ville et s'annule lorsque $P=C-ad$. La culture s'inscrit à l'intérieur d'un cercle dont le centre est occupé par la ville et le rayon est donné par la valeur de d qui annule l'expression de la rente, soit $d = (P-C)/a$.

En généralisent ce résultat à plusieurs cultures se différenciant selon le caractère intensif ou extensif de la production. Von Thünen parvient à généraliser le résultat selon lequel l'espace rural possède une structure multicentrique ou les cultures s'éloignent de la ville au fur et à mesure que les cultures deviennent extensives, comme le montre le schéma ci-dessous.

Aujourd'hui ce type de modèle est utilisé et développé dans le cadre de la nouvelle économie urbaine où l'on déduit les formes urbaines issues de la localisation des activités et des ménages d'un arbitrage entre le prix du foncier et le revenu des ménages.

Figure 1.1 : Localisation optimale de plusieurs cultures selon Von Thünen



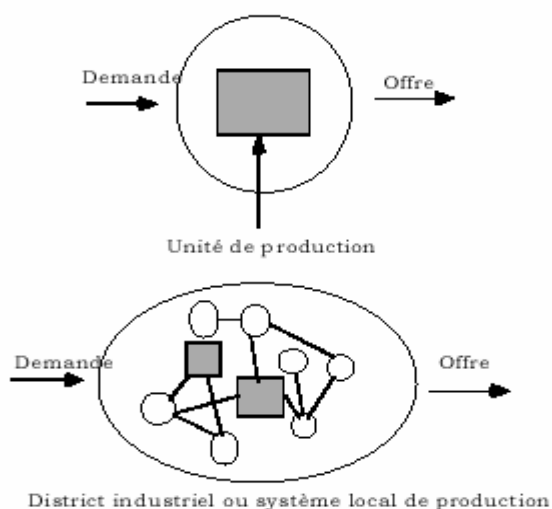
Source : M.Y. FERFERA « Analyse spatiale et régionale » cours polycopier, faculté de sciences économiques d'Alger et de Tizi-Ouzou, 1987

I-2 Le modèle de Marshall dans un système local de production

L'ouvrage de Marshall, en 1890 « les principes de l'économie », insiste sur la notion de district industriel, défini comme un regroupement d'entreprises et un marché du travail spécialisé localisés dans une aire géographique. Alors que l'efficacité de la production était présentée par les économistes comme la conséquence d'une division du travail et d'une croissance au sein de la firme, Marshall montre que cette efficacité peut être acquise au sein d'un espace. A la recherche d'économie d'échelle localisée sur un espace. Par définition l'économie d'échelle est le fait d'obtenir une diminution des coûts de production par unité, grâce à une augmentation des quantités produites. Par exemple, plus on produit un bien x, plus on va répartir sur un nombre important d'unité les coûts et ainsi on réduit le coût unitaire de production de ce bien.

Lorsque l'on dépasse le cadre de la firme dont la taille est souvent positivement corrélés aux problèmes d'organisation pour l'intégration des activités, et que l'on passe au niveau territoire regroupant plusieurs entreprises en situation d'interdépendances productives et techniques et dont la structure est plus flexible.

Figure 1.2 : système local de production



Source : M.Y. FERFERA « Analyse spatiale et régionale » cours polycopier, faculté de sciences économiques d'Alger et de Tizi-Ouzou, 1987

Selon Marchal ces districts industriels présentent une atmosphère industrielle qui favorise la maîtrise et le développement des compétences au niveau local et est une source continue d'innovation. Un siècle plus tard, les économistes s'interrogent encore sur les mécanismes formels qui donnent lieu à l'émergence de districts suivant un processus de concentration.

I-3 Les modèles de concentration

Pour expliquer le processus de concentration qui pousse les acteurs économiques à se rapprocher les uns des autres, il faut partir de deux forces de localisation dans l'espace. La première correspond à la force de répulsion et l'autre à la force d'attraction.

I-3-1 La force de répulsion

La force de répulsion est une distance maintenue entre les fournisseurs de service afin de ne pas se concurrencer.

I-3-1-1 Le modèle de Christaller et Lösch : Centralité et périphérie

Le modèle de Christaller et Lösch explique que les fournisseurs de service maintiennent entre eux une distance égale à deux fois leur portée limite fin de ne pas se concurrencer. Cela aboutit à une distribution régulière dans l'espace. L'idée mise en avant c'est la volonté de chaque acteur d'établir un monopole spatial.

La théorie est fondée sur la distinction entre centre qui sont le siège d'une offre de biens et de services, et des périphéries (qui viennent compléter le centre) où réside la demande, la population utilisatrice, les consommateurs.

La notion de centralité justifie le regroupement en un même lieu de la production des services de même niveau et de même portée destinés à la population dispersée, répartie dans la région complémentaire (la périphérie) qui est sous influence du centre.

Le modèle de Christaller et Lösch montre qu'il y a une organisation spatiale qui repose sur un réseau de ville à distance équivalente avec autour de chaque centre une périphérie sous influence. Cela aboutit à un maillage de l'espace qui se structure sur un réseau de villes, un réseau hiérarchisé entre villes et entre centres et périphéries.

Cependant ce modèle repose sur deux hypothèses qui ne sont pas toujours vérifiées à savoir :

1. l'hypothèse d'un comportement rationnel des consommateurs qui fréquentent le centre le plus proche.
2. hypothèse d'une concurrence entre les centres qui se partagent la clientèle.

Ce qui a pour conséquence une régularité de l'espacement des villes et par une hiérarchisation par niveau de services. Or dans la réalité il est bien évident qu'on ne trouve pas cela.

Le problème malgré tout c'est que ce modèle ne tient pas compte des transformations économiques de la production et de la distribution, deux éléments qui abolissent de fait le lien de proximité producteur / client sur lequel repose justement la théorie des lieux centraux.

Le système capitaliste évolue, se transforme sans cesse, notamment du fait de la concurrence qui en est un des moteurs pour maximiser les parts de marchés et les profits. Car le but du système capitaliste est d'accroître sans cesse les parts de marché

et c'est là où les logiques de rapprochement interviennent et renversent la logique d'ensemble de la localisation.

I-3-2 la force d'attraction

La force d'attraction pousse les marchands à se placer au même endroit, de préférence au centre du marché, précisément pour gagner des clients sur leurs concurrents.

I-3-2-1 Le modèle de Hotelling : la concurrence monopolistique localisée

Le modèle de Hotelling (1929) cherche à expliquer comment cette concentration peut prendre trois formes essentielles qui renvoient en fait à la même explication : augmenter les gains de productivité.

1) principe d'économie d'échelle : qui détermine les rendements décroissants de la production. C'est-à-dire, que l'efficacité économique d'une grosse unité de production peut être meilleure que celle de plusieurs petites unités.

Il y a moins de sites de production mais ils sont plus importants en taille afin de réaliser des économies d'échelles.

On retrouve cette approche dans la géographie de l'économiste Paul Krugman par exemple qui mêle dans son approche : localisation des entreprises, coût du transport et ce qu'il appelle « l'effet du marché interne », autrement dit l'intérêt à être à proximité du principal marché pour limiter les coûts de transport.

La résultante spatiale des économies d'échelle est la création d'usines colossales, de grande capacité.

2) principe d'externalités : se sont les avantages procurés à une entreprise sans qu'elle n'ait besoin d'agir ou de payer directement pour les obtenir. C'est-à-dire sur le fait que les entreprises ont intérêt à se situer près d'autres entreprises afin de tirer avantage d'une localisation dans une même zone où seront déjà présents des réseaux routiers ainsi qu'un bassin de main d'oeuvre.

Tout ceci permet à un acteur économique de profiter de l'action des autres acteurs sans que cela ne lui coûte. Et toutes ces externalités sont liées à un lieu, à la proximité et à la présence d'autres acteurs, ce qui aboutit à créer des économies de localisation.

3) Economies d'urbanisation : Les unités de production ont intérêt à se situer dans des espaces où les activités et la population sont concentrées. Les fortes densités garantissent la présence d'une main d'oeuvre abondante, d'un marché important, mais aussi d'infrastructures de bon niveau.

En dessous d'une masse de population (ou de production) certaines activités économiques ne sont pas rentables. Il faut donc un certain seuil pour que l'intérêt de la

concentration soit effectif. Ce seuil varie mais il est, par exemple, généralement de plusieurs centaines de milliers d'habitants pour une université, de plusieurs millions pour un aéroport international...etc.

En conséquence les formes spatiales de cette concentration sont celles des métropoles, des grands bassins de population.

Finalement de ces trois types de concentration, on observe un processus de combinaison entre les trois, ces formes se conjuguant et se renforçant mutuellement, cela donne des espaces où l'on trouve à la fois de grandes unités de production et de nombreuses unités à proximité les unes des autres au sein d'un espace de concentration où existent des économies d'agglomération, c'est-à-dire l'addition des économies de localisation et des économies d'urbanisation.

En conséquence la décision de localisation d'un producteur dépend, non seulement de la distribution spatiale des consommateurs, mais aussi de la décision de localisation des autres producteurs, si bien que c'est la théorie des jeux qui est aujourd'hui la plus souvent mobilisée pour expliquer les interactions stratégiques et leurs conséquences en terme de force d'agglomération et de dispersion.

I -4 Application des modèles gravitaires

Les travaux sur les modèles gravitaires constituent la base même des théories de l'analyse spatiale.

I -4 -1 Les lois de RAVENSTEIN ¹(fin XIXe)

- Le nombre de migrants diminue quand la distance augmente; la plupart ne vont pas très loin, ceux qui se déplacent sur de grandes distances se dirigent de préférence vers les grands centres commerciaux et industriels.
- Le processus se fait de la façon suivante : une ville à croissance rapide attire les gens des régions environnantes, les vides ainsi créés sont comblés par les migrants de districts plus éloignés ; la force d'attraction des grandes villes dynamiques se fait donc sentir de proche en proche en diminuant d'intensité.
 - Le nombre de migrants de la zone d'accueil est donc proportionnel à la population de la zone d'origine et inversement proportionnel à la distance qui les sépare.
 - Chaque courant principal de migration suscite un contre-courant compensatoire.
 - Les citadins ont une mobilité plus faible que les ruraux.
 - L'intensité des migrations augmente avec le développement du commerce, de l'industrie et des transports.
- Les facteurs déterminant la migration sont nombreux, mais le plus important est le facteur économique.

¹ (Ravenstein E. géographe. 1885 & 1889, "The Laws of Migration", *Journal of Royal Statist. Society*, London)

I -4 -2 L'analogie gravitaire (PARETO, STEWART début XXe siècle)

Sous des formes voisines, plusieurs auteurs proposent de transposer aux migrations humaines la loi de la gravitation universelle de Newton :

Tableau 1.1 Modèle de Newton vs modèle de Stewart

<i>Modèle de Newton</i>	<i>Modèle de Stewart</i>
$A_{ij} = g \cdot M_i \cdot M_j / (D_{ij})^2$	$F_{ij} = k \cdot P_i \cdot P_j / (D_{ij})^2$
A_{ij} : force d'attraction entre i et j	F_{ij} : nombre de migrants de i vers j
g : constante gravitaire universelle	k : niveau global de mobilité
M_i, M_j : masse des planètes	$P_i \cdot P_j$: population des lieux i et j
D_{ij} : distance entre les planètes	D_{ij} : distance entre les lieux i et j

Les auteurs suivants introduisent des modifications et proposent plusieurs variantes :

$$F_{ij} = k \cdot (P_i)^{\beta_1} \cdot (P_j)^{\beta_2} \cdot (D_{ij})^{-\alpha} \quad : \text{modèle gravitaire généralisé (Pareto)}$$

$$F_{ij} = k \cdot (P_i)^{b_1} \cdot (P_j)^{b_2} \cdot \exp^{-\alpha D_{ij}} \quad : \text{modèle gravitaire généralisé (exponentiel)}$$

En clair, l'interaction entre les deux unités géographiques est proportionnelle au produit de leurs masses respectives et inversement proportionnelle à la distance les séparant. Si les masses et les distances ne sont pas les éléments créateurs de l'interaction, ils sont empiriquement un excellent descripteur des montants des interactions. En introduisant la notion de complémentarité, on peut qualifier les masses P_i et P_j et donc privilégier la qualité des unités aux dépens de la quantification de leurs masses. La notion d'occasions interposées améliore les qualités descriptives et explicatives du modèle gravitaire et définit un nouveau type de modèle dit de substitution ou relatif.

Première hypothèse : la mobilité (professionnelle) n'est pas proportionnelle à la distance mais est directement proportionnelle au nombre d'occasions qui existent au point d'arrivée et inversement proportionnelle au nombre d'occasions interposées. Empiriquement, on a pu remplacé la distance par les occasions interposées.

Deuxième hypothèse : il faut faire entrer dans le modèle l'effet d'absorption et de concurrence des masses des autres unités géographiques, effet qui « absorbe » l'interaction. Donc postuler que l'intensité de l'interaction entre i et j est influencée par l'existence d'autres unités ayant une masse pouvant plus ou moins se substituer à la masse de l'unité j . La substitution dépend de la position géographique et/ou des aptitudes économiques des

unités entre i et j et le pouvoir d'attraction de j diminue par le renforcement du pouvoir d'absorption des autres unités géographiques du territoire.

I -4 -3 La remise en cause de l'analogie gravitaire (fin XXe siècle)

De nombreux auteurs n'acceptent pas la transposition au comportement humain des lois de la physique et tentent d'expliquer autrement la décroissance des interactions avec la distance :

I -4 -3 -1 L'explication économique (ZIPF)

L'effet de la distance doit s'interpréter comme l'ajustement de l'offre et de la demande de mobilité en fonction du coût de déplacement.

La distance n'agit que si elle est proportionnelle au coût de mise en relation. Les acteurs obéissent au principe du moindre effort (rationalité économique).

I -4 -3 -1 L'explication sociologique (S. STOUFFER)

Plus la distance augmente, plus le nombre d'occasions interposées (*intervening opportunities*) est élevé. Ce n'est donc pas la distance qui agit directement mais la population du fait que le migrant trouve d'autres occasions de migrer entre son lieu de départ et son lieu d'arrivée.

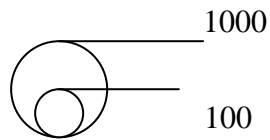
I -4 -3 -1 L'explication psychologique (T. HAGERSTRAND)

Toute migration est précédée d'un flux d'information (connaissance du lieu de destination). La décroissance des migrations avec la distance est liée à l'ignorance des migrants sur les lieux de destination éloignés [8].

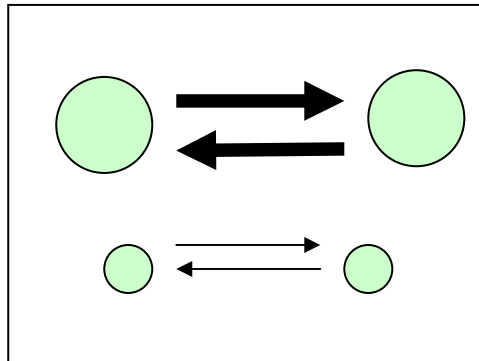
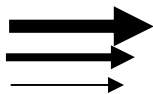
I-5 Les trois hypothèses fondamentales pour modéliser les flux

Hypothèse 1 : Les échanges entre deux lieux sont proportionnels à leurs capacités d'émission et de réception.

Figure 1.7 : Hypothèse 1



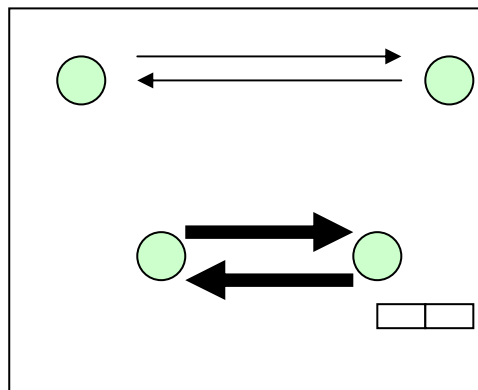
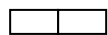
Capacité des lieux à générer ou à attirer des flux



Hypothèse 2 : l'importance des échanges entre deux lieux diminue lorsque la distance augmente.

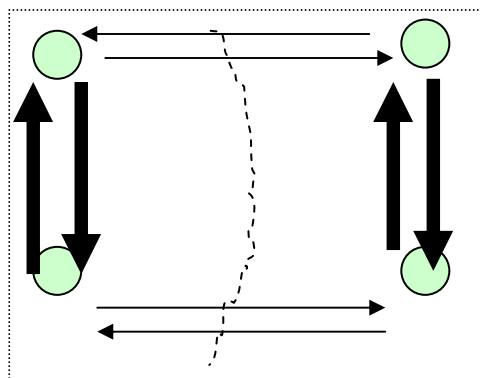
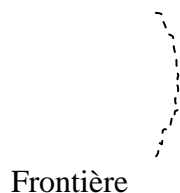
Figure 1.8 : Hypothèse 2

Distance



Hypothèse 3. Deux lieux appartenant à la même entité territoriale ont des flux plus importants que deux lieux séparés par une frontière.

Figure 1.9 : Hypothèse 3



Conclusion

Les premiers modèles d'analyse des flux migratoires se réfèrent à cette théorie et prennent en compte la distance entre les points de départ et le point d'arrivée, ainsi que le pouvoir de répulsion du point de départ et celui d'attraction de la zone d'arrivée.

Si ils étaient à l'origine assez sommaire (modèle de Stouffer et modèle Zipf), les nombreux modèles développés par la suite sont plus complexes et permettant de saisir avec plus de précision le phénomène de mobilité spatiale.

Cette marginalisation de l'espace n'est pas très claire, On peut penser que l'espace s'est éloigné de la scène quand les économistes ont commencé à vouloir appréhender de façon plus systématique et formelle l'ensemble des interdépendances économiques, en reculant les limites du non démontré, en recherchant à endogénéiser la plupart des phénomènes. En prenant comme exemple Adam Smith, dans ses *Recherches sur la nature et les causes de la richesse des nations* (1776), il s'est plus penché sur les aspects aspatiaux du problème que sur les facteurs proprement géographiques.

CHAPITRE II

Introduction aux modèles d'économétrie spatiale

Introduction

Il existe plusieurs manières d'intégrer l'espace dans l'analyse économique. La plus simple est celle proposée par le paradigme Arrow-Debreu : les choix des agents porteraient sur des paniers de biens où des biens physiques identiques seraient considérés comme deux biens différents s'ils étaient localisés en des lieux différents, moyennant quoi la plupart des résultats obtenus en microéconomie générale se généralise en microéconomie spatiale. Il semble toutefois que l'on puisse difficilement obtenir ainsi une véritable intégration de l'espace et de ses problèmes spécifiques. L'espace ne peut être traité par une simple analogie avec le temps. Mais la principale critique est que le stratagème se heurte à l'hypothèse de convexité de l'ensemble de consommation. Cela révèle que l'espace n'est pas seulement une dimension supplémentaire de l'analyse et que l'économie spatiale n'est pas une simple généralisation de résultats obtenus dans une représentation a-spatiale de l'économie. Il est reconnu que l'espace modifie, voire remet totalement en cause un certain nombre de propositions théoriques microéconomiques traditionnelles, ce qui cause à la fois de la difficulté et de l'intérêt.

Intégrer l'espace dans l'analyse, c'est d'abord prendre en compte une ou plusieurs composantes fondamentales de l'espace géographique (Beguin et Thisse (1979)). Il y a géographie quand on a défini un ensemble de lieux. Une représentation et une mesure de distance entre ces lieux. Éventuellement une mesure de surface pour chacun d'eux et quand on a caractérisé ces lieux par des attributs.

Intégrer l'espace dans l'analyse, c'est aussi retenir les cinq principes de Paelinck¹ (1982).

- 1) Les modèles spatiaux doivent être fondés sur une interdépendance spatiale.
- 2) Ces relations d'interdépendance spatiale sont vraisemblablement asymétriques
- 3) Selon le principe d'allotropie, les phénomènes localisés en un lieu sont souvent expliqués par des causes localisées en d'autres lieux."allotropie"² : ce principe concerne l'influence à distance de variables exogènes.
- 4) L'interaction ex post est différente de l'interaction ex ante.
- 5) L'espace contenant bidimensionnel doit apparaître explicitement.

L'interaction entre des agents différemment localisés est ainsi au centre des recherches les plus actuelles en économie spatiale.

L'économie spatiale est une autre façon de poser et penser les mêmes problèmes ; mais c'est aussi un autre ensemble de problèmes. On peut dire qu'une question tout à fait centrale est de comprendre pourquoi et comment les agents et les activités économiques se répartissent dans l'espace de façon non uniforme, créant des lieux de plus ou moins grande densité, engendrant des centres et des périphéries, des concentrations productives, des villes et des villages.

Déjà, Koopmans (1957) suggérait le lien étroit entre les indivisibilités et la question de la répartition spatiale des activités. Le théorème de Starett (1978), nous enseigne plus

¹ Emeritus professor, Erasmus University Rotterdam 2000

² mot d'origine grec, signifiant respectivement "autre" et "site"

précisément que sans ces indivisibilités et les rendements croissants qui en découlent, on n'a en général qu'une solution possible au problème de l'équilibre général spatial. Celle qui répartit les activités dans l'espace (on produit tous les biens dans tous les lieux), ce qui revient à nier l'espace. Les indivisibilités et les rendements croissants se voient ainsi reconnaître un rôle clé dans la microéconomie spatiale et dans l'émergence des économies d'agglomération qui mesurent les avantages obtenus par les agents du fait de leur proximité géographique.

Dans les travaux empiriques, l'économiste est confronté à l'utilisation de données localisées, c'est-à-dire au traitement d'observations d'une variable mesurée en des localisations différentes réparties dans l'espace. Il est souvent admis que ces données spatiales observées en coupe transversale sont indépendantes alors que cette hypothèse est rarement justifiée et devrait être systématiquement testée. L'introduction de l'espace dans les modèles économétriques n'est ni neutre, ni immédiate, et les techniques de l'économétrie spatiale visent à prendre en compte la présence de deux effets spatiaux importants : l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre observations géographiques, et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace.

II-1 L'état de l'art de l'économétrie spatiale

« Historiquement, c'est à Cliff et Ord (1973) que l'on doit, après une série d'articles à la fin des années 60 et au début des années 70, un ouvrage présentant de manière synthétique l'état des savoirs en statistique et en économétrie spatiales. La fin des années 70 et les années 80 sont marquées par le raffinement du cadre original d'analyse et plus particulièrement par le développement de la théorie de l'estimation et des tests (Ord, 1975 ; Paelinck et Klassen, 1979 ; Anselin, 1980). Un ensemble d'ouvrages rendent compte de ces développements, comme par exemple ceux de Cliff et Ord (1981), Anselin (1988a), Griffith (1988a), Cressie (1993), Jayet (1993) ou Anselin et Florax (1995a).

Jusqu'au début des années 90, les techniques permettant de spécifier, d'estimer et de tester la présence de l'autocorrélation spatiale dans les modèles économétriques étaient principalement publiées dans les revues spécialisées et appliquées à des problèmes d'économie régionale, spatiale ou urbaine. Ainsi, Can (1992), Can et Megboluge (1997), Pace et Gilley (1997) ont abordé le traitement de l'autocorrélation spatiale dans les modèles hédoniques³ de prix immobiliers, alors qu'Anselin *et al.* (1997) et Varga (1998, 2000) ont étudié les externalités spatiales d'information dues à la recherche universitaire et à la R&D.

Toutefois, les méthodes de l'économétrie spatiale sont maintenant appliquées à d'autres sujets tels que l'analyse de la demande (Case, 1991) ; l'économie internationale (Aten, 1997) ; l'économie publique (Case *et al.*, 1993 ; Brueckner, 1998), l'économie rurale (Benirschka et Binkley, 1996) ; les lois de Kaldor-Verdoorn (Fingleton, 2000, 2001, 2002) ou encore les

³ Théorie qui attribue rigoureusement un prix à des objets qui n'en ont pas.

phénomènes de croissance et de convergence (Moreno et Trehan, 1997 ; Fingleton, 1999 ; Rey et Montouri, 1999). Ces méthodes sont, en effet, potentiellement applicables à toutes les études empiriques nécessitant l'utilisation de données spatiales.

Selon Anselin et Bera (1998) et Anselin (2001b), deux raisons principales peuvent être attribuées au regain d'attention porté à la prise en compte des effets spatiaux dans les études empiriques. La première est le développement de nouveaux courants théoriques, tels que l'économie géographique et ses ramifications en économie industrielle, économie internationale, théories de la croissance ou économie du travail..., qui ont généralisé la prise en compte des interactions et des externalités spatiales dans l'analyse des décisions économiques des agents. La modélisation de l'autocorrélation spatiale permet en effet de capter l'existence, l'ampleur et l'influence des effets de débordement géographiques. La deuxième raison, d'ordre empirique, est la disponibilité croissante des données spatialisées, le fort développement actuel des logiciels de Systèmes d'Information Géographiques (SIG) et l'apparition de logiciels, tels que *SpaceStat* (Anselin, 1999), ou la librairie pour Matlab (LeSage, 1999a), proposant l'estimation des principaux modèles spatiaux.

II-2 Introduction aux modèles d'économétrie spatiale

L'objet de ce chapitre est de présenter les outils nécessaires à une démarche économétrique visant à prendre en compte les effets d'autocorrélation et d'hétérogénéité spatiale.

II-2-1 Les effets spatiaux : définition et détection

Les données spatiales sont des observations d'une variable mesurée en des localisations différentes réparties dans l'espace. Par conséquent, les données spatiales sont « spéciales » : elles combinent une information sur les attributs ainsi qu'une information sur la localisation de ces attributs, mesurée par exemple par une adresse postale ou des coordonnées géographiques (latitude et longitude). Les données spatiales possèdent souvent des propriétés particulières et doivent être analysées différemment des données a-spatiales. Le but de cette section est de préciser la nature particulière de ces données.

La multitude de données spatiales existante appelle des traitements statistiques et économétriques différents. En économie, on est généralement confronté à trois types d'observations. Il peut s'agir, tout d'abord, de points représentant, par exemple, des localisations d'unités de production ou de distribution... Ces points sont souvent mesurés par leur latitude et leur longitude. Ensuite, ces localisations peuvent être des lignes, connectées entre elles ou non, comme un réseau routier ou fluvial. Enfin, les données sont parfois fournies pour des aires géographiques comme des régions ou des pays, nous nous restreignons au 2^{ème} type d'observations.

Dans tous les cas, le nombre de ces points, de ces lignes ou de ces zones est *fini*. Cette caractéristique permet de distinguer les techniques de l'économétrie spatiale de celles de la géostatistique. L'économétrie spatiale est principalement utilisée lorsqu'on est en présence

d'un *ensemble fini* (régulier ou non) de points ou de zones reliés entre eux par des relations de voisinage. La géostatistique concerne essentiellement les *données issues d'un processus spatial sous-jacent continu* sur l'espace étudié. L'utilisation de l'une ou l'autre méthodologie dépend de la disponibilité des données et du problème étudié.

Les caractéristiques des données spatiales sont: l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre des observations géographiques, et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace. Des outils spécifiques sont nécessaires à la modélisation des interdépendances entre les observations : les matrices de poids et les variables spatiales décalées. Les méthodes d'analyse exploratoire des données spatiales ont pour but de caractériser les distributions ainsi que la présence et l'ampleur des effets spatiaux.

II-2 -2 Autocorrélation et hétérogénéité spatiales

Depuis Anselin (1988a), l'habitude a été prise de classer les effets spatiaux en deux grandes catégories : tout d'abord, la dépendance spatiale qui se réfère à l'absence d'indépendance entre observations géographiques, et ensuite l'hétérogénéité spatiale qui est liée à la différenciation spatiale des comportements.

II-2 -2 -1 L'autocorrélation spatiale

Dès 1914, Student suspecte la présence d'une relation entre différentes observations géographiques, une idée qui entraîne l'abandon de l'hypothèse statistique fondamentale d'observations indépendantes. Ce phénomène, appelé autocorrélation spatiale, est défini par Anselin et Bera (1998) de la façon suivante :

« *Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity.* » En d'autres termes, *l'autocorrélation spatiale positive* se traduit par une tendance à la concentration dans l'espace de valeurs faibles ou élevées d'une variable aléatoire. En revanche, *l'autocorrélation spatiale négative* signifie que chaque localisation tend à être entourée par des localisations voisines pour lesquelles la variable aléatoire prend des valeurs très différentes. Enfin, *l'absence d'autocorrélation spatiale* indique que la répartition spatiale des valeurs de la variable est aléatoire. Détecter de l'autocorrélation spatiale dans une série spatiale fournit alors une information supplémentaire par rapport aux statistiques traditionnelles (telles que la moyenne ou l'écart-type) sur la façon dont les différentes valeurs sont disposées géographiquement.

La présence d'autocorrélation spatiale pour une variable signifie qu'il y a une relation fonctionnelle entre ce qui se passe en un point de l'espace et ce qui se passe ailleurs. Tobler (1979) l'avait déjà souligné en suggérant la première loi de la géographie suivante :

« *Everything is related to everything else, but closer things more so.* » Les concepts de « proximité » et de « distance » sont pris en compte à travers l'utilisation de matrices de poids. L'autocorrélation spatiale est multidirectionnelle et « tout est relié à tout » ; elle diffère donc de l'autocorrélation temporelle qui est unidirectionnelle.

Cette interdépendance généralisée a pour conséquence de rendre plus complexes les méthodes de traitement de l'autocorrélation spatiale. Par exemple, certaines méthodes d'estimation valables pour les séries temporelles ne sont pas directement transposables au cas spatial.

L'autocorrélation spatiale a deux sources principales :

1/ Elle peut provenir du fait que les données sont affectées par des processus qui relient des lieux différents et qui sont à l'origine d'une organisation particulière des activités dans l'espace (Odland, 1988 ; Haining, 1990).

En effet, les processus d'interaction sont source d'autocorrélation spatiale lorsque les événements ou les circonstances en un lieu donné affectent les conditions en d'autres lieux si ces derniers interagissent, d'une manière ou d'une autre, par des mouvements de biens, de personnes, de capitaux, par des externalités spatiales ou par toutes les formes de comportements où un acteur économique réagit aux actions d'autres acteurs. Ainsi, la diffusion d'un phénomène (comme la diffusion technologique) à partir d'un ou de plusieurs lieux d'origine implique que l'intensité de la mesure de ce phénomène dépend de la distance à l'origine. Aux localisations proches les unes des autres, et donc à des distances comparables de l'origine, sont donc associées des intensités similaires pour le phénomène étudié.

2/ Elle peut également provenir d'une mauvaise spécification du modèle, comme des variables omises spatialement autocorrélées, d'une forme fonctionnelle incorrecte ou d'erreurs de mesure. C'est, en particulier, le cas lorsque l'étendue spatiale du phénomène étudié ne coïncide pas avec les unités spatiales d'observation. Elle est alors considérée comme un outil de diagnostic et de détection d'une mauvaise spécification du modèle.

II-2 -2 -1-1 Autocorrélation spatiale et modèles économétriques

Sur le plan économétrique, la non-vérification de l'hypothèse d'indépendance des observations conduit à remettre en cause l'inférence statistique basée sur l'estimation d'une régression par les MCO. Le test de l'autocorrélation spatiale s'avère alors être une précaution indispensable à toute étude empirique mobilisant des données géographiques. La démarche à adopter s'articule autour de deux opérations : détecter et modéliser l'autocorrélation spatiale.

Les tests de l'autocorrélation spatiale permettent à la fois de détecter la présence ainsi que la forme de l'autocorrélation dans les régressions, orientant ainsi l'économiste vers la spécification économétrique appropriée. Cependant, dans les études empiriques, on peut aussi disposer *a priori* d'informations théoriques conduisant à préférer une forme particulière de dépendance spatiale entre les observations pour capter différents types d'effets de débordement géographiques. Ainsi, les tests et la modélisation de l'autocorrélation spatiale apparaissent être des opérations interdépendantes.

II-2-2-1-2 Définition de l'autocorrélation spatiale et modèle économétriques

Comme point de départ, considérons le modèle de régression linéaire classique en coupe transversale :

$$y = X\beta + \varepsilon \quad (2-1)$$

Nous adopterons les conventions d'écriture suivantes :

N est le nombre total d'observations, ici des zones géographiques.

K est le nombre de coefficients inconnus à estimer.

y est le vecteur $(N,1)$ des observations de la variable dépendante.

X est la matrice (N,K) des observations des K variables explicatives

β est le vecteur $(K,1)$ des K coefficients inconnus à estimer

ε est le vecteur $(N,1)$ des termes d'erreurs.

Les hypothèses du modèle linéaire général sont les suivantes :

H_1 : Sur les variables explicatives.

X est une matrice non-stochastique de rang complet $N = K$ et lorsque la taille de l'échantillon devient infiniment grande, $\lim_{n \rightarrow \infty} (1/N)X'X = Q$ où Q est une matrice finie et non-singulière.

H_2 : Sur les erreurs.

Le vecteur d'erreur consiste en des erreurs non-observables qui satisfont les propriétés

$E(\varepsilon) = 0$ et $E(\varepsilon \varepsilon') = \sigma^2 I_N$ où $E(\cdot)$ désigne l'espérance mathématique et I_N la matrice identité d'ordre N .

Si ces différentes conditions sont vérifiées, alors l'estimateur des MCO défini par $\hat{\beta} = (X'X)^{-1} X'y$ est un estimateur BLUE⁴ (théorème de Gauss-Markov) : il s'agit de l'estimateur centré et à variance minimale dans la classe des estimateurs linéaires.

L'introduction de l'autocorrélation spatiale dans le modèle de régression linéaire conduit au relâchement de certaines de ces hypothèses. Cette introduction peut s'effectuer de plusieurs manières : par des *variables spatiales décalées*, endogènes, ou exogènes, ou par une *autocorrélation spatiale des erreurs*. Enfin, des processus plus complexes dans lesquels ces deux aspects sont combinés sont également envisageables.

⁴ *best linear unbiased estimator*

II-2-2 -2 L'hétérogénéité spatiale et ses deux facettes

Lorsque l'on étudie les espaces urbains, on fait souvent apparaître une segmentation des marchés immobiliers : les caractéristiques et les prix des logements diffèrent substantiellement selon leurs localisations. Cette segmentation, provenant entre autres de l'inélasticité de la demande des ménages pour certaines caractéristiques des logements ou encore de diverses barrières institutionnelles, conduit à des variations persistantes et significatives des caractéristiques des logements et de leurs prix dans les différents sous marchés.

Dans ces conditions, estimer une relation « globale » entre le prix du logement et ses caractéristiques - relation s'appliquant de la même façon sur toute l'aire urbaine étudiée - ne permet pas de capter les différences importantes de prix dans l'espace.

L'instabilité dans l'espace des relations économiques illustrée par cet exemple est appelée *hétérogénéité spatiale*. Ce phénomène se retrouve à plusieurs échelles : les comportements et les phénomènes économiques ne sont pas les mêmes dans le centre d'une ville et dans sa périphérie, dans une région urbaine et dans une région rurale, dans le « Nord » et dans le « Sud », etc. Plus précisément, deux aspects sont liés à cette notion : l'instabilité et l'hétéroscédasticité.

Le premier, *l'instabilité*, provient de l'absence de stabilité dans l'espace des comportements ou d'autres relations étudiées : les formes fonctionnelles et les paramètres varient selon leurs localisations et ne sont donc pas homogènes. Il est donc nécessaire de mobiliser des modélisations prenant en compte les caractéristiques particulières de chaque localisation de l'échantillon.

Le deuxième aspect de l'hétérogénéité spatiale est *l'hétéroscédasticité*. Dans les modèles économétriques, elle peut venir de variables manquantes ou de toute autre forme de mauvaise spécification. Par exemple, les unités spatiales elles-mêmes ne sont généralement ni de formes régulières, ni homogènes : des régions peuvent avoir des formes et des aires différentes, des niveaux de développement technologique variables, des populations plus ou moins importantes, etc.

En conclusion, pour décrire des phénomènes quantitatifs ou qualitatifs ayant une dimension spatiale, les interactions spatiales sont fréquentes. Ainsi, il est nécessaire de considérer non seulement les dimensions et les structures des observations mais aussi leurs positions relatives. Ceci s'effectue à l'aide des matrices de poids.

II-2 -3 Les matrices de poids et les variables spatiales décalées

Pour modéliser les interactions spatiales, il est nécessaire d'imposer une structure sur l'étendue de ces interactions. En effet, on ne peut pas estimer $N(N-2)/2$ termes d'interaction avec N observations. En économétrie spatiale, on définit alors un ensemble de voisins pour chaque région de l'échantillon, ce qui aboutit aux matrices de poids. D'une façon générale, les matrices de poids permettent de spécifier de manière exogène la topologie du système spatial.

Une matrice de poids W est une matrice carrée, ayant autant de lignes et de colonnes qu'il y a de zones géographiques (on note N le nombre de régions) et où chaque terme w_{ij} représente la façon dont la région i et la région j sont connectées spatialement. Elles ne contiennent donc pas d'éléments à estimer.

II-2 -3-1 Les matrices de contiguïté

Les matrices les plus utilisées sont les *matrices de contiguïté*. La contiguïté à l'ordre 1 entre deux régions se définit par le fait qu'elles ont une frontière commune. Chaque terme w_{ij} est alors défini comme suit :

$$w_{ij} = \begin{cases} 1 & \text{Si les régions } i, j \text{ sont contiguës d'ordre 1} \\ 0 & \text{Sinon} \end{cases}$$

Par convention, une région n'est pas contiguë avec elle-même : $w_{ii} = 0, \forall i$.

Si l'on désire connaître le nombre de régions contiguës à une région i , il suffit de calculer la somme des éléments de la ligne i de la matrice de contiguïté soit $L_i = \sum_j w_{ij}$. Le nombre total de liens existant dans le système régional est alors égal à $A = (1/2) \sum_i L_i$.

Dans le cas d'une disposition spatiale régulière, la définition de la contiguïté n'est pas unique. Ainsi, si l'on considère la figure 2.1, trois notions de contiguïté sont possibles. Ces dernières font référence au jeu d'échecs et au déplacement de la tour, du fou, de la dame. La tour se déplace horizontalement ou verticalement alors que le fou se déplace sur les diagonales. La dame peut se déplacer quant à elle dans toutes les directions. Les définitions de la contiguïté se déduisent de ces différents types de déplacement.

Tableau 2.1 : Les définitions de la contiguïté pour une configuration spatiale régulière

c	b	c
b	a	b
c	b	c

Dans le cas de la tour, la région a compte 4 voisines : les régions de type b qui partagent avec la région a des côtés communs. Dans le cas du fou, la région a compte 4 voisines : les régions de type c qui partagent avec la région a des coins communs. Enfin, dans le cas de la dame, la région a compte 8 voisines : les régions de type b et de type c qui partagent un côté ou un coin commun. Ces diverses définitions sont souvent utilisées dans des exercices de simulation visant à établir les propriétés des différents tests.

II-2 -4 Le modèle autorégressif spatial : variable endogène décalée

« La première façon de prendre en compte l'autocorrélation spatiale peut se faire grâce au modèle autorégressif spatial : une « variable endogène décalée » est incluse dans le modèle (2.1). Le modèle structurel s'écrit dans ce cas de la façon suivante :

$$y = \rho Wy + X\beta + \varepsilon \quad (2.2)$$

Wy est la variable endogène décalée pour la matrice de poids W , ρ est le paramètre spatial autorégressif indiquant l'intensité de l'interaction existant entre les observations de y .

Dans ce modèle, l'observation y_i est, en partie, expliquée par les valeurs prises par y dans les régions voisines : $(Wy)_i = \sum_{j \neq i} w_{ij} y_j$. En effet, $(Wy)_i$ (s'interprète comme la moyenne des valeurs de y sur les observations voisines à i lorsque W est standardisée. Cette standardisation facilite aussi la comparaison de l'ampleur de ρ lorsque (2.2) est estimé pour plusieurs matrices de poids.

L'introduction de Wy dans le modèle (2.2) est un moyen d'apprécier le degré de dépendance spatiale alors que les autres variables sont contrôlées. Symétriquement, il permet de contrôler la dépendance spatiale pour évaluer l'impact des autres variables explicatives.

Lorsqu'une variable endogène décalée est ignorée dans la spécification du modèle, mais présente dans le processus générateur des données, les estimateurs des MCO dans le modèle aspatial (2.2) sont biaisés et non convergents.

Cette spécification possède plusieurs propriétés.

1/ Filtrage spatial de la variable dépendante

Le modèle (2.2) se réécrit de la façon suivante :

$$(I - \rho W)y = X\beta + \varepsilon \quad (2.3)$$

Dans ce modèle : $y^* = (I - \rho W)y$ s'interprète comme la variable dépendante « filtrée », dans laquelle les effets de l'autocorrélation spatiale ont été éliminés.

2/ Effets de multiplicateur et de diffusion

Supposons la matrice $(I - \rho W)$ non-singulière. Ceci est vrai lorsque ρ est différent de 0 et lorsque l'inverse de ρ n'est pas une valeur propre de W . Dans ce cas, (2.3) se réécrit sous la forme réduite suivante :

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \quad (2.4)$$

L'espérance mathématique de y s'écrit alors :

$$E(y) = (I - \rho W)^{-1} X\beta$$

La matrice inverse $(I - \rho W)^{-1}$ est une matrice pleine qui, lorsque $w_{ij} < 1$ et $|\rho| < 1$.

Implique une série infinie pour les variables explicatives et pour le terme d'erreur associé à toutes les localisations :

$$y = (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \rho^4 W^4 + \dots) X \beta + (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \varepsilon$$

A partir de cette expression, on peut définir deux types d'effets : *un effet de multiplicateur spatial* affectant les variables explicatives et *un effet de diffusion spatiale* affectant les erreurs.

Concernant les variables explicatives, cette expression signifie qu'en moyenne la valeur de y dans une région i n'est pas seulement expliquée par les valeurs des variables explicatives associées à cette région, mais aussi par celles associées à toutes les régions (voisines de i ou non) à travers la transformation spatiale inverse $(I - \rho W)^{-1}$. Cet effet de *multiplicateur spatial* décline avec l'éloignement.

Concernant le processus des erreurs, cette expression signifie qu'un choc aléatoire dans une région i affecte non seulement la valeur de y de cette région, mais a également un impact sur les valeurs de y dans les autres régions à travers la même transformation spatiale inverse.

C'est *l'effet de diffusion*, effet qui décline aussi avec l'éloignement.

3/ Matrice des variances-covariances

De (2.4), on déduit la matrice des variances-covariances de y :

$$V(y) = (I - \rho W)^{-1} E(\varepsilon \varepsilon') (I - \rho W')^{-1}$$

$$V(y) = \sigma^2 [(I - \rho W)' (I - \rho W)]^{-1}$$

Cette matrice des variances-covariances est une matrice pleine, ce qui implique que chaque localisation est corrélée avec chaque autre localisation du système mais que cette corrélation diminue avec l'ordre de proximité.

4/ Variable endogène décalée et hétéroscédasticité

Posons $u = (I - \rho W)^{-1} \varepsilon$. La matrice de variances-covariances de u s'écrit alors :

$$V(u) = \sigma^2 [(I - \rho W)' (I - \rho W)]^{-1} \quad (2.5)$$

De (2.5), il s'ensuit que la covariance entre chaque paire d'erreurs est non nulle et décroissante, mais décroît avec l'ordre de proximité. Par ailleurs, les éléments de la diagonale de $V(u)$ ne sont pas constants, ce qui implique l'hétéroscédasticité des erreurs u , que ε soit hétéroscédastique ou non (McMillen, 1992).

Ce modèle autorégressif est utilisé, par exemple, pour modéliser les interactions stratégiques et la concurrence fiscale entre communes (Case *et al.*, 1993 ; Brueckner, 1998), les externalités de voisinage dans les modèles hédoniques de prix immobiliers (Can, 1990, 1992 ; Macedo, 1998), les externalités spatiales influençant la distribution de la

population dans les espaces urbains (Griffith, 1981 ; Griffith et Can, 1996) ou influençant la production de firmes concurrentes (Griffith, 1999).

II-2 -4-1 Modèle avec autocorrélation spatiale des erreurs

Une dernière façon d'incorporer l'autocorrélation spatiale dans un modèle de régression est de spécifier un processus spatial pour les erreurs. Différentes possibilités existent qui mènent à différentes covariances des erreurs, avec des implications variables pour la nature et l'ampleur de l'autocorrélation spatiale dans le modèle (Anselin, 2001).

II-2-4-1-1 Les erreurs suivent un processus autorégressif

« La spécification la plus utilisée est un processus spatial autorégressif dans les erreurs.

Le modèle structurel s'écrit alors :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \end{aligned} \tag{2.6}$$

Le paramètre λ reflète l'intensité de l'interdépendance entre les résidus de la régression, u est le terme d'erreur tel que : $u \rightarrow iid(0, \sigma^2 I)$. Omettre à tort une autocorrélation spatiale des erreurs produit des estimateurs non biaisés mais inefficients et l'inférence statistique basée sur les MCO est biaisée.

Cette spécification possède plusieurs propriétés.

1/ Diffusion spatiale

Tout d'abord, si la matrice $(I - \lambda W)$

est non-singulière, le modèle (2.6) se réécrit sous la forme réduite suivante :

$$y = X\beta + (I - \lambda W)^{-1} u$$

Cette expression fait apparaître un effet de diffusion spatiale comme pour le modèle (2.3) mais comme $E(y) = X\beta$, il n'y a pas d'effet de multiplicateur spatial.

2/ Filtrage spatial de la variable dépendante et des variables explicatives

Le modèle (2.6) peut aussi être exprimé en termes de variables spatialement filtrées.

Après avoir pré-multiplié les deux termes de (2.6) par $(I - \lambda W)$, on obtient l'expression suivante :

$$(I - \lambda W) y = (I - \lambda W) X\beta + u$$

Ce modèle de régression est un modèle où la variable expliquée ainsi que les variables explicatives sont filtrées spatialement, le terme d'erreur u étant homoscédastique.

3/ Matrice des variances-covariances

De la formulation [2.6], il s'ensuit que :

$$v(y) = v(\varepsilon) = E[(I - \lambda W)^{-1} u u' (I - \lambda W')^{-1}]$$

$$v(y) = v(\varepsilon) = \sigma^2 [(I - \lambda W)']^{-1} (I - \lambda W)^{-1}$$

On trouve, pour ε et pour y , une structure identique à celle du modèle spatial autorégressif. Par conséquent, la covariance entre chaque paire d'erreurs et d'observations est non nulle et décroissante : elle décroît avec l'ordre de proximité. La structure spatiale de covariance induite par le modèle avec autocorrélation des erreurs est donc *globale*, puisqu'elle relie toutes les localisations du système à toutes les autres (Mur, 1999 ; Anselin, 2001c).

Par ailleurs, la structure d'erreurs induit des éléments de la diagonale de $v(y)$ non constants, ce qui implique l'hétéroscédasticité des erreurs ε , que u soit hétéroscédastique ou non (McMillen, 1992).

II-2 -5 Estimation des modèles spatiaux

En présence d'autocorrélation spatiale, les estimateurs obtenus par les MCO ne sont pas convergents (cas d'une variable endogène décalée) et sont inefficients (cas d'une autocorrélation spatiale des erreurs). Pour trouver des estimateurs convergents et asymptotiquement normaux et efficaces, la méthode la plus couramment utilisée est celle du maximum de vraisemblance. Il est également possible d'avoir recours à la méthode des variables instrumentales ou à celle des moments généralisés.

II-2 -5-1 La non-convergence des MCO et ses conséquences

Considérons le modèle spatial le plus général, incluant à la fois une variable endogène décalée et une autocorrélation spatiale des erreurs (modèle 2.6). Ce modèle se réécrit de la façon suivante :

$$y = (I - \rho W_1)^{-1} X \beta + (I - \rho W_1)^{-1} \varepsilon$$

$$\varepsilon = (I - \lambda W_2)^{-1} u$$

Par conséquent, la matrice des variances-covariances de l'erreur Σ est :

$$V(\varepsilon) = E(\varepsilon \varepsilon') = \sigma^2 [(I - \lambda W_2)']^{-1} (I - \lambda W_2)^{-1}$$

La corrélation entre la variable explicative $W_1 y$ et l'erreur ε s'écrit :

$$E(W_1 y \varepsilon') = E[W_1 (I - \rho W_1)^{-1} X \beta \varepsilon' + W_1 (I - \rho W_1)^{-1} \varepsilon \varepsilon']$$

$$E(W_1 y \varepsilon') = W_1 (I - \rho W_1)^{-1} E(\varepsilon \varepsilon')$$

$$E(W_1 y \varepsilon') = \sigma^2 W_1 (I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} (I - \lambda W_2')^{-1} \quad (2.7)$$

L'expression (2.7) est non nulle en général, les éléments de la variable endogène décalée sont corrélés avec ceux des erreurs. Les paramètres du modèle (2.6) ne peuvent donc pas être estimés d'une façon convergente par les MCO.

Pour le modèle spatial autorégressif (2.6) les estimateurs des MCO ne sont pas convergents car la variable endogène décalée Wy est corrélée avec l'erreur ε , quelle que soit la distribution de cette erreur : si $\lambda = 0$ et $W_1 = W_2$ dans (2.7), alors $E(W_1 y \varepsilon') = \sigma^2 W_1 (I - \rho W_1)^{-1}$, expression non nulle en général. La matrice $(I - \rho W_1)^{-1}$ est une matrice pleine et non triangulaire, le décalage spatial pour une observation donnée i , $(Wy)_i$ n'est pas seulement corrélé avec le terme d'erreur en i , mais aussi avec tous les termes d'erreurs de toutes les localisations. Ce résultat contraste avec une propriété de séries temporelles. Dans ce cas, en effet, les estimateurs des MCO restent convergents en présence d'une ou de plusieurs variables retardées tant que les erreurs ne sont pas corrélées.

Pour le modèle avec autocorrélation spatiale des erreurs, l'estimateur de β par les MCO est sans biais mais inefficace puisque les erreurs ne sont pas sphériques. La solution théorique dans ce cas consiste dans l'application des Moindres Carrés Quasi-Généralisés (MCQG). Cependant, cette procédure n'est pas applicable au cas spatial puisque l'estimateur de λ par les MCO n'est pas convergent (Anselin, 1988a, chap. 6) » [13].

Finalement, pour aboutir à des estimateurs des paramètres dans les modèles spatiaux qui sont convergents, asymptotiquement normaux et efficaces, il faut utiliser d'autres méthodes d'estimation. La plus utilisée aujourd'hui est la méthode par le maximum de vraisemblance et la méthode d'estimation Bayésienne.

II-3 Méthodes d'estimation Bayésienne

« L'utilisation des estimateurs du Maximum de Vraisemblance a été dominante en économétrie spatiale jusqu'à la fin des années 90 (Anselin et Bera, 1998). Mais cette méthode nécessite le calcul du Hessien⁵, qui peut se révéler coûteux en termes de temps de calcul lorsque l'échantillon est de grande taille. De plus, la normalité des erreurs pour les problèmes spatiaux est sujette à caution (LeSage, 1997). Les modèles théoriques en économie spatiale (et notamment en économie géographique) exigent également l'introduction de nouvelles méthodes d'estimation capables de tenir compte de phénomènes d'interdépendance beaucoup plus complexes : effets asymétriques, variation des effets spatiaux entre régions, incertitude concernant les mesures de distance ou de proximité, etc. Nous avons vu dans le chapitre précédent que des estimateurs alternatifs ont été proposés. Ils sont basés sur les variables instrumentales ou encore les méthodes des moments généralisés qui ne nécessitent pas la normalité des erreurs. S'interrogeant sur les hypothèses de stationnarité et d'isotropie requises par les approches des moments généralisés, Conley (1996) suggère une approche non paramétrique pour l'estimation de la matrice de variance intégrant les interdépendances spatiales. Cependant, Anselin (2001) souligne que les propriétés de stationnarité ne sont pas appropriées pour l'estimation du modèle SAR⁶ du fait de l'hétéroscédasticité. Il est aussi de plus en plus courant de modéliser les interactions spatiales à partir de distances économiques (Pinske et Slade, 1998). L'approche que nous allons présenter permet d'estimer les modèles spatiaux en tenant compte à la fois des problèmes d'hétérogénéité, des observations extrêmes, et en modélisant les interactions spatiales à partir de fondements économiques.

On expose les différentes méthodes de simulation appelées MCMC (Monte Carlo Markov Chain) qui permettent l'estimation des modèles Bayésiens.

L'analyse statistique Bayésienne nécessite le calcul d'estimateurs issus de distributions de probabilité *a posteriori* généralement très complexes. Des techniques d'estimation Monte Carlo permettent alors de mettre en œuvre le calcul pratique de ces estimateurs. Le principe de base de ces méthodes est d'utiliser des séquences d'échantillons distribués suivant les lois *a posteriori* d'intérêt. Ces distributions sont souvent complexes et ne peuvent donc être simulées par des techniques classiques.

Les algorithmes de simulation par chaînes de Markov sont alors introduits. Ils permettent de simuler des distributions très générales et rendent possible le calcul des estimateurs issus des procédures d'estimations Bayésiennes. Les méthodes de simulation MCMC deviennent alors efficaces pour résoudre de nombreux problèmes d'estimation liés à la dépendance temporelle et spatiale et à la présence de distributions multimodales.

⁵ Matrice des dérivées partielles secondes.

⁶ Spatial Auto Regressif.

II-3-1 Introduction à l'approche Bayésienne

Depuis l'ouvrage fondateur de Zellner (1971), l'économétrie Bayésienne fait partie des domaines de recherche les plus attractifs en économétrie. De nombreux ouvrages spécialisés ont été publiés ces dernières années: Bauwens et al. (1999) se focalisent sur les séries temporelles, Poirier (1995) développe les aspects théoriques de l'économétrie Bayésienne par une comparaison avec l'approche fréquentiste, Koop (2003) présente un cadre généralisé de l'économétrie Bayésienne et LeSage (1997) introduit ces méthodes en économétrie.

II-3-1-1 Principes de l'économétrie Bayésienne

Différents types de modèles peuvent être estimés selon une approche fréquentiste ou probabiliste, Bayésienne. La distinction réside en fait dans la démarche même de ces deux approches. Dans l'approche fréquentiste les données observées sont considérées comme des observations de variables aléatoires.

À partir de ces observations $y \in Y \subset \mathbb{R}^n$ nous portons des inférences sur les paramètres $\theta \in \Theta \subset \mathbb{R}^p$ ayant dirigé leur mécanisme de génération. L'information provenant des données observées y est ainsi l'unique source d'information. Par ailleurs, elle repose sur des théories asymptotiques sous l'hypothèse de la répétition de circonstances similaires. En revanche l'approche Bayésienne considère les paramètres θ comme étant des variables aléatoires pour lesquelles nous spécifions une distribution a priori, noté $\pi(\theta)$. Une fois les données observées, l'inférence est alors faite à partir de leur distribution a posteriori, notée $\pi(y|\theta)$, qui est proportionnelle au produit de l'a priori et de la fonction de vraisemblance, notée $\pi(y|\theta)$. En utilisant le théorème de Bayes, nous obtenons:

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)} \quad (2.8)$$

Où le dénominateur est une constante d'intégration représentant la densité prédictive de y et ne dépendant pas de θ .

Nous pouvons donc écrire que :

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta) \quad (2.9)$$

Autrement dit, la loi a posteriori représente une actualisation de l'information a priori au vu de l'information apportée par les observations. Les méthodes Bayésiennes se sont considérablement développées ces dernières années (Koop, 2003), notamment en économétrie spatiale (Lesage, 1997). Ceci vient principalement de l'essor des méthodes de simulation permettant de générer la distribution a posteriori lorsque celle-ci est d'écriture complexe, et donc difficilement calculable analytiquement. En effet, il n'est pas toujours possible d'obtenir une écriture analytique simple de la distribution a posteriori. Cependant, lorsque les lois

utilisées appartiennent à la famille exponentielle, nous pouvons avoir recours aux lois dites conjuguées :

Si $\pi(y/\theta)$ appartient à la famille exponentielle, alors il existe une loi conjuguée naturelle pour $\pi(\theta)$ et $\pi(\theta/y)$. Quelques exemples de lois conjuguées naturelles sont donnés dans le tableau 2-1. Les démonstrations et les propriétés de ces lois sont exposées dans des ouvrages de référence tels que Berger (1985), Robert (1992) ou encore Bernardo et Smith (1994).

Tableau .2.2. Exemples de lois conjuguées naturelles pour des lois appartenant à la famille exponentielle

$\pi(y/\theta)$	$\pi(\theta)$	$\pi(\theta/y)$
Normale (θ, σ^2)	Normale (μ, τ^2)	Normale ($\rho(\sigma^2\mu + \tau^2y), \rho\sigma^2\tau^2$) avec $\rho^{-1} = \sigma^2 + \tau^2$
Poisson (θ)	G(α, β)	G($\alpha + y, \beta + 1$)
G(ν, θ)	G(α, β)	G($\alpha + \nu, \beta + y$)
Normale ($\mu, 1/\theta$)	G(α, β)	G($\alpha + 0.5, (\mu - y)^2 / 2$)

Lorsqu'il n'y a pas d'écriture analytique simple de $\pi(\theta/y)$ ou que la distribution *a posteriori* est multidimensionnelle, nous pourrions avoir recours à l'utilisation de méthodes de simulation pour générer des échantillons de variables aléatoires suivant la distribution $\pi(\theta/y)$. Ces méthodes de simulation reposent sur des algorithmes stochastiques faisant intervenir des procédures itératives d'échantillonnage. Nous allons ainsi échantillonner successivement les composants du vecteur de paramètres et créer une suite dépendante de valeurs de ces composants, mais possédant des propriétés markoviennes.

Les algorithmes les plus couramment utilisés appartiennent à la famille des algorithmes de Monte Carlo par Chaînes de Markov (MCMC). Il s'agit de l'échantillonneur de Gibbs (Tanner et Wong, 1987), ainsi que de l'algorithme de Metropolis (Metropolis et al, 1953), ou de son extension par Hastings (Hastings, 1970). Les principes de l'algorithme de Metropolis et de l'échantillonneur de Gibbs sont détaillés dans le paragraphe II-3-2.

L'essor de l'utilisation des méthodes Bayésiennes en économétrie est principalement dû : aux résultats de l'analyse Bayésienne hiérarchique, qui décompose la loi a priori en distributions conditionnelles indépendantes, ce qui revient à modéliser l'incertitude portée sur la loi a priori par une loi sur les paramètres de cette loi; grâce aux développements des techniques de simulation utilisant les algorithmes MCMC. Une des principales critiques à l'encontre des méthodes Bayésiennes tient au choix qui doit être fait de l'a priori puisque celui-ci sera déterminant dans l'inférence qui s'en suivra. Dans certains cas, seule l'expérience antérieure dans le domaine étudié peut apporter une information suffisamment précise pour permettre de déterminer la loi a priori. Dans les autres cas, nous ferons appel à des lois dites

conjuguées (tableau 2.2) ou à des lois non informatives qui, souvent au détriment de l'information a priori, pourront rendre le calcul mathématique plus simple. Cependant, plus la loi a priori est non informative, plus la loi a posteriori se rapproche de la vraisemblance et donc de l'information apportée par les données. Il peut donc y avoir une certaine subjectivité dans la détermination de la loi a priori et nous voyons bien que l'utilisation de lois non informatives rapproche l'inférence Bayésienne de l'inférence fréquentiste puisque nous ne considérons plus que l'information apportée par les observations.

De même, lorsque les échantillons deviennent d'une taille suffisamment grande, l'information a posteriori rejoint l'inférence classique c'est à dire que l'information a priori a moins d'importance. En revanche, pour des échantillons de petite taille l'approche Bayésienne peut fournir des estimations plus précises puisque nous tiendrons compte d'une information a priori.

Il existe cependant des méthodes permettant d'étudier la robustesse par rapport à la loi a priori, c'est-à-dire de mesurer les conséquences d'une imprécision de la loi *a priori* sur l'inférence *a posteriori* (Berger, 1985).

Du point de vue du calcul, les deux approches sont confrontées à des difficultés différentes : la méthode du maximum de vraisemblance se heurte à des problèmes d'optimisation (modes multiples, résolution des équations de vraisemblance ...) alors que ce sont des problèmes d'intégration qui sont rencontrés dans l'approche Bayésienne. Comme il en a déjà été fait mention, ces derniers sont contournés grâce aux méthodes d'échantillonnage Bayésien.

Enfin, la théorie de la décision fait intervenir dans le choix du décideur le degré de croyance qu'il attache à la réalisation de l'événement étudié. Il convient alors de pouvoir comparer les procédures décisionnelles au moyen d'un critère d'évaluation qui décrit les conséquences de chaque décision en fonction des paramètres du modèle. Nous verrons par la suite que la meilleure estimation des paramètres en question sera alors celle qui maximisera la fonction d'utilité et qui décrira les préférences du décideur. La nécessité d'un tel critère d'évaluation est là encore sujette à de nombreuses controverses (Robert, 1992).

Il faut également noter que l'utilisation des méthodes Bayésiennes demande certaines connaissances en statistiques (par exemple, propriétés de base sur les lois de probabilité, principes des méthodes MCMC) et en informatique, dans la mesure où leur mise en oeuvre peut faire appel à de la programmation.

II-3-1-2 Choix de distributions a priori

L'approche Bayésienne mène souvent, comme nous venons de le voir, à la considération de la distribution a posteriori dont la densité est proportionnelle au produit de la vraisemblance et de la densité a priori de la distribution d'intérêt. La vraisemblance, dépendant explicitement de la structure du modèle, sera toujours considérée comme étant définie. En revanche, la distribution a priori des paramètres n'est pas toujours fixée explicitement dans la modélisation, le choix de cette distribution n'est pas forcément unique car il est laissé à l'utilisateur.

Lorsque l'information a priori est très vague, il est nécessaire de considérer des distributions pouvant refléter le mieux ce manque d'information. Une solution consiste à

utiliser des distributions dites non-informatives. Nous utiliserons par la suite des distributions uniformes et conjuguées tout en sachant que celles de Jeffeys (construites à partir de la matrice d'information de Fisher) sont couramment employées. En effet, la solution consistant à introduire une loi uniforme permet d'affecter les mêmes probabilités à toutes les valeurs possibles de la variable. Cette distribution doit être utilisée avec précaution surtout dans le cadre de lois a priori impropres.

Robert (1996) met en garde contre l'utilisation des lois conjuguées (voir tableau 2.1). Nous nous attacherons par la suite à justifier le choix des lois a priori à partir de la théorie économique.

Le choix des distributions conjuguées a pour but d'exploiter de manière efficace la distribution de vraisemblance de façon à donner une forme explicite de la distribution a posteriori. Lorsque la vraisemblance est une distribution usuelle, une distribution a priori conjuguée (Robert, 1992) permet d'explicitier la distribution a posteriori. Dans les différents modèles que nous allons analyser, la vraisemblance $p(y|\theta)$ sera de type exponentiel :

$$p(y|\theta) = g(y) \exp(\theta y - \phi(\theta)) \quad (2.10)$$

Une distribution a priori de type

$$p(\theta | y, a) \propto g(y) F(\lambda, \mu) \exp(\theta(y + \mu) - (\lambda + 1)\phi(\theta)) \quad (2.11)$$

Avec $a = (\lambda, \mu)$. La distribution a priori (2.11) est alors dite conjuguée pour la vraisemblance (2.10).

Cette approche est très utilisée en pratique car de nombreuses distributions usuelles sont de type exponentiel (loi de Poisson, loi exponentielle, loi normale, etc.). Il est naturel de remarquer le caractère artificiel du choix de ce type de distribution pour prendre en compte l'information *a priori* sur le modèle. La flexibilité de cette approche permet cependant d'obtenir une variance importante pour des densités a priori.

II-3-2 Définition des méthodes de Monte Carlo par Chaînes de Markov

L'approche Bayésienne nécessite le calcul d'estimateurs à partir d'une loi a posteriori d'intérêt. Or il est souvent impossible d'établir une expression analytique de ces lois a posteriori. Nous allons introduire des méthodes de simulation de la loi d'intérêt visant à pallier aux problèmes d'intégration et d'optimisation rencontrés lors de la construction de ces estimateurs.

En effet, dans le cas de l'intégration numérique, il s'agit d'approcher :

$$I = \int_0^1 g(x) dx.$$

La méthode de Monte Carlo consiste à écrire cette intégrale sous la forme

$I = E[g(u)]$ avec U , une variable aléatoire suivant une loi uniforme sur $[0,1]$. Cette méthode est basée sur la loi des grands nombres. En d'autres termes, si u_1, \dots, u_n sont des nombres tirés au hasard dans l'intervalle $[0,1]$, $\frac{1}{n}(g(u_1) + \dots + g(u_n))$ est une approximation de $\int_0^1 g(x) dx$.

Dans un cadre plus général, nous chercherons par la suite à approcher l'intégrale :

$$I = \int_{R^n} g(x) f(x) dx .$$

Où $f(x) \geq 0$ et $\int_{R^n} f(x) dx = 1$ alors $I = E[g(x)]$ où X est une variable aléatoire de densité f par rapport à l'intégrale de Lebesgue⁷ sur R^n . Toujours par la loi des grands nombres, si $(X_i)_{i \in N}$ est une suite de variables indépendantes sur R^n de loi de densité f ,

$$\frac{1}{n} \sum_{i=1}^n g(X_i) \xrightarrow{p.s.} E[g(x)]$$

Et donc si (x_1, x_2, \dots, x_n) est une réalisation de (X_1, X_2, \dots, X_n) , $\frac{1}{n}((g(x_1) + g(x_2) + \dots + g(x_n)))$ sera une approximation de I .

Le principe de base des méthodes décrites dans cette section est l'utilisation d'une Chaîne de Markov ergodique de loi stationnaire f : partant d'une valeur arbitraire x_0 , nous générons une chaîne (X_t) à partir d'un noyau de transition de loi stationnaire f ; qui garantit la convergence en loi de la chaîne vers f .

Par comparaison avec les techniques classiques d'optimisation, nous verrons comment la recherche de modes peut être améliorée » [26].

II-3-2-1 Méthodes de simulation classiques

Le principe de simulation des événements discrets est celui d'un tirage aléatoire à l'aide d'un mécanisme analogue à celui d'une roulette. Pour des raisons évidentes, ce type de simulation est généralement connu sous le nom de simulation de Monte Carlo. Le premier point essentiel est celui de la génération de nombres dits aléatoires, c'est-à-dire de nombres qui seraient le résultat d'une succession de tirages aléatoires indépendants suivant une même loi de distribution uniforme sur $[0, 1]$. Dès lors qu'un générateur de variables aléatoires, ou pseudo-aléatoires, uniformément distribuées est disponible, la plupart des lois usuelles

⁷ Lebesgue ne subdivise pas l'intervalle d'intégration J , mais son image $f(J)$. à la différence de celle de Riemann et de Cauchy.

peuvent être simulées à partir de transformations déterministes d'échantillons simulés uniformément.

II-3-2-1-1 Simulation des méthodes MCMC

Les méthodes MCMC permettent de simuler des VA en construisant une Chaîne de Markov. Des estimateurs permettent ensuite d'intégrer ou d'optimiser des fonctions.

II-3-2-1-1-a Définition d'une chaîne de Markov

Une Chaîne de Markov est définie par la propriété suivante :

$$p(x_i | x_0, \dots, x_{i-1}) = p(x_i | x_{i-1})$$

Une Chaîne de Markov est définie par deux composantes:

La distribution initiale $p(x_0)$

le noyau de transition $T(x, A) = p(x_{i+1} \in A | x_i = x)$

Propriétés importantes des chaînes de Markov

- Invariance/Stationnarité : si x_i distribué suivant π , alors x_{i-1} et les suivants sont distribués suivant π .
- Irréductibilité : tous les ensembles de probabilité non nulle peuvent être atteints à partir de tout point de départ.
- Récurrence : les trajectoires (X_i) passent une infinité de fois dans tout ensemble de probabilité non nulle.
- Apériodicité : aucun noyau n'induit un comportement périodique des trajectoires.

II-3-2-1-1-b Algorithme de Metropolis-Hastings

Le schéma de simulation est d'autant plus efficace que la distribution $q(x_i / x^*)$ est simulable rapidement.

1. initialiser $x_0 \sim \pi_0(x_0)$, $l = 0$

2. à l'itération l :

Proposer un candidat x^* pour x_{l+1} c'est-à-dire simuler :

(a) simuler $x^* \sim q(x^* / x_l)$

Accepter le candidat x^* avec la probabilité

$$p(x^*, x_l) = \min \{1, r(x^*, x_l)\}$$

Simuler $u \sim U_{(0,1)}$

Si $u \leq p(x^*, x_l)$, alors prendre $x_{l+1} = x^*$, sinon poser $x_{l+1} = x_l$

3. le taux d'acceptation $p(x^*, x_l)$, fait intervenir le rapport $r(x^*, x_l)$ ci-après :

$$r(x^*, x_l) = \frac{\pi(x^*) q(x_l / x^*)}{\pi(x_l) q(x^* / x_l)}$$

Lorsqu'il est possible d'échantillonner à partir des probabilités conditionnelles, l'échantillonneur de Gibbs est le meilleur.

II-3-2-1-1-c Échantillonneur de Gibbs

Cet algorithme a été proposé par Geman et Geman (1984) pour résoudre des problèmes de traitement d'images puis a été généralisé à des problèmes de simulation statistique par Gelfand et Smith (1990) et plus particulièrement pour le calcul d'estimateurs Bayésiens par Smith et Roberts (1993). Le schéma de simulation est donné ci-dessous :

Schéma de simulation

1. initialisation $x^{(0)} = (x_1^{(0)}, \dots, x_N^{(0)})$

2. iteration i : simuler

$$x_1^{(i+1)} \sim \pi_1(\cdot / x_2^{(i)}, \dots, x_N^{(i)})$$

$$x_2^{(i+1)} \sim \pi_2(\cdot / x_1^{(i+1)}, x_3^{(i)}, \dots, x_N^{(i)})$$

⋮

$$x_N^{(i+1)} \sim \pi_N(\cdot / x_2^{(i+1)}, \dots, x_{N-1}^{(i+1)})$$

3. $i \leftarrow i + 1$ et aller en 2

Conclusion

L'utilisation des estimateurs du Maximum de Vraisemblance a été dominante en économétrie spatiale jusqu'à la fin des années 90 (Anselin et Bera, 1998). Mais cette méthode nécessite le calcul du Hessien, qui peut se révéler coûteux en termes de temps de calcul lorsque l'échantillon est de grande taille. De plus, la normalité des erreurs pour les problèmes spatiaux est sujette à caution (LeSage, 1997). Les modèles théoriques en économie spatiale (et notamment en économie géographique) exigent également l'introduction de nouvelles méthodes d'estimation capables de tenir compte de phénomènes d'interdépendance beaucoup plus complexes : effets asymétriques, variation des effets spatiaux entre régions, incertitude concernant les mesures de distance ou de proximité.

L'approche que nous avons présentée a permis d'estimer les modèles spatiaux tout en tenant compte à la fois des problèmes d'hétérogénéité, des observations extrêmes, et en modélisant les interactions spatiales à partir de fondements économiques.

Nous avons donc exposé dans ce chapitre les différentes méthodes de simulation appelées MCMC (Monte Carlo Markov Chain) qui permettent l'estimation des modèles Bayésiens.

L'analyse statistique Bayésienne nécessite le calcul d'estimateurs issus de distributions de probabilité a posteriori généralement très complexes. Des techniques d'estimation Monte Carlo permettent alors de mettre en œuvre le calcul pratique de ces estimateurs. Le principe de base de ces méthodes est d'utiliser des séquences d'échantillons distribués suivant les lois a posteriori d'intérêt. Ces distributions sont souvent complexes et ne peuvent donc être simulées par des techniques classiques.

Les algorithmes de simulation par Chaînes de Markov sont alors introduits. Ils permettent de simuler des distributions très générales et rendent possible le calcul des estimateurs issus des procédures d'estimations Bayésiennes. Les méthodes de simulation MCMC deviennent alors efficaces pour résoudre de nombreux problèmes d'estimation liés à la dépendance temporelle et spatiale et à la présence de distributions multimodales.

CHAPITRE III

Structure migratoire et spatiale de l'Algérie

Introduction

Les migrations de population peuvent être dues à des logiques économiques, démographiques ou politiques. Le mouvement de la population est l'un des éléments qui est à l'origine de la transformation de l'espace. Seul le changement de résidence entraînant le passage d'une zone à une autre est enregistré, le volume des migrants varie, en conséquence selon le découpage considéré.

La migration mesure un stock et comprend la migration volontaire et la migration forcée. Les statistiques montrent que les très grandes vagues migratoires ont récemment diminué, au profit d'une tendance à l'immigration choisie, hélas favorable à la fuite des cerveaux et compétences des pays pauvres vers les pays riches.

Dans les mouvements migratoires, on distingue les migrations internationales et les migrations internes. Les premières concernent les personnes qui migrent de l'étranger vers la l'Algérie (immigration) et de l'Algérie vers l'étranger (émigration).

III 1-1 Migration internationale

Une migration humaine est un déplacement du lieu de vie d'individus. C'est un phénomène probablement aussi ancien que l'humanité. Les statistiques officielles évaluent entre 185 et 192 millions le nombre de migrants internationaux pour les années 2000, pour les personnes ayant quitté leur pays pour vivre et se fixer dans un autre pays pour au moins un an. Ce chiffre augmente de 2 % par an, malgré les restrictions à l'immigration qui ont vu le jour dans de nombreux pays.

Les caractéristiques du phénomène migratoire actuel sont la diversification des pays de provenance et de destination, ainsi que les formes prises par la migration. On estime que l'argent injecté dans les pays d'origine en provenance des pays d'accueil est au moins égal si ce n'est très supérieur à la quantité d'aides financières apportées par les pays dits « riches » aux pays plus pauvres. Les démographes considèrent que les migrations seront une importante variable d'ajustement d'ici 2050, échéance à laquelle 2 ou 3 milliards d'individus supplémentaires sont attendus sur la planète.

L'Algérie est un pays d'immigration, en 1995 selon les immatriculations consulaires algérienne 1 073 000 Algériens vivaient à l'étranger, dont 92.5% en Europe, plus précisément en France (86.9%). cependant, les statistiques montrent que l'Algérie n'est plus le pays maghrébin ayant une grande population émigrée

III 1-2 Migration inter-wilayas

L'émigration algérienne est intimement liée à son histoire politique et économique récente. De pays à dominante rurale (près de 85% au début du siècle), l'Algérie a connu des bouleversements démographiques importants qui ont marqué son visage actuel.

L'Algérie demeure un pays à urbanisation rapide. Accentué par une forte intensité des migrations internes notamment de l'exode rural, résultat de l'industrialisation, ce phénomène

a atteint les proportions les plus élevées pendant les années 1960 et 1970 et a conduit à l'étouffement des villes. Un habitat précaire est né et s'est développé, ce qui a entraîné un besoin pressant de l'extension des centres urbains.

La forte croissance démographique, l'urbanisation rapide de la population et enfin le processus d'industrialisation accentuent la pression sur les infrastructures, et en l'absence suffisante de l'offre, le développement de la couverture des besoins des populations tend à se dégrader.

Partant du postulat que les urbains ont une mobilité plus faible que les ruraux, nous allons étudier l'évolution de la population urbaine en Algérie.

III 1-3 Evolution de la population urbaine en Algérie de 1886 à 1998

La population urbaine en Algérie n'a cessé d'augmenter par rapport à la population rurale. De 13,9 % en 1886, la population urbaine a atteint une moyenne de 58.30% en 1998 (tableau1). L'accroissement des villes en Algérie attire une population de plus en plus importante. En même temps, la population rurale a diminué pour ne former qu'une faible minorité dans certaines villes. L'espace urbain s'étend de plus en plus sur les terres agricoles. Ces éléments favorisent la fuite des ruraux vers les villes.

Tableau 3-1: Evolution de la population urbaine en Algérie de 1886 à 1998

	Population			%
	Urbaine	Rurale	Total	Population Urbaine
1886	523 431	3 228 606	3 752 037	13,9
1931	1 247 731	4 654 288	5 902 019	21,1
1966	3 778 482	8 243 518	12 022 000	31,4
1987	1 444 249	11 594 693	23 038 942	49,7
1998	16 966 937	12 133 916	29 100 863	58,3

Source : Office National des Statistiques

III 1-3-1 Répartition de la population selon les trois grands ensembles de l'espace physique algérien

« La structure spatiale du pays entièrement polarisée sur le Nord, et plus particulièrement sur la bande côtière est retracée dans le tableau III-2. La densité moyenne dans la bande littorale où se concentrent plus d'un tiers de la population (37.8%) est de 245 hab. /km², alors que la densité moyenne du pays est de 12.22.

Neuf algériens sur dix vivent dans le nord du pays (du littoral à la limite nord de l'Atlas saharien), sur un peu plus d'un dixième de la superficie du pays (12.6%).

Inversement un dixième seulement de la population vit dans plus de 2 000 000 km², soit 88% du territoire.

Tableau 3-2 : Structure géographique de l'Algérie

	Superficie approximative (Km ²)	Population en millions hab.		Densités hab. /km ²	
		1987	1998	1987	1998
BANDE LITTORALE	45 000 (1,9%)	8 900 (38,6%)	11 000 (37,8 %)	222,5	245
TELL ET STEPPE	255 000 (10,7%)	12 139 (52,7%)	15 300 (52,6 %)	49,6	60
SUD	2 080 000 (87,4%)	2 000 (8,7%)	2 800 (9,7 %)	0,9	1,35
ALGE RIE	2 381 741 (100%)	23 039 (100%)	29100 (100%)	9,6	12,22

Source : Office National des Statistiques

Cette structure géographique profondément déséquilibrée donne la mesure du problème d'aménagement du territoire auquel l'Algérie doit faire face. L'avenir proche est en effet engagé sur cette partie du territoire, inéluctablement son caractère urbain va s'accroître au risque de devenir exclusif. En conséquence, le déséquilibre entre l'intérieur et le Nord a toutes les chances de peser d'avantage.

La variabilité des densités de la population est constatée non seulement entre les trois grands ensembles de l'espace physique algérien, mais elle est également présente à l'intérieur de chacune de ces zones :

Dans la bande littorale exception faite pour la wilaya d'Alger 3144hab/km², les densités de population varient de 572hab/km² dans la wilaya d'Oran, et 497hab/km² dans la wilaya de Blida et 105hab/km² pour la wilaya de Tlemcen.

Dans les hautes plaines, le déséquilibre est encore plus important : on passe d'une densité de plus 370hab/km² pour la wilaya de Constantine à 12 hab. /km² pour la wilaya de Djelfa ou encore seulement 02hab/km² pour la wilaya d'El Bayadh.

Même cas pour les wilayas du Sud puisque la densité varie de 27.5hab/km² pour les wilayas de Biskra à moins d'un habitant pour la wilaya de Tindouf 0.17 hab. /km² ou la wilaya d'Ilizi 0.12hab/km²» (H.Remaoun).

III 1-4 La logique de la migration internationale de la population algérienne

« L'émigration algérienne est très ancienne et remonte à la seconde moitié du XIX^{ème} siècle. Mais c'est surtout en réaction au processus de colonisation que des tribus entières quittèrent le pays pour s'installer dans les pays limitrophes, cherchant ainsi à fuir, le gouvernement et l'ordre coloniaux.

Durant la Guerre de Libération Nationale 1954-62, l'armée et les autorités coloniales françaises ont planifié et mis en oeuvre une politique de regroupements de population et de création de «zones interdites», engendrant ainsi des déplacements importants de populations civiles des zones rurales et montagnaises vers les villes, ce qui a eu pour effet de déstructurer l'économie traditionnelle nationale.

Ce mouvement a été amplifié par l'exode rural qui a suivi l'indépendance, exode encouragé par la politique nationale d'industrialisation et par le salariat qui représentait un revenu stable et régulier pour les travailleurs qui étaient habitués à une autre forme de revenu. Les campagnes se sont vidées de leur substance au profit des villes du pays et de l'étranger (la population urbaine est estimée à près de 50% en 1987 alors qu'elle n'était que de 31% en 1966).

La migration interne a concerné près de 2.5% de la population de 1987 (environ 570 000 personnes ont effectué une migration durant la période (1977-1987) contre 5.6% de la population en 1977.

Nous pouvons schématiser en disant que l'émigration algérienne commence à devenir importante avec le début du XX^{ème} siècle, et en particulier suite aux deux guerres mondiales qui ont vu la participation, de nombreux jeunes algériens enrôlés malgré eux dans l'armée française. C'est à l'occasion de cette expérience que de nombreux Algériens découvrent l'Europe.

L'effort de reconstruction des pays européens et en particulier de la France a drainé vers ce pays une main d'oeuvre nombreuse et bon marché surtout d'origine rurale. Cette main d'oeuvre a été essentiellement confinée aux travaux les plus pénibles ; ce mouvement continua et fut favorisé par le patronat européen jusque vers le début des années 1970.

Après cette date, un coup de frein sévère fut donné aux mouvements migratoires à destination de l'Europe. La France adoptera une politique restrictive qui marquera un arrêt brutal de la migration des Algériens qui, traditionnellement allaient, vers l'ancienne métropole coloniale.

Commença alors un mouvement de regroupement familial qui aura pour conséquence de maintenir en vie un flux migratoire qui était, avant la crise pétrolière de 1973-74, très dynamique. Mais la possibilité d'émigrer au moyen du regroupement familial ne devait pas durer indéfiniment (les entrées en France pour raison de regroupement familial passent ainsi de plus de 9000 personnes en 1982 à moins de 3000 en 1996).

Et c'est ainsi que vers le milieu des années 1980, ce mouvement de regroupement qui était déjà faible commença à se tarir et nous assistâmes alors, à un redémarrage de la migration individuelle masculine mais non plus seulement en direction de la France bien que ce

mouvement reste encore dominant mais également vers d'autres destinations, en particulier le Canada, dont la politique migratoire est plus souple.

Nous pouvons donc résumer le mouvement migratoire algérien en trois phases historiques, une première phase qui commence avec le début du siècle, puis une seconde qui s'étale grosso modo sur les années 1980, puis une troisième phase qui est en cours actuellement. Ces trois phases historiques s'accompagnent pour les migrants de caractéristiques démographiques et sociales très distinctes.

En effet dans la 1ère phase, ce sont surtout des paysans illettrés qui pour certains n'ont vu d'une ville dans leur propre pays que la gare maritime ; dans la seconde phase, ce sont leurs épouses et leurs enfants restés au pays qui les rejoignent presque dans les mêmes conditions. Cependant, dans cette seconde phase, la présence des enfants est un élément important car, à la différence de leurs parents, ils ont connu l'école publique. Le regroupement familial aura d'ailleurs des conséquences importantes pour les émigrés algériens en ce sens qu'ils ont franchi un pas important dans le processus d'intégration à travers leurs enfants qui ont grandi et se sont imprégnés des valeurs des pays d'accueil.

La troisième phase qui est la phase actuelle, est radicalement différente des deux autres (il faut cependant souligner que ce phasage n'a rien d'hermétique et que nous pouvons certainement trouver ça et là des chevauchements de ces processus migratoires) en ce qu'elle concerne non plus de jeunes paysans algériens qui ont connu l'exode rural, mais surtout des citadins porteurs de diplômes souvent universitaires. Cette troisième phase coïncide avec le démantèlement de l'économie administrée engagée par notre pays et le passage vers l'économie de marché. Dans cette phase, l'émigration n'est pas totalement masculine comme dans la 1ère phase même si elle est encore dominée par l'élément masculin. Elle n'a pas, comme par le passé, une destination unique privilégiée, à savoir la France » [30].

III 1-5 La migration inter-wilayas de la population

« La migration est un phénomène démographique complexe qui, en faisant intervenir deux dimensions, le temps et l'espace, pose des problèmes de définition et d'analyse.

Selon le dictionnaire démographique, le mot migration désigne le déplacement d'une personne entraîné par un changement de résidence. La précision de cette définition dépend de celle avec laquelle est définie la notion de résidence.

Le résident, est la personne, le membre de ménage habitant dans le logement qui est sa résidence habituelle depuis plus de 6 mois. Exception faite pour les nouveaux nés, nouveaux mariés, ou les personnes nouvellement installées avec l'intention d'habiter définitivement le logement.

La notion de résidence est ainsi définie à l'aide de deux critères ; on parle de résidence habituelle : tout individu même s'il est en déplacement temporaire, ou saisonnier, est rattaché

au lieu où il a coutume d'habiter. On fixe une durée de plus de 6 mois, à partir de laquelle le déplacement temporaire devient définitif » [34].

L'exploitation de cette question portant sur les mouvements migratoires nationaux de la période intercensitaire est l'objet de notre application, dont nous dégageons ici les principaux résultats.

III 1-5 -1 Application du flux spatial Origine-Destination

Le recensement général de la population et de l'habitat réalisé en juin 1998, constitue la source principale de données concernant les migrations internes de la population algérienne. Il comporte plusieurs questions permettant de se situer dans le temps et dans l'espace.

Ainsi, il était notamment demandé à toutes les personnes recensées en juin 1998 où elles résidaient à la date du précédent recensement 1987.

Sur les 29 millions de personnes recensées en 1998, 21,5 millions de personnes sont nées avant 1987, 84% n'avaient pas changé de wilaya de naissance (sédentaires), 14.7% ne résident pas au 25 juin 1998 dans leurs wilaya de naissance et 4.5% soit 940.000 des ménages ordinaires et collectifs ont changé de wilaya de résidence depuis 1987 (migrants intercensitaires), l'objet de notre étude se limite uniquement au changement de résidence.

On considère comme migrant intercensitaire 1987, toute personne ayant déclaré une wilaya de résidence en 1987 différente de celle de 1998 ; cette définition exclut :

- La migration des personnes décédées entre 1987 et 1998 ou ayant émigré à l'étranger.
- La migration des enfants nés entre 1987 et 1998.
- La migration des personnes à l'intérieur de la wilaya.
- La migration brève qui été suivie d'un retour à la wilaya de résidence d'origine.

Tableau 3-3 : Flux migratoires (1987-1998)

		Vers (j)					
De (i)	Wilaya	ADRAR	LAGHOUAT	...	AIN TEMOUCHENT	RELIZANE	Total
	ADRAR	0	245	...	124	115	8248
	LAGHOUAT	160	0	...	26	136	10635
	⋮	⋮	⋮		⋮	⋮	⋮
	AIN TEMOUCHENT	251	59	...	0	588	13474
	GHARDAIA	838	554	...	32	87	8194
	RELIZANE	338	86	...	132	0	11725
	Total	30092	11572	...	7580	27668	938 537

L'analyse du vecteur flux migratoires Origine-Destination de dimension $N=n^2$, où n représente les 48 wilayas, $N=2304$. L'étude de ce grand nombre de flux n'a pu être accomplie que grâce à l'introduction du logiciel ArcGis (logiciel de cartographie), la carte ci-dessous (Carte III-1) fait apparaître les principaux flux migratoires de 1987 à 1998. Toutefois nous donnons un aperçu du vecteur des flux migratoires.

Tableau 3-4 Les trois principaux flux migratoires

Wilaya d'Origine	Wilayas de Destination	Flux migratoire	Pourcentage
Médéa	Djelfa	7 833	0.83 %
	Blida	17 342	1.85 %
	Alger	29 389	3.13 %
Alger	Sétif	5 191	0.55 %
	Tipaza	6 220	0.66 %
	Bejaia	7 326	0.78 %
	Tizi-Ouzou	11 658	1.24 %
	Blida	19 470	2.07 %
	Boumèrdes	21 678	2.31 %
Blida	Tipaza	3 786	0.40 %
	Alger	20 621	2.20 %

L'examen des flux migratoires entre les wilayas fait apparaître un fort déséquilibre dans les échanges entre les wilayas.

On remarque que la capitale semble avoir atteint un point de saturation, provoquant des départs vers les wilayas limitrophes, comme Blida, Boumerdes, Tizi-Ouzou. Mais Alger continue à être une wilaya attractive malgré les efforts visant à la désengorger.

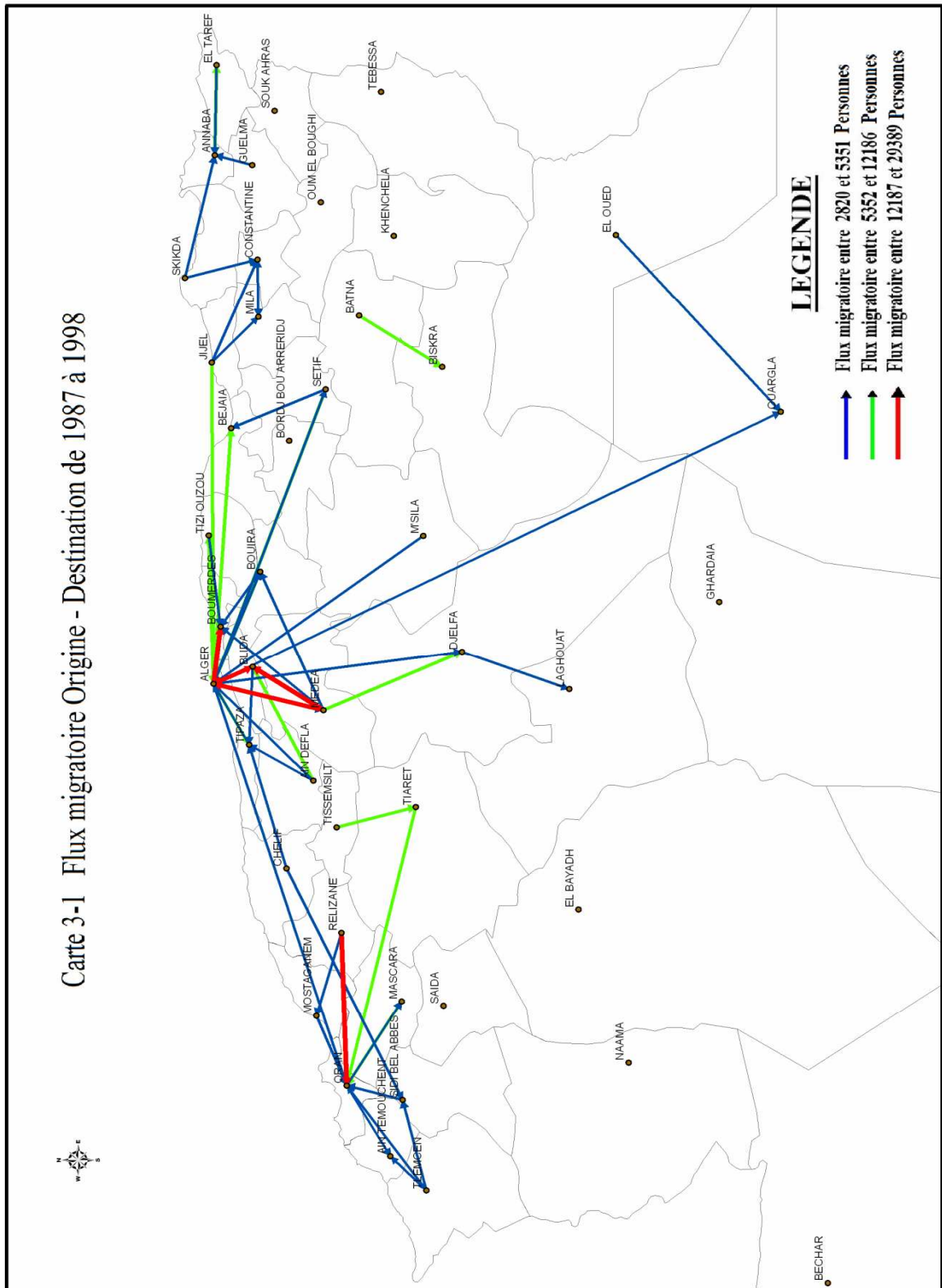
En remarque aussi que les wilayas de Médéa et de Relizane connaissent un fort taux de déficit migratoire¹, car ce sont des wilayas enclavées, à prédominance rurale.

La métropole de l'est, Annaba, comme le cas de la capitale, semble avoir atteint le point de saturation, et provoque des départs vers une unique destination : El Tarf.

En conclusion, la redistribution de la population sous l'effet des migrations se fait au profit d'un nombre limité de wilayas accentuant de ce fait, les déséquilibres existants.

¹ Voir annexe D

Carte 3-1 Flux migratoire Origine - Destination de 1987 à 1998



III-2 Structure spatiale de l'Algérie

Au-delà de l'évidence que l'homme déploie ses activités économiques dans l'espace. Ce dernier n'apparaît pas seulement comme une dimension, un support de ses activités. Il en est une variable souvent déterminante, imprégnant les décisions et les interactions jusqu'à engendrer des phénomènes économiques qu'on ne pourrait même pas évoquer si on n'avait pas une vue essentiellement spatiale de l'économie. Car la concurrence entre les collectivités territoriales repose essentiellement sur le découpage du territoire.

III 2-1 Le découpage du territoire algérien de 1515 à nos jours

La régence turque qui s'établit en Algérie, à partir de 1515, impose pour la première fois à l'espace algérien du Nord, un découpage régional précis.

Le Nord est divisé en trois beyliks (entités administratives), du Ponant, du Titteri et du Levant. Ces trois entités administratives sont placées sous l'autorité d'un Bey dépendant directement du Dey d'Alger, nommé lui-même par le pouvoir turc. Le beylik du Levant, est constitué de l'Algérie orientale, avec pour centre politique Constantine, et comme point d'ouverture sur la Méditerranée, le port de Bône. Le beylik du Titteri, en position centrale, a pour capitale Médéa et pour port Dellys. Enfin le dernier beylik, du Ponant, localisé dans l'Algérie occidentale, s'organise autour de Mascara, et d'Arzew comme base portuaire. Alger et la plaine de la Mitidja, constituent un quatrième territoire, appelé Dar-es-Soltan, sous l'autorité d'un Agha relevant directement du Dey. La capitale du pouvoir turc est installée à Alger sur le littoral, ce qui constitue une rupture avec les précédents royaumes qui s'appuyaient sur des capitales localisées préférentiellement dans le Tell et les Hautes Plaines. L'autorité turque a ainsi mis en place un découpage tripartite, divisions qui furent reprises par l'autorité française au XIXe. Ce découpage adhérent à l'architecture physique du pays, s'avéra relativement fonctionnel, et reste encore aujourd'hui un élément prégnant de la réalité algérienne. La fonctionnalité de ce découpage est due également à la longévité de l'empreinte turque en territoire algérien du XVIe au XIXe.

Avec l'affirmation d'un découpage, les frontières Est et Ouest du pays purent s'affirmer. Par contre, au Sud, aucune délimitation ne fut imposée, l'espace saharien était perçu comme flou et lointain. Il est vrai que le déclin du commerce transsaharien associé à la difficulté de contrôler des populations nomades, n'incita pas le pouvoir turc à installer plus à l'intérieur du pays ses structures administratives et militaires. Pour pouvoir aborder le Sahara algérien en termes d'espace construit, aux sens juridique et géographique, il fallut attendre la période française, qui donna corps à cette entité.

« Le Sahara algérien est pour la première fois délimité, ses frontières sont créées par l'autorité militaire française qui trace à la règle ces limites au travers d'espaces désertiques. Les territoires militaires, sont bornés au Nord par l'Atlas Saharien ce qui constitue une barrière naturelle, rendant quelque peu légitime ou compréhensible ce découpage.

Par contre en ce qui concerne le sud, le découpage semble ne répondre à aucune logique quelle soit physique ou encore ethnique. En effet, le massif que l'Ahaggar, se prolonge au-delà de la frontière algérienne notamment vers le Niger, dans la région l'Aïr. Et la frontière, scinde le territoire occupé par le groupe de population Touareg en plusieurs espaces nationaux (Mali, Niger), fractionnant ainsi des parcours nomades, des liens tribaux ou encore économiques.

Ainsi, en cette première moitié du XX^{ème} siècle, les frontières algériennes sont fixées, clairement tracées et donnent à l'espace algérien un fort ancrage africain de par sa pénétration importante vers le sud. L'espace saharien couvre dès lors environ deux millions de kilomètres carrés et représente donc les 4/5 du territoire algérien » [03].

III 2-2 Le découpage du territoire algérien durant la période coloniale (1830-1962)

La colonisation française a transposé le découpage spatial turc, qui a perduré un siècle dans les trois départements de Constantine, Alger et Oran. Cependant, « le principe d'organisation changea de nature car l'administration coloniale française avait mis au point un système politique dont le fondement reposait sur une distinction sociale et spatiale. Le résultat fut une dichotomie administrative d'ordre politique et colonial : il s'agissait d'intégrer et d'aménager des espaces occupant les meilleures terres pour la population européenne et de cantonner la population autochtone sur les secteurs plus pauvres » (Raham D., 2004, p256).

La subdivision de chacun des trois départements en territoires civils, avait créé puis entretenu le processus de désintégration des tribus et du système tribal et une certaine dissociation des métropoles régionales et de leur arrière-pays méridional. Vers la fin de l'époque coloniale, la population ayant doublé et l'économie s'étant complexifiées, la gestion nécessita une organisation administrative démultipliée. En 1956, le nombre de départements fut multiplié par cinq, un statut unique régissant toutes les communes. Cette densification du maillage fut accompagnée d'un vaste programme d'expansion économique et social.

Tableau 3-5 Évolution des circonscriptions administratives en Algérie

Années	Wilayas	communes
1958	15	1525
1963	15	676
1974	31	704
1984	48	1540

III 2-3 Les réorganisations territoriales de la période post-indépendance

Durant la Guerre de Libération, le territoire algérien est découpé en six wilayas : l'Oranais, l'Algérois, la Kabylie, le Constantinois, l'Aurès, et le Sud. Alger retrouve un statut proche de celui qui lui était conféré par l'administration turque, le Dar es Soltan, et constitue une zone autonome.

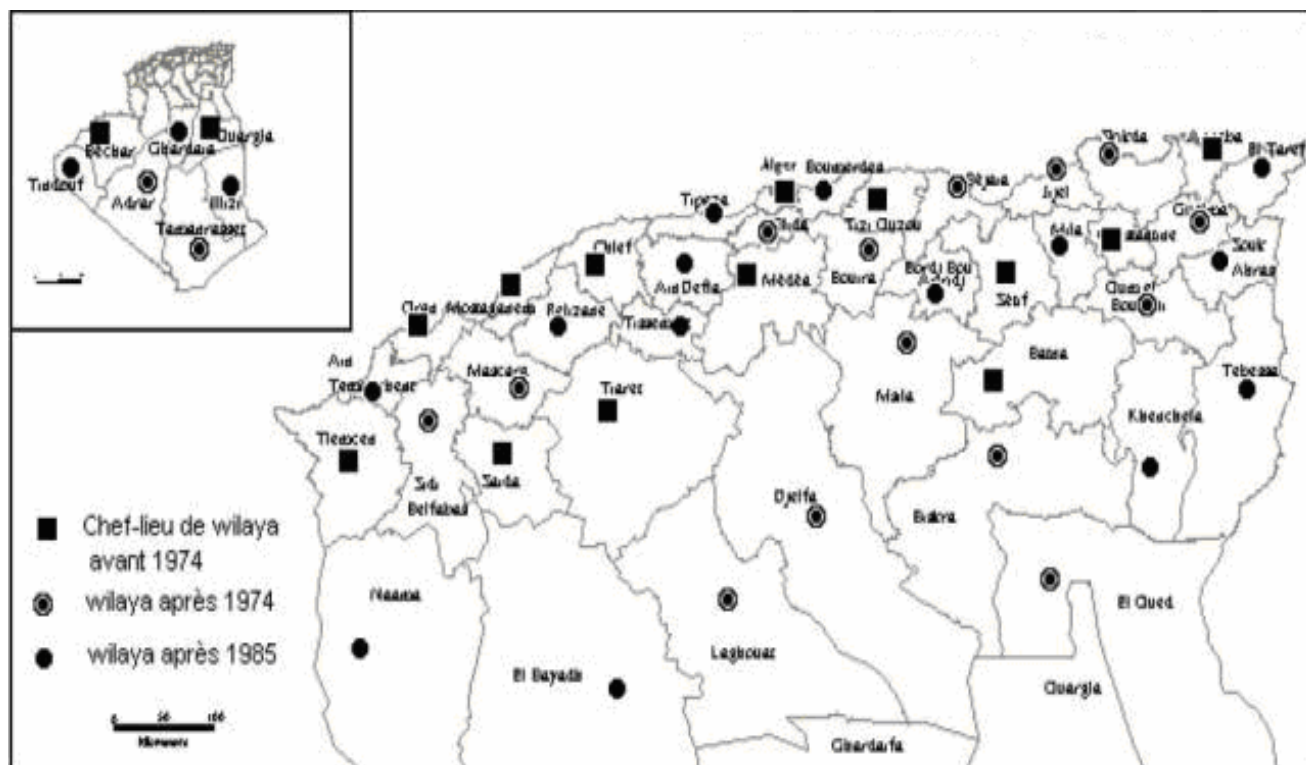
Au lendemain de l'Indépendance, l'intervention sur le maillage territorial s'inscrivait dans une conjoncture où la maîtrise du territoire l'emportait sur la logique spatiale : « sitôt l'indépendance acquise, l'une des préoccupations majeures de l'Algérie fut d'asseoir son autorité et de renforcer le contrôle de son territoire, (Troin, 1985, p. 312). Le nombre de départements, rebaptisés wilayas, fut maintenu mais les communes regroupées (voir tableau 3-5).

III 2-3-1 Les réformes administratives de 1974

Un nouveau découpage intervient en 1974 (voir Carte 3-2), avec pour vocation de créer un maillage plus cohérent et efficace, afin de favoriser le développement économique du pays. L'Etat algérien voulait ainsi gommer davantage les disparités créées par l'ancienne puissance coloniale et engendrer un développement solidaire entre wilayas prospères et wilayas économiquement en difficulté. Le nombre de wilayas passe de quinze à trente et une et les dénominations des wilayas se simplifient, chaque wilaya prend dès lors le nom de son chef lieu.

On a donc cherché à réduire les disparités régionales héritées du système macro régional de la période coloniale. La logique de la nouvelle partition territoriale visait une homogénéisation de la trame administrative : promouvoir des wilayas plus comparables par la taille et appliquer une seule grille d'équipements. Il en est résulté une augmentation sensible du nombre de circonscriptions et une diminution de la taille des mailles, y compris de la maille communale.

Carte 3-2 Evolution des wilayas (1974 -1984)



Source : Mohamed Chadli et Ali Hadjiedj, « L'apport des petites agglomérations dans la croissance urbaine en Algérie », *Cybergeo*, Espace, Société, Territoire, article 251.

III 2-3-2 Les réformes administratives de 1984

Malgré un effort de réflexion, le découpage de 1974 présente encore quelques imperfections (voir Carte 3-2), certaines limites de wilaya ne sont basées que sur des particularités physiques, ou encore sur des spécificités économiques comme c'est le cas pour la wilaya de Ouargla. Il y a donc parfois incohérence entre découpage administratif et espaces fonctionnels. Le découpage de 1984, appliqué à partir du 1er janvier 1985, tente de remédier à ces incohérences en multipliant le nombre de wilaya. Celles-ci sont désormais quarante-huit, le nombre de daïra augmente de 160 à 229, ainsi que le nombre de communes qui passe alors à 1541. Le territoire algérien devient plus finement maillé, ce qui permet une action en matière d'aménagement plus ciblée, et davantage en cohérence avec les préoccupations et attentes locales. Ces deux derniers découpages étaient marqués par une volonté d'un développement harmonieux du territoire, volonté qui s'est exprimée notamment par le biais des politiques de promotion administrative et industrielle.

Conclusion

Derrière cet objectif de découpage équitable du territoire national, on pourrait également lire les intentions politiques de l'Etat algérien. Un niveau intermédiaire a, en outre, été introduit : celui de la *daïra*. Le découpage du territoire algérien en trois niveaux, wilaya, daïra et commune, emboîtés et hiérarchisés, tend à masquer et à réduire les particularismes régionaux et locaux .

« Le niveau de la « région » n'existant pas en tant qu'entité économique et sociale en soit. En effet, la wilaya est une représentation purement politique et administrative de l'autorité centrale. Les 48 wilayas se rattachent directement à l'Etat. Bien que l'effet négatif de l'absence du niveau régional dans les programmes de développement local soit reconnu par la classe politique dirigeante, il reste néanmoins posé pour les politiques d'aménagement futures » [05].

CHAPITRE IV

**Application des outils de l'économétrie
spatiale à la modélisation des flux
Origine_Destination**

Introduction

Les principales bases de la théorie du processus Origine-Destination sont :

Chaque région ou wilaya joue le rôle d'une force sélective ou magnétique, attirant à elle les éléments de population appropriés et repoussant les éléments inassimilables. Elle va ainsi dans le sens des divisions biologiques et culturelles de la population d'une ville.

Le développement de ce type de modèle a nécessité l'utilisation du logiciel MATLAB à environnement interactif, c'est un langage de haut niveau permettant l'exécution de tâches nécessitant une grande puissance de calcul et dont la mise en œuvre est plus rapide. L'objet de ce mémoire est l'estimation des approches alternatives à l'économétrie classique qui peuvent être utilisées lorsque l'on traite avec des échantillons de données spatiales.

Ce chapitre repose essentiellement sur l'article de J.LeSage, article publié en 2006 dans Social Science Research [17].

IV Champs d'application du modèle économétrique inter wilaya Origine-Destination

Notre application repose sur l'étude d'un groupe de trois méthodes économétriques pour la modélisation de l'origine et la destination (OD), en utilisant les matrices contenant le flux interrégional.

Elles ont des structures de données utilisées généralement dans de nombreux domaines, l'économie, le commerce international, le transport, la migration, la recherche du flux d'information et de communication, les mouvements pendulaires, ainsi que les modèles économiques régionaux et interrégionaux. Par contre dans le modèle économétrique spatial où l'échantillon est formé de n régions, où chaque région devient une observation, ces modèles nécessitent $n^2=N$ paires origine-destination où chaque paire origine-destination devient une observation.

Le terme « modèle à interaction spatiale » a été utilisé dans la littérature pour caractériser les modèles qui se basent sur les flux entre origines et destinations (Sen et Smith 1995). Ce type de modélisation cherche à expliquer la variation dans le niveau des flux par rapport à l'échantillon de taille N des paires OD. Ces modèles dépendent de la fonction de distance entre l'origine et la destination aussi bien que des variables expérimentales pertinentes qui permettent de caractériser aussi bien les régions d'origine que les régions de destination.

L'objet de notre étude est de fournir une méthode spécifique de régression spatiale qui puisse être appliquée à des modèles à interaction spatiale. Les travaux de Porojon (2001) pour le cas du flux du commerce international et ceux de Lee and Pace (2004) pour les ventes au détail, ont pointé du doigt que les résidus qu'ils ont trouvés, dévoilent une forme de dépendance spatiale.

Les modèles autorégressifs spatiaux standards, qui sont reliés à des structures de poids spatial construites pour modéliser la dépendance entre n régions. Mais la nature même d'un modèle qui prend en charge un échantillon de données de flux interrégional, engendre ($N=n^2$)

paires d'Origine-Destination, ce qui est un obstacle pour la modélisation, pour outre passer ce problème on propose une structure de poids spatial qui soit en adéquation avec les modèles autorégressifs spatiaux. Ce sont une extension des modèles de régression spatiale introduits par Anselin (1988).

IV-1 Origine-Destination : Notation et ordre

Soit Y une matrice ($n \times n$) mesurant le flux interrégional de chaque n région d'origine à chaque n régions de destinataire où les n colonnes représentent les différentes origines et les n lignes représentent les différentes destinations.

En conséquence, $n^{-1} Y \mathbf{1}_n$ est la moyenne du flux de chaque n d'origine pour toutes les destinations, où $\mathbf{1}_n$ est un vecteur unitaire ($n \times 1$).

D'une manière similaire, $n^{-1} Y' \mathbf{1}_n$ est la moyenne du flux pour chaque n destinations pour toutes les origines.

On peut produire un vecteur $N (=n^2)$ par 1 de ces flux de deux manières, la première reflète un ordre « Origin-centric » comme (4.1), et l'autre reflète l'ordre « Destination-centric » comme en (2). Les indices l, o, d dénotent dans l'ensemble les indices de $1, \dots, N$ et l'origine, les indices de $1, \dots, n$ des destinations.

Commençons par la matrice Y dont les colonnes reflètent les origines et les lignes les destinations ; on obtient un ordre « Origin-centric » avec $y = \text{vec}(Y)$, et l'ordre « Destination-centric » par le paramétrage $y^{(d)} = \text{vec}(Y')$. Ces deux ordres sont reliés par la matrice $\text{vec} - \text{permutation}$ donc $P y = y^{(d)}$, ainsi que par les propriétés de la permutation des matrices $y = P^{-1} y^{(d)}$. Pour la suite des démonstrations, on se focalisera sur l'ordre Origin-Centric où les n premiers éléments dans la sommation du vecteur y reflètent le flux de l'origine 1 aux n destinations. Les derniers n éléments de ce vecteur représentent les flux de n origine aux autres destinations de 1 à n .¹

$$\begin{array}{ccc}
 1^{(0)} & 1 & 1^{(0)} \\
 & (0) & \\
 1 & 1 & 1 \\
 \vdots & \vdots & \vdots \\
 n & 1 & n \\
 \vdots & \vdots & \vdots \\
 N-n+1 & N & 1 \\
 \vdots & \vdots & \vdots \\
 N & N & n
 \end{array} \quad (4.1)$$

$$\begin{array}{ccc}
 1^{(d)} & 1 & 1^{(d)} \\
 & (d) & \\
 1 & 1 & 1 \\
 \vdots & \vdots & \vdots \\
 n & N & 1 \\
 \vdots & \vdots & \vdots \\
 N-n+1 & 1 & n \\
 \vdots & \vdots & \vdots \\
 N & N & n
 \end{array} \quad (4.2)$$

IV-2 Les modèles de gravités avec indépendance des observations

Le modèle de gravité conventionnel avec l'approche de la régression des moindres carrés pour expliquer les variations dans le vecteur du flux Origine-Destination repose sur la construction d'une matrice $(n \times k)$ des variables explicatives que l'on va étiqueter X , qui contient k caractéristique pour chaque n régions. Et sans nuire à la généralité, soit chaque colonne de X possédant 0 comme moyenne-différentielle. Etant donné le format du vecteur y , où les observations de 1 à n reflètent le flux de l'origine 1 pour toutes les destinations.

Cette matrice est répétée n fois pour produire une matrice N par k qui représente les caractéristiques des destinations que l'on va étiqueter X_d . On note que X_d est égal à $\iota_n \otimes X$, où ι_n est un vecteur unitaire $(n \times 1)$. La deuxième matrice représente les caractéristiques des origines que l'on va étiqueter X_0 .

Ceci fait répéter les caractéristiques de la première région n fois pour former les premières n lignes de X_0 , les caractéristiques de la deuxième région n fois pour les prochaines n lignes de X_0 ainsi de suite, jusqu'à la construction d'une matrice $(N \times k)$, que l'on va étiqueter $X_0 = X \otimes \iota_n$.

La distance de chaque origine pour chaque destination est aussi incluse comme vecteur variable explicatif dans le modèle de gravité. On a G qui représente une matrice n par n , de la

distance entre l'origine et la destination, ainsi $g = \text{vec}(G)$ est un vecteur $(N \times 1)$. Puisque G est symétrique, $g = \text{vec}(G')$ donne le même résultat). Sans nuire à la généralité soit g un vecteur ayant comme moyenne des zéro. Ce résultat est montré dans le modèle de régression suivant :

$$y = \alpha \iota_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon \quad (4.3)$$

Dans (4.3) les matrices des variables explicatives X_d , X_o de dimension $(N \times k)$ représentent des matrices qui contiennent les caractéristiques respectivement des destinations et des origines, associé aux vecteurs des paramètres k par 1 qui sont β_d et β_o . Le paramètre du scalaire γ reflète l'effet de la distance g , et α dénote le terme constant sur ι_N . Le vecteur $\varepsilon = \text{vec}(E)$ de dimension $(N \times 1)$ représente une perturbation, on suppose que $\varepsilon \approx N(0, \sigma^2 I_N)$.

Quelques manipulations élémentaires sur les moments de la matrice peuvent nous éclairer sur la simplicité du modèle de gravité basé sur l'indépendance des observations pour le cas de la matrice carrée, où chaque origine est aussi une destination. L'Algèbre du produit de Kronecker¹, peut être utilisé pour former une matrice des moments sans traiter directement avec une matrice N par N . soit les matrices arbitraires suivantes A , B , C , $(C \otimes A)\text{vec}(B) = \text{vec}(ABC)$ ².

En utilisant $Z = [\iota_N \ X_o \ X_d \ g]$ cela nous conduit au moment de la matrice (4.4).

$$Z'Z = \begin{bmatrix} N & 0 & 0 & 0 \\ 0 & nX'X & 0 & X'G\iota_n \\ 0 & 0 & nX'X & X'G\iota_n \\ 0 & \iota_n'G'X & \iota_n'G'X & \text{tr}(G^2) \end{bmatrix} \quad (4.4)$$

On note que $\text{tr}(G^2)$ peut être calculé d'une manière plus efficace en utilisant $\iota_n'(G \Theta G)$ ι_n , où l'opérateur Θ fait référence à la multiplication (d' Hadamard)³, d'une manière similaire $Z'y$ peut s'exprimer comme suit :

$$Z'y = [\iota_n'Y\iota_n \ X'Y\iota_n \ X'Y'\iota_n \ \text{tr}(GY)] \quad (4.5)$$

Où le calcul de $\iota_n'(G \Theta Y) \iota_n$ est plus efficace à calculer que $\text{tr}(GY)$.

Le moment de la matrice peut être réécrit pour obtenir un aperçu des estimations des moindres carrées $\hat{\beta} = \left(\frac{Z'Z}{N} \right)^{-1} \left(\frac{Z'y}{N} \right)$.

¹ (Voir annexe C.2)

² (Horn and Johnson, 1991, p.255, Lemma)

³ (voir Annexe C.1)

$$\frac{Z'Z}{N} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{nXX}{n} & 0 & \frac{X' Gt_n}{n} \\ 0 & 0 & \frac{nXX}{n} & \frac{X' Gt_n}{n} \\ 0 & \frac{t_n'G' X}{n} & \frac{t_n'G' X}{n} & \frac{tr(G^2)}{N} \end{bmatrix} \quad (4.6)$$

$$Z'y = \begin{bmatrix} \frac{t_n'Yt_n}{N} & \frac{X' Yt_n}{n} & \frac{X' Yt_n}{n} & \frac{tr(GY)}{n} \end{bmatrix} \quad (4.7)$$

Les quantités $\frac{X' G' t_n}{n}$ et $\frac{t_n' G' X}{n}$ mesurent la covariance entre la variable explicative et la distance. Dans le cas où chaque origine est aussi une destination pour n'importe laquelle des deux régions i et j il y a deux OD paires, ij paires et ji paires si la valeur des variables explicatives est différente pour i et j il y aura deux valeurs pour les variables explicatives associées à la même distance pour le ij et ji OD paire, et ceci sera aussi le cas pour toute les paires $n(n-1)/2$ ou $i \neq j$ ceci devait être le résultat d'une petite variance entre la distance et les variables explicatives. On note que pour $\frac{X' G' t_n}{n}$ et $\frac{t_n' G' X}{n}$ ont des valeurs suffisamment petite, pour que l'on puisse considérer les éléments de la diagonale $Z'Z$.

Où l'inverse peut être obtenu bloc par bloc.

On peut exploiter la structure du block diagonal qui peut se présenter si l'origine, la destination, et les variables de distance exhibent une petite covariance pour produire une estimation associée avec les caractéristiques de l'origine :

$$\hat{\beta}_o = (X'X)^{-1} X' [n^{-1} Y t_n] \quad (4.8)$$

Où $n^{-1} Y t_n$ représente la moyenne des flux de n origines vers toutes les destinations.

D'une manière similaire, une estimation des caractéristiques de destination est :

$$\hat{\beta}_d = (X'X)^{-1} X' [n^{-1} Y' t_n] \quad (4.9)$$

Où $n^{-1} Y' t_n$ représente la moyenne des flux de n destinations qui proviennent de toutes les origines.

Alternativement, l'estimation du paramètre d'origine $\hat{\beta}_o$ peut être vue comme une moyenne d'estimations à travers n régressions des flux d'origine et leurs caractéristiques d'origines, et $\hat{\beta}_d$ comme une moyenne de n régressions comprenant les flux de destination et les leurs caractéristiques de destination.

Soit $\hat{\beta}_{oj} = (X'X)^{-1} X' Y_j$ (où j est l'indice des colonnes de Y), et $\hat{\beta}_{di} = (X'X)^{-1} X' [Y_i']$ (ou i est l'indice des lignes de Y). Ce sont les estimations locales de l'effet de l'origine et les

caractéristiques de destination. Par exemple, si la troisième colonne de Y correspond aux observations qui ont comme origine Oran, $\hat{\beta}_{o3}$ (ce sont les estimations obtenues par le flux de régression dont l'origine est Oran), alors que $\hat{\beta}_{d3}$ (sont les estimations obtenues par la régression impliquant les flux dont la terminaison est Oran. Par conséquent,

$$\hat{\beta}_o = (X'X)^{-1} X' [n^{-1} Y \iota_n] = \left(\frac{1}{n}\right) \sum \beta_{oj} \quad (4.10)$$

$$\hat{\beta}_d = (X'X)^{-1} X' [n^{-1} Y' \iota_n] = \left(\frac{1}{n}\right) \sum \beta_{di} \quad (4.11)$$

Donc, l'estimation globale de l'effet des caractéristiques de l'origine et de la destination est égale à la moyenne de l'estimation locale des mêmes effets.

Si la covariance entre la distance et les variables explicatives est petite, l'estimation approximative par moindres carrés du paramètre de la distance γ est : $\text{tr}(GY)/\text{tr}(G^2) = g'y/g'g$. C'est l'estimation du paramètre de la pente de la régression simple de y en g. Par conséquent, une faible corrélation de la distance avec les variables explicatives veut dire que l'addition de la distance a ce modèle pourrait améliorer l'ajustement, mais n'affecte pas les estimations des variables explicatives.

En conclusion, le carré ($n^2 = N$) du modèle de gravité sous l'hypothèse de l'indépendance entre le flux de l'origine et de la destination, a une structure très simple qui relie l'estimation globale et locale des caractéristiques de l'origine et de la destination. De plus, la distance a probablement un effet sur l'ajustement, mais pas sur l'estimation associée des variables indépendantes.

La simplicité du modèle de gravité qui suppose l'indépendance des observations, revient à occulter le flux spatial Origine-Destination. Pour enrichir ce modèle simple on le complète avec des formes de dépendance spatiale.

IV-3 La dépendance spatiale dans les flux Origine-Destination

Bien que le modèle de gravité modélise l'interdépendance des observations en utilisant la distance, cette tentative s'avère inadéquate pour plusieurs types de données du flux où chaque région devrait affecter ses régions voisines. Par exemple le voisinage des régions d'origine et de destination doit montrer plus d'erreurs similaires. Où la décision de la migration produite dans une région affecte les décisions de la migration produite au voisinage des régions.

La matrice de poids spatiale fournit un chemin commode et parcimonieux pour spécifier la dépendance spatiale des observations.

Dans les modèles croisés à n régions, où chaque région représente une observation, la matrice de poids spatial W représente une matrice creuse non négative. Cette matrice capture

les relations de dépendances entre les observations (régions). Par exemple $W_{ij} > 0$ si la région i est contiguë avec la région j .

En plus de la contiguïté, des mesures différentes de la proximité comme distance cardinale (ex: kilomètres), et la distance ordinal (ex: les dix voisins proches), par convention $W_{ii} = 0$, pour qu'une observation ne soit définie comme étant voisine à elle-même, et la matrice W est standardisé pour avoir la somme des lignes égale à l'unité.

La clef du problème est de construire une matrice de poids spatiale significative dans le cas où le vecteur des observations ($N \times 1$) reflétant le flux de toutes les origines vers toutes les destinations, plutôt que le cas où chaque observation représente une région. On commence par une ligne standardisé avec un cas typique de matrice W de contiguïté d'ordre ($n \times n$) qui reflète la relation entre les n destinations / aux régions d'origine. Ceci peut être répété en utilisant $(I_n \otimes W)$. On crée une matrice de poids spatiale N par N lignes standardisées que l'on étiquettera W_0 , représentée en (4.12) et où 0_n représente une matrice ($n \times n$) de zéros.

$$W_0 = \begin{pmatrix} W & 0_n & \dots & 0_n \\ 0_n & W & 0_n & \vdots \\ \vdots & 0_n & \ddots & 0_n \\ 0_n & \dots & 0_n & W \end{pmatrix} \quad (4.12)$$

En utilisant cette matrice pour former un décalage spatial de la variable dépendante, $W_0 y$, (où $W_0 = (I_n \otimes W)$, avec W ligne standardisée), capture la relation spatiale de dépendance de « l'origine de base » en utilisant une moyenne des flux des voisins de chaque région d'origine vers chaque destination.

D'une manière triviale, le flux de n'importe quelle origine vers une région de destination particulière doit créer un flux similaire au voisinage de cette origine vers la même destination, c'est l'effet capturé par le décalage spatiale $W_0 y$.

On prend un exemple : on considère une ligne particulière i du vecteur de décalage spatial $W_0 y$ qui représente les flux de l'origine, wilaya d'Annaba vers la destination, wilaya d'Alger. En prenant comme matrice de contiguïté du premier ordre, les voisins de la wilaya d'Origine Annaba qui sont : El Tarf, Guelma, Skikda, et les voisins de la wilaya de destination Alger qui sont : Boumerdes, Blida, Tipaza.

Le décalage spatial $W_0 y$ représente le flux moyen d'El Tarf, Guelma, Skikda (voisins de l'Origine) vers la wilaya de destination Alger.

D'une manière similaire, lorsqu'on applique le décalage spatial $W_0 y$ sur les autres lignes. Par exemple, on examine le flux de la wilaya d'origine Oran vers la wilaya de destination Alger. Le décalage spatial, $W_0 y$ représente le flux moyen de Mostaganem, Mascara, Sidi Bel Abbes, Ain Témouchent (voisins de l'Origine) vers la wilaya de destination Alger.

Un deuxième type de dépendance spatiale qui apparaît dans le modèle de gravité est la dépendance de 'destination de base'. D'une manière triviale, les forces qui conduisent à créer un flux d'une wilaya d'origine vers une wilaya de destination, vont générer un flux similaire à celui créé par les wilayas voisines à la wilaya de destination.

Le poids spatial de la matrice que l'on étiquettera W_d peut être construit pour capturer ces types de dépendances en utilisant $W \otimes I_n$, qui produit (N x N) matrice de poids spatial qui capture les relations entre les flux de wilaya d'origine vers les voisinages des wilayas de destination.

Pour développer les notions de dépendance de destinations de base, on emploie la matrice vec-permutation P introduite précédemment. Si on adopte l'ordre de la « Destination-centric », la spécification de la matrice de poids de la destination serait $I_n \otimes W$ par la même logique comme introduite dans le développement de l'origine de la matrice de poids de base. Par conséquent le poids de la matrice de destination sous l'ordre de « l'Origine-centric » serait $P(I_n \otimes W)P$. quelques résultats du produit de Kronecker avance la simplification de $P(I_n \otimes W)P$. sachant que P est la matrice vec-permutation⁴, $P(I_n \otimes W)P = W \otimes I_n$ ainsi $W_d = W \otimes I_n$.

En utilisant l'exemple du flux de l'origine de la wilaya d'Annaba vers la wilaya de destination Alger, le vecteur de décalage spatial $W_d y$ qui représente les flux moyen d'Annaba vers Boumerdes, Blida, Tipaza, wilayas voisines de la wilaya d'Alger.

Dans un autre exemple où l'origine est la wilaya d'Oran vers la wilaya de destination Alger, le vecteur de décalage spatial $W_d y$ représente les flux moyens d'Oran vers Boumerdes, Blida, Tipaza, wilayas voisines de la wilaya d'Alger.

⁴ Corollaire 4.3.10 dans Horn et Johnson (1991, p.290)

IV-4 Les spécifications du modèle spatial pour les flux Origine-Destination

On propose le modèle général d'autorégression spatiale suivant, qui prend en compte les dépendances Origine-Destination.

$$y = \rho_o W_o y + \rho_d W_d y + \alpha I_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon \quad (4.13)$$

Ce modèle général, conduit à un nombre de modèle plus spécifiques. Nous recensons quatre modèles issus de la restriction du paramètre $\rho_i, i=1,2$.

1. la restriction $\rho_o = \rho_d = 0$, produit le modèle des moindres carrées ou il n'y a pas d'existence de dépendance autorégressive spatiale.

2. la restriction $\rho_d = 0$, résulte du modèle basé sur le poids de la matrice singulière W_o qui reflète la dépendance autorégressive spatiale de l'origine.

3. la restriction $\rho_o = 0$, produit un modèle basé sur le poids de la matrice singulière W_d qui reflète la dépendance autorégressive spatiale de destination.

4. En l'absence de restriction, ceci nous conduit au modèle (4.13) impliquant les deux matrices W_o, W_d .

IV-5 Estimation du modèle spatial des flux O-D

La vraisemblance est le point de démarrage pour la méthode du maximum de vraisemblance et de l'estimation Bayésienne. On note que la fonction de log-vraisemblance pour le modèle spécifié sachant les paramètres β et σ a la forme suivante :

$$\text{LnL}(\rho_1, \rho_2) = C + \ln |I_N - \rho_o W_o - \rho_d W_d| - \frac{N}{2} \ln(S(\rho_1, \rho_2)) \quad (4.14)$$

Où $S(\rho_1, \rho_2)$ représente le carré des erreurs qui s'exprime comme une fonction scalaire du paramètre $\rho_i, i=1,2$ seul, qui ne tient pas compte des paramètres β, σ et C ou C dénote une constante qui ne dépend pas de $\rho_i, i=1,2$ (voir Lesage et Pace, 2004).

Le log-déterminant de la matrice joue un rôle très important dans les deux méthodes, maximum de vraisemblance et l'estimation Bayésienne pour la transformation de la variable aléatoire.

Notamment, le log-déterminant assure que la transformation de la variable aléatoire a une densité continue. Autrement dit la multiplication de la variable dépendante par une

transformation comme εI , où ε est un nombre positif infinitésimal, devait réduire l'amplitude de l'estimation résiduelle à un niveau négligeable.

Le terme log-déterminant sert donc à se prévenir des transformations hasardeuses qui pourrait nuire à l'estimation.

Par conséquent, la vraisemblance est invariable pour de tels changements d'échelle.

L'algorithme standard pour le maximum de vraisemblance, l'estimation Bayésienne où la méthode des moments généralisés utilisée dans l'économétrie spatiale des modèles du flux interrégional de l'Origine-Destination devient difficile quand le nombre d'observation augmente.

Par exemple, l'utilisation de la matrice du flux interrégional de l'Origine-Destination pour un échantillon approximatif de 1500 communes, conduit à avoir une matrice de poids spatial de dimension n par n où $N=n^2 = 2.250.000$ observations.

Le maximum de vraisemblance et l'estimation Bayésienne exigent le calcul du log-déterminant pour la matrice N par $N(I_N - \rho_o W_o - \rho_d W_d)$.

Tandis qu'une approche plus spécialisée pour le calcul du log-déterminant pour des matrices très larges a été proposé par Pace et Lesage (2004), Barry et Pace (1999), ainsi que par Smirnov et Anselin (2001), une approche beaucoup plus efficace existe qui permet d'exploiter la structure spéciale de la matrice $W_d = I_n \otimes W$, et $W_o = W \otimes I_n$. Le paragraphe suivant examine une manière plausible pour l'estimation de la somme des carrés des erreurs, et le log-déterminant utilisé dans le calcul du log-vraisemblance.

IV-5-1 Estimation des paramètres et calcul des moments

Plus le nombre des origines et des destinations augmente, plus la difficulté de l'implémentation du modèle O-D augmente. Et en particulier, la création et le stockage des matrices X_o et X_d de dimension $(N \times k)$, peut créer lors du calcul un problème de mémoire pour un n grand, et le même problème se pose avec plus d'acuité pour les matrices creuses W_o , W_d de dimension $(N \times N)$.

On démontre comment la structure particulière du modèle O-D peut être exploitée pour permettre l'estimation et l'inférence des données plus larges.

Dans plusieurs cas, le calcul direct des quantités $X_o' X_o$ qui nécessitait $O(Nk^2)$ se réduit à $O(nk^2)$ grâce à l'exploitation de la structure unique du modèle. Ceci veut dire que le bénéfice que l'on tire de l'exploitation de structure spéciale augmente linéairement avec n . Et en plus une formulation efficace de l'estimation du problème permet à plusieurs calculs d'être effectués une seule fois, avec une subséquente mise à jour, le calcul de quantités comme la somme des carrés des erreurs demandera un temps de calcul très court.

Premièrement, l'implémentation du modèle O-D ne requiert plus la formation de la matrice $(N \times N)$, W_o , W_d . car $W_o y = (I \otimes W) \text{vec}(Y)$, donc $W_o y = \text{vec}(WY)$, en utilisant la

relation $(C' \otimes A) \text{vec}(B) = \text{vec}(ABC)$. D'une manière similaire, $W_d y = \text{vec}(YW')$, on réécrit le modèle (4.13).

$$Y - \rho_o W_o Y - \rho_d W_d Y = \alpha \iota_n \iota_n' + X \beta_d \iota_n' + \iota_n \beta_o X' + \gamma G + E \quad (4.15)$$

La combinaison du côté gauche de l'équation (4.15) est une combinaison linéaire de trois variables dépendantes, donc la combinaison des paramètres estimés est une combinaison linéaire de trois termes séparés, que on étiquette $\hat{\beta}^{(t)} = (Z'Z)^{-1}(Z'F^{(t)}(Y))$, où $F^{(t)}(Y)$ égale Y, WY, YW' , quand $t=1,2,3$. Donc, la combinaison du paramètre estimé devient :

$$\hat{\beta}(\tau) = [\hat{\beta}^{(1)} \quad \hat{\beta}^{(2)} \quad \hat{\beta}^{(3)}] \quad \text{où} \quad \tau(\rho) = [1 \quad -\rho_o \quad -\rho_d]'$$

La variable indépendante de la matrice des moments $Z'Z$ a la forme simple du modèle des moindres carrés traité en (4.4).

$$Z'F^{(t)}(Y) = [\iota_n' F^{(t)}(Y) \iota_n \quad X' F^{(t)}(Y) \iota_n \quad X' F^{(t)}(Y)' \iota_n \quad \iota_n' (G' \Theta F^{(t)}(Y)) \iota_n] \quad (4.16)$$

Quand $t=1$ l'expression (4.16) se réduit à (4.7) et l'estimation se réduit à celle du modèle de gravité. Soit $\hat{\beta}^{(t)}$ pour $t=1, 2, 3$, on forme alors la matrice des résidus $\hat{E}^{(t)}$ en substituant les paramètres estimés dans l'équation (4.17),

$$\hat{E}^{(t)} = F^{(t)}(Y) - \hat{\alpha}^{(t)} \iota_n \iota_n' + X \hat{\beta}_d^{(t)} \iota_n' + \iota_n (\hat{\beta}_o^{(t)})' X' + \hat{\gamma}^{(t)} G \quad (4.17)$$

Où la matrice des combinaisons résiduelles est :

$$\hat{E} = \hat{E}^{(1)} - \rho_o \hat{E}^{(2)} - \rho_d \hat{E}^{(3)} \quad (4.18)$$

On introduit la matrice du produit croisé des composants variables résiduels, Q , où $Q_{ij} = \text{tr}(\hat{E}^{(i)} \hat{E}^{(j)})$, $i=1,2,3$ et $j=1,2,3$, d'une manière similaire $Q_{ij} = \iota_n' (\hat{E}^{(i)} \Theta \hat{E}^{(j)}) \iota_n$ est la somme des carrés résiduel pour le modèle OD devient $S(\rho) = \tau(\rho)' Q \tau(\rho)$ par conséquent on recalcule $S(\rho)$ pour n'importe qu'elle valeur du vecteur $\tau(\rho)$, ceci requière un nombre d'itérations restreint qui ne dépend ni de n ni de k . Ceci permet l'optimisation rapide de la fonction de vraisemblance et l'accélération de l'estimation des algorithmes de Monte Carlo par Chaînes de Markov (MCMC) (voir II-3-2).

IV-5-2 Log-déterminant pour une matrice de poids

Dans les termes de l'équation (9), le terme $\ln |I_N - \rho W_s|$ qui prendra $O(N^2) = O(n^4)$ temps de calcul, mais si l'on tire profit de la structure même du modèle O-D et en utilisant des techniques efficaces pour évaluer le log-déterminant, on peut réduire le temps de calcul à $O(n)$. Le log-déterminant de la transformation $I_N - \rho W_s$ est la trace de la matrice du logarithme de la transformation, ce qui équivaut à une série de Taylor de forme très simple, ce que montre l'expression (4.19)

$$\ln |I_N - \rho W_s| = \text{tr}(\ln(I_N - \rho W_s)) = - \sum_{t=1}^{\infty} \frac{\rho^t \text{tr}(W_s^t)}{t} \quad (4.19)$$

Pour le cas d'une matrice de destination ou d'une matrice d'origine $W_d = I_n \otimes W$ ou $W_o = W \otimes I_n$, comme démontré dans les modèles (4.3), soit $W_s, s=o, d$, et on note :

$$\text{tr}(W_s^t) = \text{tr}(I_n^t \otimes W^t) = \text{tr}(I_n^t) \cdot \text{tr}(W^t) = n \cdot \text{tr}(W^t) \quad (4.20)$$

Le résultat de la trace d'une matrice carrée W d'ordre N est simplifié à une multiplication d'un scalaire (n) avec la trace d'une matrice carrée W d'ordre n .

$$\ln |I_N - \rho W_s| = -n \sum_{t=1}^{\infty} \frac{\rho^t \text{tr}(W_s^t)}{t} = n \ln |I_n - \rho W| \quad (4.21)$$

En résumé, pour le cas d'une matrice de poids $W_s, s=o, d$, les utilisateurs peuvent employer l'algorithme pour le calcul du log-déterminant de la matrice $(n \times n) \ln |I_n - \rho W|$, quand on travaille avec le vecteur de N du flux de l'Origine-Destination. Pour l'exemple précédent de $n=1500$ communes et $N=2.250.000$ on peut résoudre ces problèmes d'estimation en une poignée de secondes lorsqu'on utilise un calcul efficace de l'algorithme. (Voir Lesage et Pace, 2004).

IV-5-3 Log-déterminant pour le modèle du filtrage successive

Pour la spécification du modèle du filtrage spatial successif dans le modèle (4.4) de la transformation de la variable dépendante soit :

$$(I_N - \rho_o W_o) (I_N - \rho_d W_d) = (I_N - \rho_d W_d) (I_N - \rho_o W_o) \quad (4.22)$$

Le terme du log-déterminant apparaît dans la vraisemblance du modèle (4.4) et prend une expression simple dans (4.23), parce que le log-déterminant du produit est la somme des logarithmes des déterminants.

$$\ln |(I_N - \rho_o W_o)(I_N - \rho_d W_d)| = n \ln |(I_n - \rho_o W)| + n \ln |(I_n - \rho_d W)| \quad (4.23)$$

En résumé, le problème potentiel du calcul des flux Origine-Destination qui faisait intervenir ($N=n^2$) observations, pesé considérablement sur le temps de calcul, cet handicap a été levé car désormais le calcul du log-déterminant se réduit simplement au calcul d'une trace d'une matrice ($n \times n$).

IV-6 Application des outils de l'économétrie spatiale au modèle O-D

L'analyse que nous avons menée, nous permet de distinguer les facteurs explicatifs des mobilités des individus, car selon Berry et Ray (1987.57), l'une des conséquences de la migration interne est la modification de la composition par âge, par genre et du taux d'analphabétisme, des régions de départ et d'arrivée.

Pour les régions de départ cela conduit entre autres, à une érosion démographique (vieillesse de la population).

Pour les régions d'accueil, cela stimule la croissance démographique et la consommation, mais la migration peut aussi contribuer à la création de problèmes économiques et sociaux.

IV-6-1 Spécification du modèle économétrique spatiale O-D appliqué au flux migratoires intercentraux

Pour caractériser un mouvement migratoire on doit s'intéresser à deux aspects :

Le Premier : Quelles sont les variables explicatives qui poussent les migrants à sortir d'une wilaya (effet origine)

Le deuxième : Quelles sont les variables explicatives qui poussent les migrants à entrer dans une autre wilaya (effet destination).

Le modèle économétrique que nous proposons est le suivant :

$$y = \rho_o W_o y + \rho_d W_d y + \alpha_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon \quad (4.24)$$

y est la variable **endogène** qui contient les flux migratoires de dimension $(N \times 1)$ où n représente les 48 wilayas contiguës et $N = n^2 = 2304$.

Les flux migratoires ont été saisis du RGPH 1998, où il a été notamment demandé à toutes les personnes recensées en juin 1998 où elles résidaient à la date du précédent recensement 1987. Sur les 29 millions de personnes recensées en 1998, 940.000 d'entre elles soit 4.5% ont changé de wilaya de résidence depuis 1987 (migrants intercentraux), c'est sur cette base que le vecteur y est construit.

W est la matrice de poids spatiale qui capture les relations de contiguïté entre les wilayas, de dimension $(n \times n)$.

$W_o = I_n \otimes W$ de dimension $(N \times N)$, qui capture les relations entre les flux de wilaya d'origine vers les voisinages des wilayas de destination, où \otimes est le produit de Kronecker (voir annexe C.2).

$W_o y$ est la variable **endogène décalée** pour la matrice de poids W_o , car dans ce modèle, l'observation y est en partie expliquée, par les valeurs prises par y dans les régions voisines de l'origine.

$W_d = W \otimes I_n$ de dimension $(N \times N)$, qui capture les relations entre les flux de wilaya de destination vers les voisinages des wilayas d'origine.

$W_d y$ est la variable **endogène décalée** pour la matrice de poids W_d , car dans ce modèle, l'observation y est en partie expliquée, par les valeurs prises par y dans les régions voisines de la destination, voir (IV-3).

ρ_o et ρ_d sont les paramètres spatiaux autorégressifs indiquant l'intensité de l'interaction existant entre les observations de y de l'origine respectivement de la destination.

Le paramètre α (scalaire) dénote le terme constant, où $\mathbf{1}_N$ vecteur unitaire de dimension $(N \times 1)$.

Le paramètre γ (scalaire) reflète l'effet de la distance, $g = \text{vec}(G)$ où G est une matrice $(n \times n)$ qui représente la distance entre chaque point d'origine pour chaque point de destination, ainsi la dimension de g est $(N \times 1)$.

X est la variable explicative qui contient k caractéristiques pour chacune des n wilayas.

Les k caractéristiques retenues pour l'explication du phénomène de migration sont prises du RGPH 1987 où de très nombreux indicateurs sont concevables, nous en proposons plusieurs, pour lesquels les données sont disponibles tout en identifiant les caractéristiques qui influencent le comportement de migration, comme l'éducation, l'âge, le taux de chômage, et ceux qui ressemblent plus à des caractéristiques de lieux comme la superficie, le taux d'occupation des logements, le milieu urbain et rural etc. (voir annexe A1 pour le descriptif des variables explicatives).

$X_d = \mathbf{1}_n \otimes X$ où $\mathbf{1}_n$ est un vecteur unitaire $(n \times 1)$, X_d de dimension $(N \times k)$, représente les caractéristiques de destinations associées au vecteur des paramètres β_d de dimension $(k \times 1)$.

$X_o = X \otimes \mathbf{1}_n$ où $\mathbf{1}_n$ est un vecteur unitaire $(n \times 1)$, X_o de dimension $(N \times k)$, représente les caractéristiques d'origines associées au vecteur des paramètres β_o de dimension $(k \times 1)$.

Le vecteur $\varepsilon = \text{vec}(E)$ de dimension $(N \times 1)$ représente une perturbation.

C'est la présence des variables **endogènes décalées** qui rend l'utilisation des MCO biaisés et non convergents, d'autres techniques sont donc nécessaires pour l'estimation des paramètres du modèle (4.24), La plus utilisée de nos jours est la méthode d'estimation Bayésienne.

IV-6-2 Techniques d'estimations du modèle économétrique O-D

L'analyse statistique Bayésienne nécessite le calcul d'estimateurs issus de distributions de probabilité *a posteriori* généralement très complexes. Des techniques d'estimation Monte Carlo permettent alors de mettre en œuvre le calcul pratique de ces estimateurs. Le principe de base de ces méthodes est d'utiliser des séquences d'échantillons distribués suivant les lois *a posteriori* d'intérêt. Ces distributions sont souvent complexes et ne peuvent donc être simulées par des techniques classiques.

Les algorithmes de simulation par chaînes de Markov sont alors introduits. Ils permettent de simuler des distributions très générales et rendent possible le calcul des estimateurs issus des procédures d'estimations Bayésiennes. Les méthodes de simulation MCMC deviennent alors efficaces pour résoudre de nombreux problèmes d'estimation liés à la dépendance spatiale (voir chapitre II).

IV-6-3 Simulation du modèle O-D

L'échantillonnage par MCMC peut commencer par des valeurs initiales arbitraires, que l'on note δ^0 , $(\sigma^2)^0$, V^0 et ρ_o^0 , ρ_d^0 , on va procéder à un tirage séquentiel, tout en suivant les distributions *a posteriori* des paramètres (voir annexe B).

On emploie les paramètres mis à jour en remplacement des paramètres initiaux δ^0 , $(\sigma^2)^0$, V^0 et ρ_o^0 , ρ_d^0 . et ainsi de suite.

A chaque itération on collecte un paramètre tiré, qui sera utilisé pour la construction des distributions *a posteriori* pour les paramètres du modèle (voir annexe E).

Les programmes pour la modélisation des flux O-D ne sont pas disponibles dans les routines de J.LeSage, nous avons donc dû construire notre propre programme, que nous avons intitulé « spatial regression for O-D flows » programme qui comporte plus de 600 lignes.

Nous donnons ici un exemple de calcul de rhoO en langage MATLAB

```
function [rhoO,rhoD]=draw_rhoOD

nrhoO = length(detval(:,1));

bprior1 = beta_prior(detval(:,1),a1,a2);
den1 = den1 + log(bprior1);

% trapezoid rule1
isum1 = sum((y1(2:n1,1) + y1(1:n1-1,1)).*(x1(2:n1,1) - x1(1:n1-1,1))/2);
z1 = abs(x1/isum1);
den1 = cumsum(z1);
rnd1 = unif_rnd(1,0,1)*sum(z1);
ind1 = find(den1 <= rnd1);
idraw1 = max(ind1);
if (idraw1 > 0 & idraw1 < nrhoO)
rhoO = detval(idraw1,1);
end;
```

Dans le but de faire dérouler notre programme en utilisant l'algorithme MCMC, on a procédé à 121 000 itérations, tout en prenant comme précaution de rejeter les 1 000 premières itérations, afin de lever le biais concernant l'introduction des valeurs initiales arbitraires. La stabilisation des résultats s'est faite à partir de 100 000 itérations.

Comme matrice de contiguïté W nous utilisons la matrice d'ordre 3, car c'est à partir de cet ordre que les résultats deviennent convergents.

IV-6-4 interprétation des résultats du modèle O-D

Des études antérieures menées par (Greenwood 1985,527 ; Maynard et al .1997 ; Long et Cadwallader, 1992) ont montrées que les variables les plus pertinentes étaient entre autres: l'éducation, l'âge, le chômage, les compétences professionnelles, mais qu'on est-il du cas de l'Algérie? Nous avons construit un modèle de flux migratoire pour tenter d'identifier les facteurs qui expliquent ces flux.

Les résultats obtenus sont disponibles en annexe (A2), nous en tirant les principaux résultats, tout en faisant le descriptif des variables significatives.

IV-6-4-1 L'emploi

La répartition des flux migratoires sur une région donnée, n'est pas le fruit du hasard. En effet, certaines wilayas d'accueil peuvent être plus attrayantes que d'autres pour certaines catégories de personnes, on suppose que les individus recherchent la wilaya de résidence qui maximise leur bien-être. Ils évaluent les différentes alternatives et choisissent l'option qui leurs procure le plus de satisfaction sous la contrainte de leurs ressources humaines et financières, Ils décident alors de migrer ou de rester.

De manière analogue pour l'emploi, les individus évaluent les offres d'emplois que leurs proposent les entreprises (ou les wilayas) "en concurrence pure et parfaite", et se répartissent de manière non aléatoire (en connaissance de cause) à travers les emplois (ou les wilayas) disponibles. En conséquence, ces changements auront également une influence majeure sur la taille et la composition des flux migratoires.

Nous nous intéressant qu'aux variables significatives du modèle⁵ :

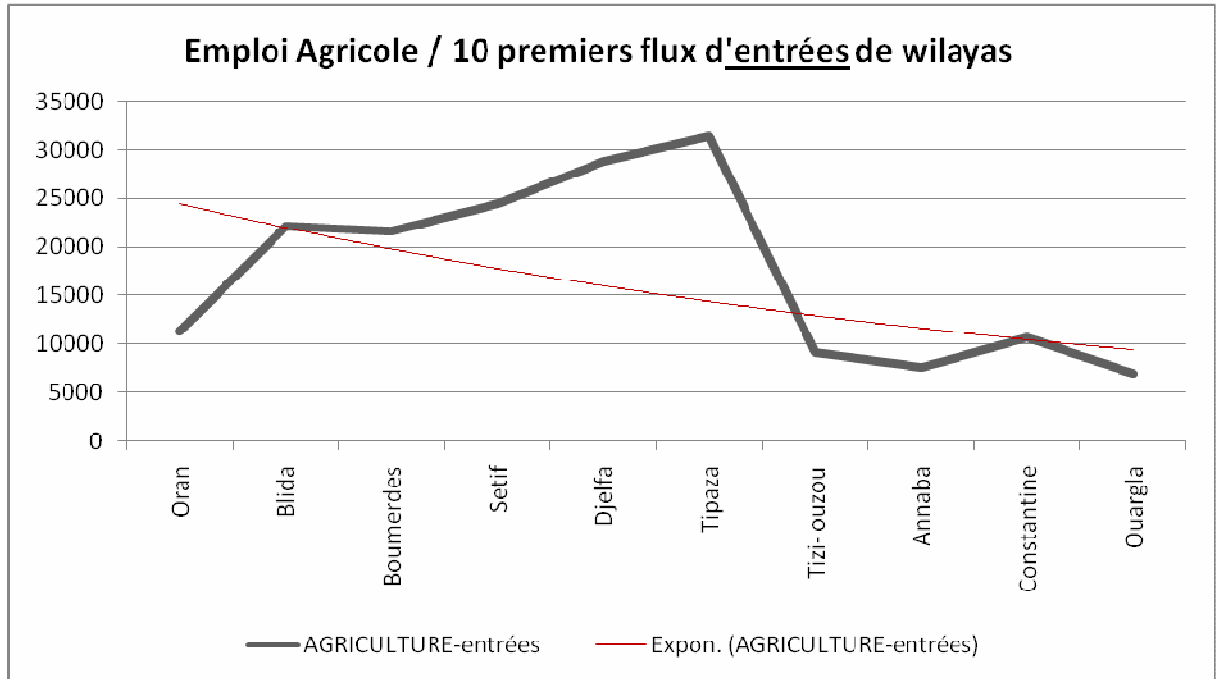


Figure 4.1

La courbe ci-dessus à une pente négative, ce qui est confirmé par les résultats du modèle (annexe A2) avec un coefficient de -0.409046, ce qui indique qu'une wilaya avec un fort taux d'emplois agricole n'attire pas les flux de migrants.

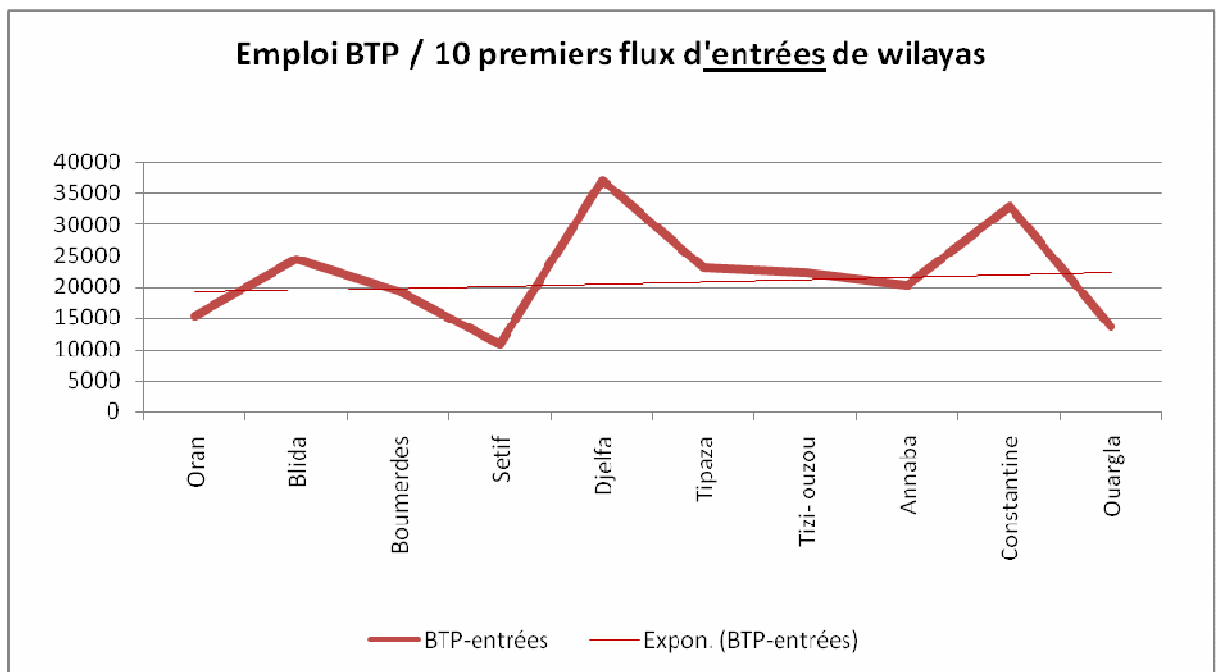


Figure 4.2

⁵ La wilaya d'Alger occupe la première place des flux d'entrées et de sorties, mais les valeurs extrêmes qu'elle prend rendent l'interprétation des résultats graphiques peu efficace; elle n'est donc pas perceptible dans les graphes proposés.

La courbe ci-dessus à une pente positive ,ce qui est confirmé par les resultats du modèle (annexe A2) avec un coefficient de 0.308612, ce qui indique qu'une wilaya avec un fort taux de BTP attire plus de flux de migrants.

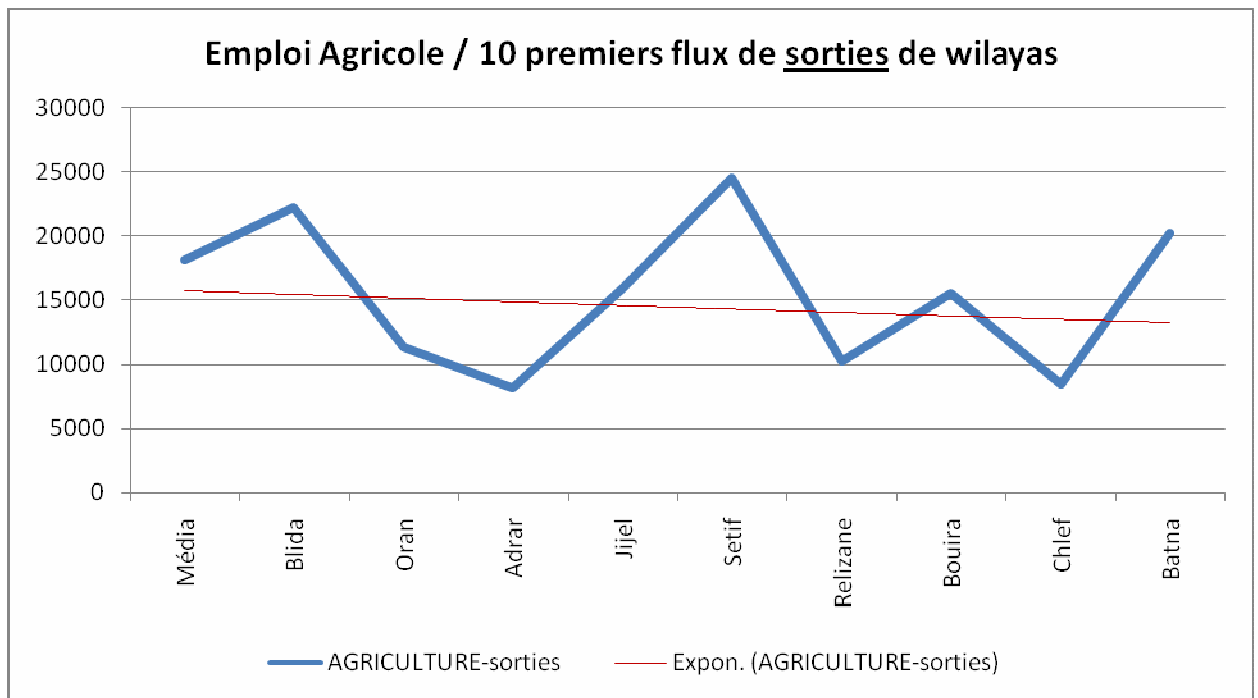


Figure 4.3

La courbe ci-dessus à une pente négative, ce qui est confirmé par les résultats du modèle (annexe A2) avec un coefficient de -0.419790, ce qui indique que les agriculteurs ne quittent pas leurs terres, ils y sont attachés.

IV-6-4-2 Le taux de chômage

En Algérie et selon l'ONS le taux de chômage en 1987 au niveau national était de 21,49%, on enregistre au niveau de la wilaya de Ghardaïa le plus bas taux avec 12,49%, le plus haut taux est enregistré à Tébessa avec 33,24%.

La variable, taux de chômage est le rapport entre la population en chômage sur la population active.

La population active comprend l'ensemble des individus qui ont un emploi rémunéré ou qui en cherche un.

La population en chômage comprend l'ensemble des individus, aptes à travailler, ayant déclaré :

- (i) être en âge d'activité (15 –59 ans).
- (ii) être sans travail (aucune activité économique marchande).
- (iii) être à la recherche d'un emploi.

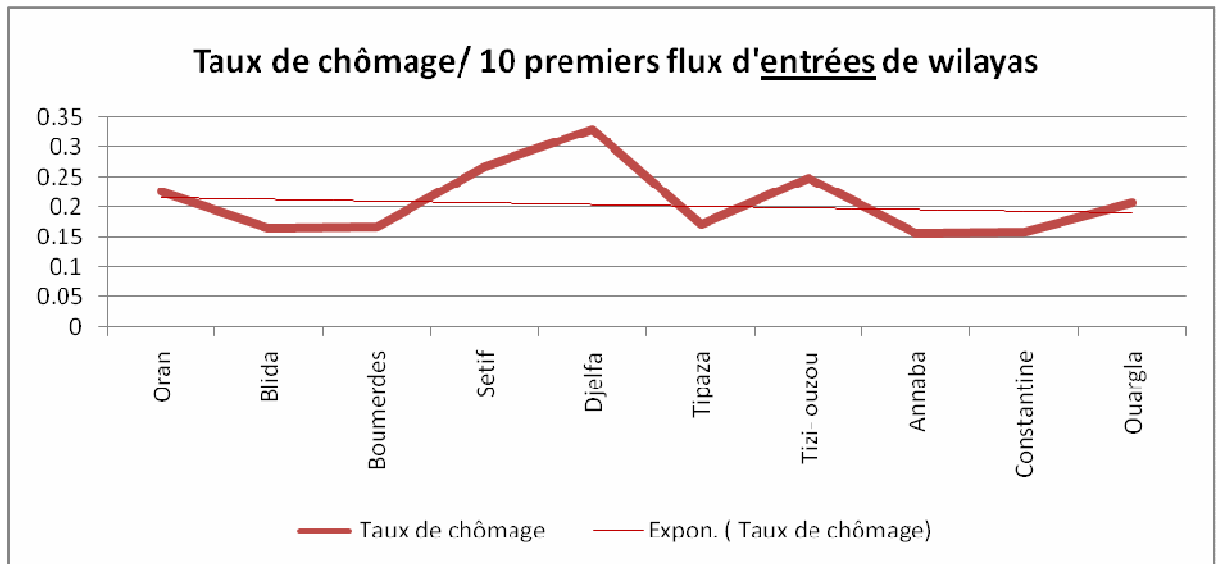


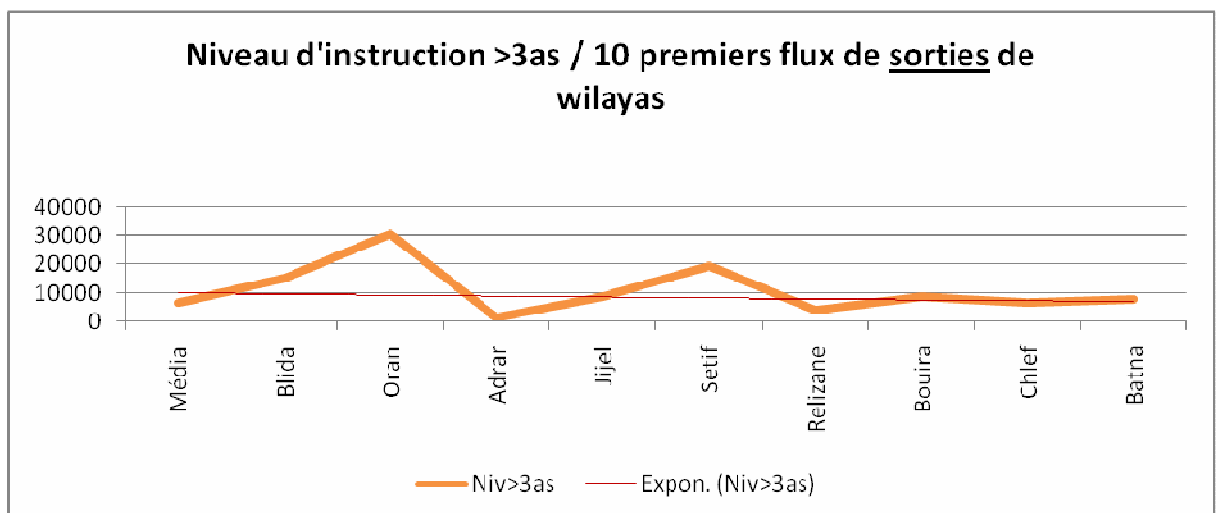
Figure 4.4

La courbe ci-dessus présente une pente négative, ce qui est confirmé par les résultats du modèle avec un coefficient de -0.137255 , ce qui indique que la migration ne se fait pas au profit des wilayas ayant un fort taux de chômage.

IV-6-4-3 Niveau d'instruction supérieur à la 3AS

Une autre régularité empirique semble être que ceux qui ont un niveau plus élevé de scolarité sont plus susceptibles de migrer, parce qu'ils reçoivent de plus grands profits pour leurs compétences lors de leurs migrations.

La migration est associée à un certain degré de "l'élitisme". Des études antérieures montrent que l'éducation a un effet sur la probabilité de migrer (Bogue 1985, Greenwood 1975), mais l'effet n'est pas nécessairement significatif dans plusieurs modèles de régression comme pour le cas de l'Algérie ; car la courbe ci-dessous a une pente négative, ce qui est confirmé par les résultats du modèle avec un coefficient de -0.287798 , ce qui indique qu'en Algérie, il n'y a pas de migration d'élite.



IV-6-4-4 Population urbaine

La population urbaine en Algérie n'a cessé d'augmenter par rapport à la population rurale. De 13,9 % en 1886, la population urbaine a atteint une moyenne de 58.30% en 1998. L'accroissement des villes en Algérie attire une population de plus en plus importante, la répartition géographique des immigrants se fait en général au profit de certains centres urbains.

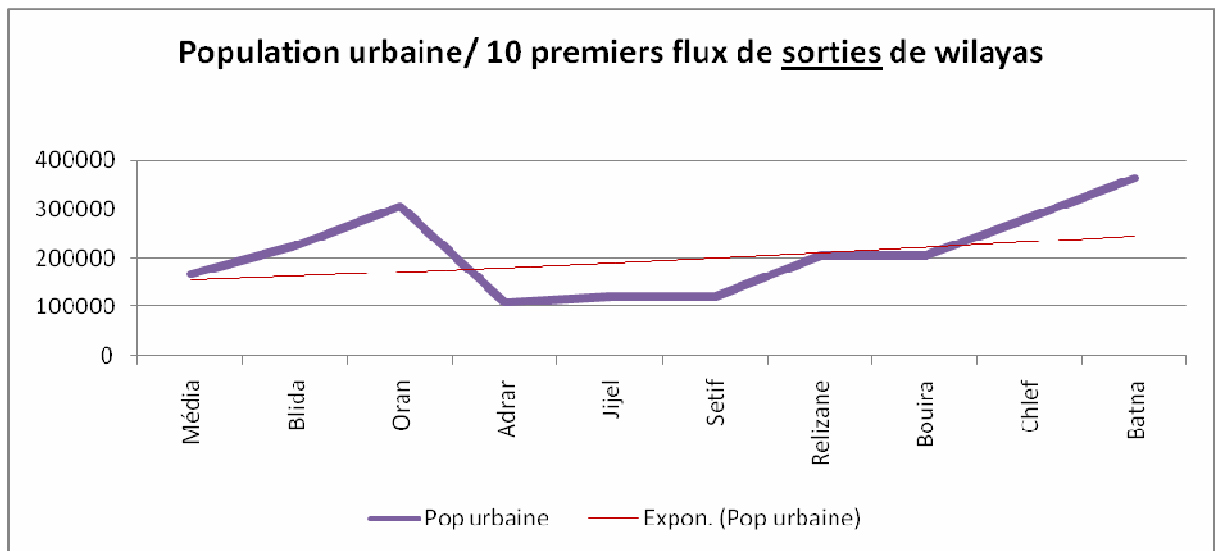


Figure 4.6

La courbe ci-dessus à une pente positive, ce qui est confirmé par les résultats du modèle avec un coefficient de 0.570663, ce qui indique que les urbains font augmenter le flux de migrants de departs.

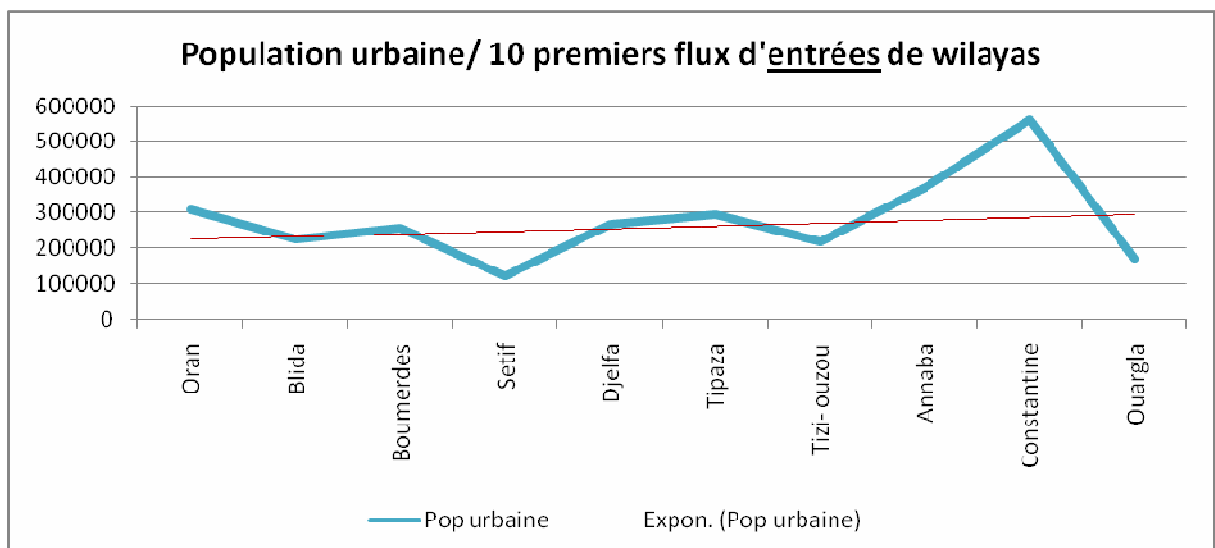


Figure 4.7

La courbe ci-dessus à une pente positive, ce qui est confirmé par les résultats du modèle avec un coefficient de 0.172841, ce qui indique que l'immigration ce se fait au profit des wilayas les plus urbaines.

IV-6-4-5 Population née dans sa wilaya de résidence

Une autre façon d'identifier les caractéristiques des migrants et l'attrait relatif d'un lieu de migration comme un objectif et de se pencher sur la proportion de ses résidents nés dans leurs wilayas de résidence.

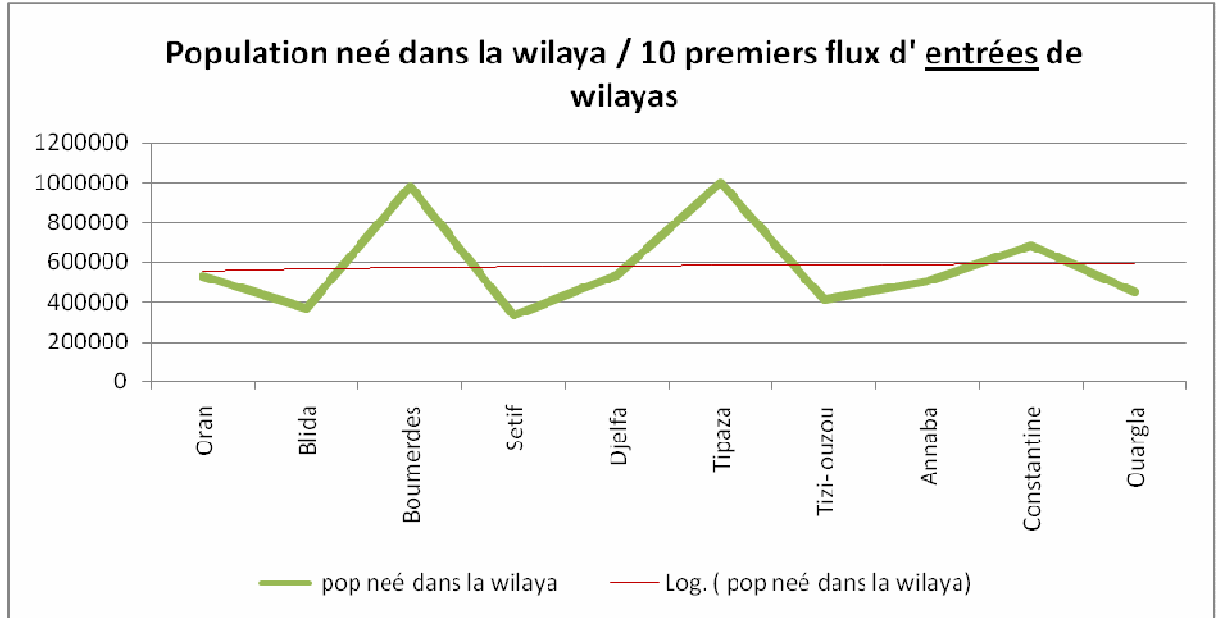


Figure 4.8

La courbe ci-dessus à une pente positive, ce qui est confirmé par les résultats du modèle avec un coefficient de 0.588823, ce qui indique qu'effectivement il y'a une migration de retour, comme l'indique le gaphe suivant :

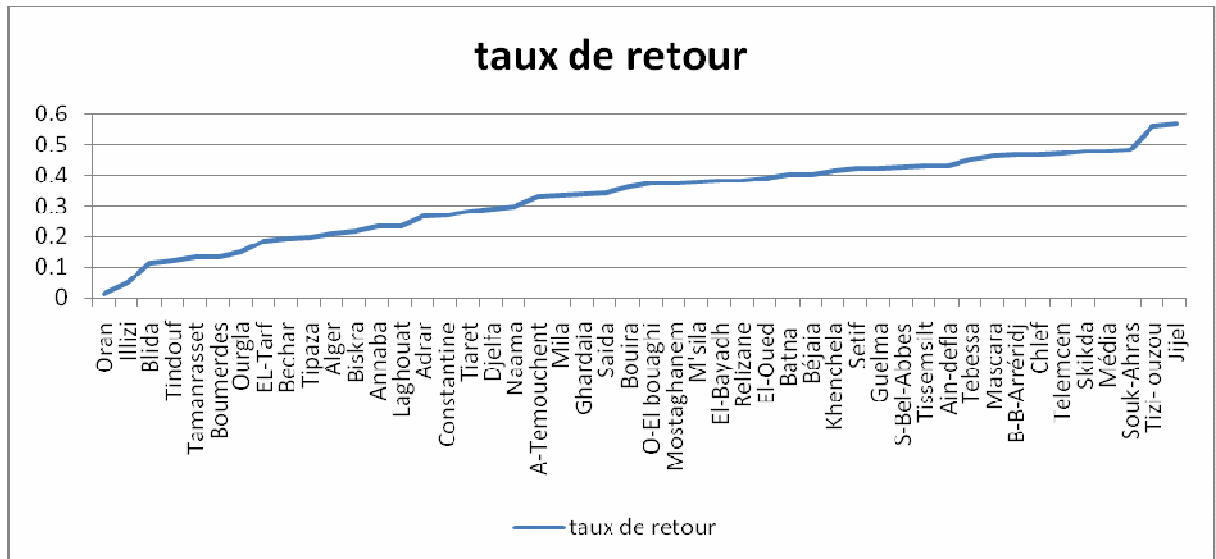


Figure 4.9

Le graphe ci-dessus a été construit sur la base du recensement de 1998, il a été demandé aux personnes entrant dans une wilaya entre 1987 et 1998 leurs lieux de naissance, pour le cas d'Alger par exemple :

Sur les 148 111 personnes qui sont entrées dans la wilaya d'Alger entre 1987 et 1998, parmi elles 31 096 sont nées à Alger.

Le graphe ci-dessus, nous renseigne sur les wilayas les plus et les moins attractives, car lorsque le taux dépasse les 0.5%, comme pour Tizi-Ouzou et Jijel cela signifie que ces wilayas ne présentent de l'intérêt que pour ceux qui y sont nés, par contre lorsque le taux est faible, comme pour Oran et Blida et Alger, cela signifie que l'on migre vers ces wilayas, même si l'on n'y est pas originaire, c'est donc des wilayas attractives.

IV-6-4-6 La densité de population

La densité de population moyenne de l'Algérie au recensement de 1987 était de 9,6 hab. /km², mais cette moyenne cache de grande disparité, à cause de la composition spatiale du pays qui est entièrement polarisée sur le Nord et en particulier dans la bande littorale, où la densité de cette dernière est de 222,5 hab. /km² alors qu'au sud elle est de l'ordre de 0,9 hab. /km².

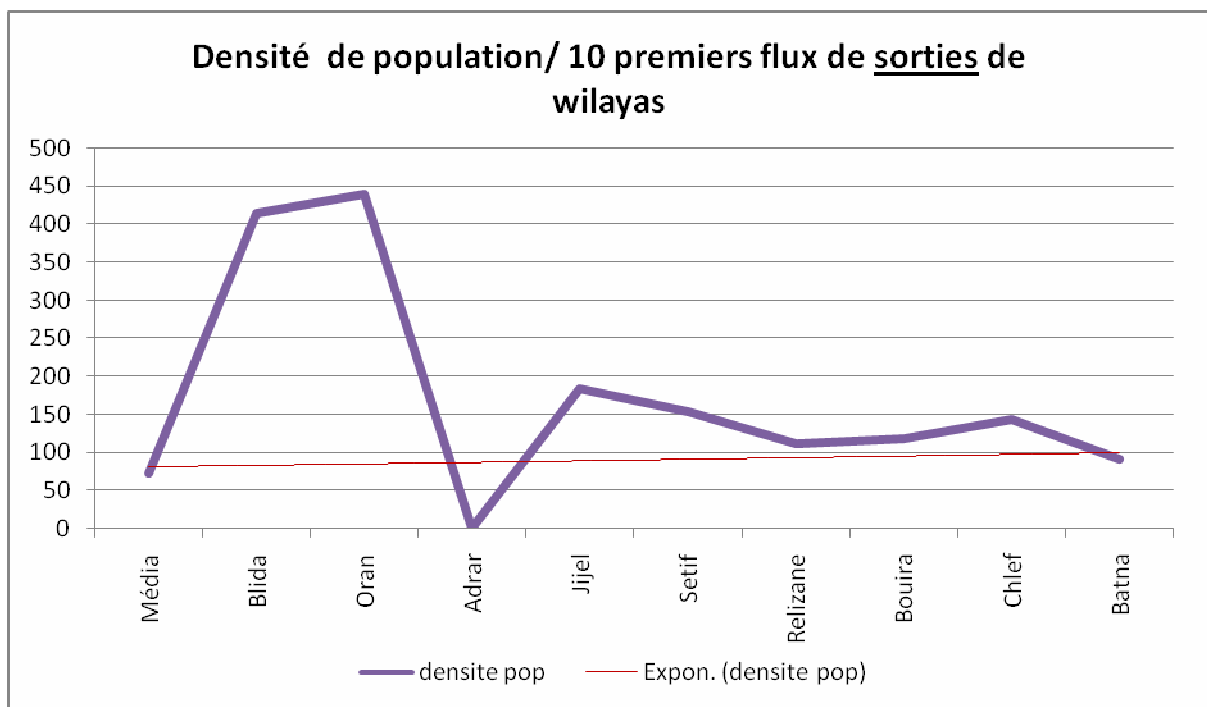


Figure 4.10

la courbe ci-dessus à une pente positive, ce qui est confirmé par les résultats du modèle avec un coefficient de 2.963496, ce qui indique que plus une wilaya a une forte densité de population et plus les flux de sortie augmentent.

IV-6-5 Interprétation globale des résultats du modèle O-D

L'application des flux migratoires entre les 48 wilayas contiguës, en intégrant l'autocorrélation spatiale, nous a prouvé l'existence des flux origine et destination, $\rho_o=0.083742$ et $\rho_d=0.130490$, ces valeurs sont significatives et relativement proches, ce qui indique une dépendance spatiale d'égale importance entre les régions voisines de l'origine et les régions voisines de la destination.

L'emploi agricole fait diminuer le flux de l'origine avec un coefficient de -0.419790 et le flux vers la destination avec un coefficient -0.409046 , ce qui indique que les agriculteurs ne quittent pas leurs terres, ils y sont attachés et cela veut dire aussi que l'on ne migre pas vers les wilayas à fort taux d'emploi agricole.

La population urbaine fait augmenter le flux de l'origine avec un coefficient 0.570663 et le flux vers la destination avec un coefficient de 0.172841 , ce qui indique que les urbains sont les plus favorables à migrer et cela veut dire aussi, que l'on migre vers les wilayas à forte proportion urbaine.

Le niveau d'instruction supérieur fait diminuer le flux de l'origine, avec un coefficient de -0.287798 , ce qui indique qu'en Algérie, il n'y a pas de migration d'élite.

La densité de population fait augmenter le flux de l'origine, avec un coefficient de 2.963496 , ce qui indique que plus une wilaya a une forte densité de population et plus les flux de sortie augmentent.

La superficie de la wilaya fait augmenter le flux de l'origine, avec un coefficient de 2.084550 , ce qui indique que plus une wilaya a une grande superficie et plus les flux de sortie augmentent.

Le taux de chômage fait diminuer le flux vers la destination, avec un coefficient de -0.137255 , indiquant que la migration ne se fait pas au profit des wilayas ayant un fort taux de chômage.

L'emploi dans le BTP fait augmenter le flux vers la destination, avec un coefficient de 0.308612 , ce qui indique qu'une wilaya avec un fort taux d'emploi dans le BTP attire plus de flux de migrants.

La variable population née dans la wilaya de résidence, fait augmenter le flux vers la destination, avec un coefficient de 0.588823 , ce qui indique que la migration permet le retour vers sa wilaya de naissance.

La distance (gamma) est positive, avec un coefficient de 0.529088, ce qui indique qu'en Algérie, ce n'est pas la proximité qui fait migrer la population, mais c'est la particularité du lieu lui-même, en se déplace par exemple d'une wilaya de l'Est vers une wilaya du centre, même si la distance qui sépare les deux wilayas est grande.

Deux variables, superficie et densité ont changé de catégorie, elles sont passées de non significatives par régression spatiale à significatives par MCO (voir annexe A2). On remarque de se fait, que la régression par MCO ne donne pas de bons résultats, car aucune des variables socio- démographiques ni socio-économiques n'est significative.

Conclusion

Nous soutenons donc que l'emploi de la régression avec la méthode des MCO, pour estimer les modèles à interaction spatiale, néglige l'éventuelle dépendance spatiale dans l'échantillon du vecteur y des flux Origine-Destination, on sait, de plus que l'utilisation des MCO en présence de dépendance spatiale dans la variable dépendante y , conduit à biaiser les estimations (voir LeSage and Pace, 2004).

Pour aborder cette question, nous avons vu comment incorporer la dépendance spatiale dans le modèle de gravité, tout en considérant un filtre spécifique pour le modèle spatial Origine-Destination : $(I_N - \rho_o W_o - \rho_d W_d)$, qui est appliqué au vecteur y des flux O-D pour capturer les deux dépendances spatiales qui peuvent survenir entre les flux Origine-Destination : Origin-based (W_o), Destination-based (W_d).

L'utilisation de ce filtre, a rendu le travail actuel sur les modèles spatiaux pertinent pour modéliser les flux origine-destination et cela malgré le nombre important d'observations dans l'échantillon ($N=n^2$).

L'approche de J.LeSage a apporté de nouveaux résultats pour rendre faisable l'estimation Bayésienne et l'estimation du maximum de vraisemblance pour des données de flux O-D. Ceux-ci exploitent la structure naturelle de la famille des modèles gravitaires, tout en appliquant la régression spatiale introduite ici, ce qui a permis d'avoir une inférence et une estimation des flux O-D.

Notre application basée sur les flux migratoires entre les 48 wilayas, a démontré l'existence des flux origine-based et destination-based, invalidant de ce fait même l'utilisation des estimations des moindres carrés. Elle a démontré également, qu'en Algérie il n'y avait pas de migration d'élite et que la migration se faisait au profit des wilayas les plus urbaines et à fort taux d'emploi dans le BTP et ne se faisait pas au profit de wilayas à fort taux d'emploi agricole et à fort taux de chômage.

Conclusion générale

Conclusion

L'espace dans la plus grande partie du XIX^e siècle et jusqu'au milieu de notre siècle a été marginalisé, mais les causes ne sont pas totalement claires. On peut penser que l'espace s'est éloigné de la scène quand les économistes ont commencé à vouloir appréhender de façon plus systématique et formelle l'ensemble des interdépendances économiques, en reculant les limites du non démontré, en recherchant à endogénéiser la plupart des phénomènes.

Les modèles théoriques en économie spatiale (et notamment en économie géographique) exigent également l'introduction de nouvelles méthodes d'estimation capables de tenir compte de phénomènes d'interdépendances beaucoup plus complexes : effets asymétriques, variation des effets spatiaux entre régions, incertitude concernant les mesures de distance ou de proximité.

L'approche que nous avons présentée a permis d'estimer les modèles spatiaux tout en tenant compte à la fois des problèmes d'hétérogénéité, des observations extrêmes, et en modélisant les interactions spatiales à partir de fondements économiques.

Nous avons exposé de nouveaux résultats pour rendre faisable l'estimation Bayésienne et l'estimation du maximum de vraisemblance pour des données de flux O-D. Ceux-ci exploitent la structure naturelle de la famille des modèles gravitaires, tout en appliquant la régression spatiale introduite ici, ce qui nous a permis d'avoir une inférence et une estimation des flux O-D.

L'application des flux migratoires entre les 48 wilayas, a démontré l'existence des flux origine-based et destination-based, invalidant de ce fait même, l'utilisation des estimations des moindres carrés. On a démontré également, qu'en Algérie il n'y avait pas de migration d'élite et que la migration se faisait au profit des wilayas les plus urbaines et à fort taux d'emploi dans le BTP et ne se faisait pas au profit de wilayas à fort taux d'emploi agricole et à fort taux de chômage.

Suite aux recherches présentées dans ce mémoire, une des voies d'exploration possibles consistera à évaluer un troisième flux qui est Origine-Destination based, qui mesurera le flux moyen des wilayas d'origine vers les wilayas voisines des wilayas de destination, ce qui permettra une meilleure estimation du modèle.

De même, il paraît utile de travailler sur un autre découpage géographique que celui de la wilaya, ensemble trop hétérogène qui masque d'importants déséquilibres, un découpage par commune devrait donc remédier à cela.

Nous apportons une dernière réserve quant à l'exploitation du flux migratoire, en fait , les données d'un recensement portent sur des effectifs de migrants (individus) et ne peuvent en aucun cas, nous renseigner sur l'ensemble des migrations (événements) faites par les personnes, de plus le recensement ne permet pas l'introduction de quelques variables pertinentes telle que : les salaires, le loyer, variables qui exercent une influence sur les décisions de migration, le recours aux enquêtes nationales peut remédier à cela.

Bibliographie

Bibliographie

- [01]Anselin L (2004) “Advances in Spatial Econometrics” Springer Berlin Heidelberg New York ISBN 3-540-43729-0.
- [02] Anselin, L (2002) “Spatial externalities, Spatial multipliers and spatial econometrics” W.P Regional economics applications laboratory.
- [03]Bendjelid A, 1982 “ *Planification et organisation de l'espace en Algérie*”, Alger, Office des Publications Universitaires, diffusion Sindbad (Paris), 135p.
- [04]Cote M., 1988 “ *L'Algérie ou l'espace retourné* ”, Paris, Flammarion, 362p.
- [05]Cote M., 1983 “ *L'Espace algérien, les prémices d'un aménagement* ”, Alger, Office des Publications Universitaires, 278p.
- [06]Derycke Pierre-Henri, Huriot Jean-Marie “Microéconomie et espace: quelle intégration ?” *Revue économique*, Vol. 47, No. 2, Microéconomie et espace (Mar., 1996), pp. 187-192.
- [07]Droesbeke J-J(2006) “Analyse spatiale des données spatiales”edition Technip,paris ISBN 978-2-7108-873-2
- [08]Grasland C.,2004 , “Les inégalités régionales dans une Europe élargie,L'Harmattan (collection « Pays de l'Est »),pp181-214.
- [09]Geweke, J (1998) “Using simulation methods for Bayesian econometrics models: inference, development and communication” W.P University of Minnesota and federal reserve bank of Minneapolis.
- [10] Geweke, J (1993) “Bayesian treatment of the independent Student t Linear model” *Journal of applied econometrics*, Vol, 8 pp S19-S40.
- [11].Hemaoun R “L'Algérie, histoire et société et culture” ED Casbah 2000,p 75.
- [12]Koop G (2003) “Bayesian Econometrics” John Wiley & Sons, Ltd, West Sussex, England.
- [13]Le Gallo J (2002), “ *Disparités géographique et convergence des régions européens : une approche par l'économétrie spatiale* ” *Thèse de Doctorat en Analyse et Politique Économiques*
- [14]Le Gallo J. (2001a) “ Econométrie spatiale, 2. Hétérogénéité spatiale”, *Document de travail du LATEC*, Dijon, n°2001-01 (Janvier 2001).
- [15]Le Gallo J., Ertur C., Baumont C. (2002) “A spatial econometric analysis of convergence across European regions”, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin,
- [16] LeSage J(2005) “ Bayesian estimation of spatial regression models” lecture note.
- [17] LeSage J and Pace K (2006) “Spatial Econometric modelling of Origin-Destination Flows”, SSRN id 924609.
- [18] LeSage, J.P. et O. Parent (2004). “Bayesian Model Averaging for Spatial Econometric Models” ,working paper.
- [19] LeSage, J.P. (2004). “Spatial Regression Models”, *Numerical Issues in Statistical Computing for the Social Scientist*, John Wiley & Sons,Inc., Micah Altman, Jeÿ Gill and Michael McDonald (eds.), pp. 199-218.

[20]LeSage J.P. (1997) “ Bayesian estimation of spatial autoregressive models”, *International Regional Science Review*, 20, 113-129.

[21]LeSage J.P. (1999a) “ *Spatial Econometrics*”, WebBook of Regional Science.

[22]LeSage J.P. (1999b) “ Bayesian estimation of limited dependent variable spatial autoregressive models”, *Geographical Analysis*, 32, 19-35.

[23]Pace, K.R. et J.P. LeSage (2002). “*Spatial Autoregressive Local Estimation*, Recent Advances in Spatial Econometrics”, Jesus Mur, Henri Zoller and Arthur Getis (eds.), Palgrave Publishers.

[24]Pace, K.R. et J.P. LeSage (2003). “ Chebyshev Approximation of Log-determinants of Spatial Weight Matrices, Computational Statistics and Data Analysis”,

[25]Pace, K.R. et J.P. LeSage (2004). “ Spatial Statistics and Real Estate”, The journal of Real Estate Finance and Economics, 29, 147-148.

[26] Parent O (2005) “Analyse Bayésienne des externalités de connaissance : Étude sur données européennes ”, Thèse de Doctorat en Sciences Économiques.

[27] Poirier, D.J. (1995). *Intermediate Statistics and Econometrics*, Cambridge, The MIT Press.

[28] Tobler W. (1979) Cellular geography, dans : Gale S., Olsson G. (eds.), *Philosophy in Geography*, Reidel, Dordrecht.

[29] Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley & Sons. NY.

Collections:

[30]ONS, 2002 communication de Mokhtari Aziz.À l’occasion de la conférence des statisticiens européens.

[31]ONS, 2000 “*Armature urbaine, au RGPH 1998*, collections statistiques n°97, Alger, ONS.

[32]ONS, 1998 “ *Données statistiques n° 315, Migrations internes inter-wilaya RGPH 1998*, Alger, ONS.

[33]ONS, 1991 “ *Evolution des Populations et des Limites Communales 1977-1987*, Alger, ONS.

[34]ONS, Mouafeg Jaennine (1984), “les migrations internes entre les wilayate de 1966 à 1977”.

Annexes

Annexe A

Annexe A1 : Description des variables explicatives

Nom des Variables	Description
Emploi agricoles	Proportion de l'emploi agricole
Emploi dans le BTP	Proportion de l'emploi dans le BTP
Emploi dans l'industrie	Proportion de l'emploi dans l'industrie
Emploi dans l'administration	Proportion de l'emploi dans l'administration
Emploi dans les hydrocarbures	Proportion de l'emploi dans les hydrocarbures
Emploi dans le commerce	Proportion de l'emploi dans le commerce
Emploi dans le transport	Proportion de l'emploi dans le transport
Population entre 25-59ans	Proportion de la population âgée entre 25-59ans
Population de plus de 60ans	Proportion de la population âgée de plus de 60ans
Niveau d'instruction > 3AS	Niveau > 3AS / population >25ans
Niveau d'instruction < 3AS	Niveau < 3AS / population >25ans
Urbain	Proportion de la population urbaine
Rural	Proportion de la population rurale
Superficie	Superficie de la wilaya
Population née dans la wilaya de résidence	Proportion de la population née dans la wilaya de résidence
Densité	Densité de la population
Taux de chômage	Taux chômage
Taux d'occupation des log	Taux d'occupation des logements
Distance	Distance entre wilaya d'origine et wilaya de destination (à vol d'oiseau)

Annexe A 2 : Estimation du modèle spatial et des MCO

nomit=1 000 ndraw=121 000 W=slag(w, 3)

Variable	Modèle spatial			MCO		
	Coefficient	t-statistique	(plevel)	Coefficient	t-statistique	(plevel)
cst	0.067159	9.249817	0.000000**	4.684454	3.520458	0.000431**
O_agriculture	-0.419790	-2.426077	0.015263 **	0.285609	0.943109	0.459206
O_BTP	0.135827	0.576246	0.564449	0.103966	0.078132	0.937723
O_Industrie	0.097258	0.570729	0.568183	-0.009304	-0.006992	0.994421
O_Administration	0.110893	0.322662	0.746951	0.258569	0.194319	0.845926
O_Hydrocarbures	0.135056	1.356145	0.175053	-0.218456	0.164174	0.869594
O_Commerce	0.026817	0.067775	0.945965	0.563080	0.423166	0.672174
O_Transport	0.050563	1.340686	0.180022	0.107192	0.080557	0.935795
O-Age25_60	0.129607	0.688733	0.490991	1.497107	1.125105	0.260545
O-Age_plus_60ans	0.047454	0.452840	0.650664	-0.220985	-0.166074	0.868099
O-inf3As	-0.037392	-0.212709	0.831554	0.454355	0.341457	0.732760
O-niv>3As	-0.287798	-2.258820	0.023895 **	-0.479109	-0.360060	0.718802
O-Urbain	0.570663	3.381201	0.000722 **	-0.270958	-0.203630	0.838643
O-Rural	-0.003686	-0.015434	0.987686	-0.764452	-0.574500	0.565629
O_Superficie	2.084550	2.518708	0.011779 **	1.745646	1.311887	0.189558
O_Né_wil_origine	-0.020452	-0.543400	0.586854	1.546348	1.162110	0.245191
O_Densité	2.963496	2.455730	0.014060 **	-0.011869	-0.008920	0.992883
O_Taux_chômage	-0.006327	-0.070502	0.943794	-0.059708	-0.044871	0.964210
O_Taux_occupation_logements	-0.106845	-3.846386	0.120123	0.217575	0.163511	0.870116
D_agriculture	-0.409046	-4.124838	0.000037 **	-0.009750	-0.007328	0.994153
D_BTP	0.308612	2.191225	0.028436 **	0.195058	0.146590	0.883456
D_Industrie	0.056403	0.547624	0.583950	0.220157	0.165452	0.868588
D_Administration	-0.095615	-0.511082	0.609293	0.268172	0.201536	0.840279
D_Hydrocarbures	0.063620	1.198941	0.230551	3.069422	2.306730	0.210670
D_Commerce	-0.108714	-0.514995	0.606557	-1.522362	-1.144085	0.252589
D_Transport	-0.399957	-2.080132	0.307513	3.972671	2.985539	0.283451
D-Age25_60	-0.054005	-0.533816	0.593469	0.268356	0.201675	0.840171
D-Age_plus_60ans	-0.005577	-0.099462	0.920771	-0.055211	-0.041492	0.966904
D-inf3As	-0.293151	-3.062482	0.219558	-0.186196	-0.139930	0.888715
D-niv>3As	-0.223642	-3.322739	0.891123	-0.291103	-0.218769	0.826830
D-Urbain	0.172841	1.956131	0.050450 *	0.711722	0.534873	0.592738
D-Rural	0.085363	0.645075	0.518879	0.853603	0.450765	0.188759
D_Superficie	0.030891	1.714724	0.86396 1	0.388981	1.147274	0.0396 61**
D_Né_wil_origine	0.588823	2.833380	0.004606 **	0.888253	2.333860	0.048066
D_Densité	0.093011	1.288391	0.197610	0.331190	1.883921	0.049760 **
D_Taux_chomage	-0.137255	-2.553347	0.010669 **	0.372515	1.533407	0.106769
D_Taux_occupation_logements	-0.064749	-2.334634	0.19563 1	-0.647549	-3.334634	0.95683 1
gamma	0.529088	9.858921	0.000000 **	0.290858	7.589421	0.000000 **
rhoO	0.083742	1.690021	0.091024 *			
rhoD	0.130490	2.488000	0.012846 **			

*Significatif à 10%

**Significatif à 5%

Annexe B:

Distribution conditionnel pour le modèle SAR hétéroscédastique.

On présente la distribution conditionnel pour les paramètres δ , σ , ρ_o , ρ_d et $V = \text{diag}(v_{ij})$, $i=1, \dots, n$, ou $\delta = [\alpha \ \beta_0 \ \beta_d \ \gamma]'$.

La présentation est pour un modèle général où la matrice diagonale V d'ordre N contient les variances (scalaire) des paramètres qui distingue ce modèle du modèle du maximum de vraisemblance, soit $Z = [i_N \ X_0 \ X_d \ g]'$.

On présente les distributions conditionnelles des paramètres δ , et σ^2 lorsque la loi a priori non informative est appliquée au paramètre δ et une loi indépendante a priori gamma (v_1, v_2) est appliquée à σ^2 .

On assigne une loi a priori uniforme pour le paramètre ρ_i où $-1 < \rho_i < 1$ les conditions de stabilité sont i) $\sum_i \rho_i > -1$, $\sum_i \rho_i < 1$, $i=1, 2$ chaque $|\rho_i| < 1$, on utilisant la méthode de rejet.

On applique la loi a priori (iid) khi-deux Geweke qui se base sur r degré de liberté pour les scalaires v_{ij} .

On prend r comme un hyper paramètre dégénérés, mais pas comme Koop(2003) l'a exposé, en faisant une extension où une distribution a priori exponentielle est placée en r .

On expose, d'une manière formelle, les lois a priori qui ont l'expression suivante :

$$\delta \propto \text{constant} \quad (1)$$

$$r/v_{ij} \sim \text{ID}\chi^2(r) \quad (2)$$

$$1/\sigma^2 \sim \Gamma(v_1, v_2) \quad (3)$$

$$\rho_i \sim U(-1, 1), \quad i = 1, 2 \quad (4)$$

La loi a posteriori, pour la distribution du paramètre δ prend la forme d'une normale multi variée.

$$\delta \setminus \rho_o, \rho_d, \sigma^2, V \sim N(\bar{\delta}, \sigma^2 \bar{D}) \quad (5)$$

$$\bar{\delta} = \beta^{(1)} - \rho_o \beta^{(2)} - \rho_d \beta^{(3)}$$

$$\beta^{(i)} = (Z^T V^{-1} Z)^{-1} Z^T V^{-1} F^{(i)}(Y)$$

$$F^{(i)}(Y) = Y, WY, YW^T, \quad i=1, \dots, 3$$

$$\bar{D} = (Z^T V^{-1} Z)^{-1}$$

La loi a posteriori pour le paramètre σ , est basé sur la loi a priori $1/\sigma^2 \sim \Gamma(v_1, v_2)$ qui est proportionnelle a l'inverse d'une distribution gamma .

$$\sigma^2 \setminus \rho_o, \rho_d, \delta, V \propto IG[(N+v_1)/2, (\tau(\rho) + v_2)/2] \quad (6)$$

$$\tau(\rho) = (1 - \rho_o - \rho_d)^T$$

$$Q_{ij} = \text{tr}[(E^{(i)})^T R (R^T E^{(j)})] \quad i, j = 1, \dots, 3$$

$$R = \begin{pmatrix} v_{11}^{-\frac{1}{2}} & \dots & v_{1n}^{-\frac{1}{2}} \\ \vdots & \ddots & \vdots \\ v_{n1}^{-\frac{1}{2}} & \dots & v_{nn}^{-\frac{1}{2}} \end{pmatrix}$$

$$E^{(i)} = F^{(i)}(Y) - \alpha^{(i)} i_n i_n^T - X \beta_d^{(i)} i_n^T - i_n (\beta_0^{(i)})^T X^T - \gamma^{(i)} G$$

$$E = E^{(1)} - \rho_o E^{(2)} - \rho_d E^{(3)}$$

$$\beta^{(i)} = (\alpha^{(i)} \quad \beta_d^{(i)} \quad \beta_0^{(i)} \quad \gamma^{(i)}) = (Z^T V^{-1} Z)^{-1} Z^T V^{-1} F^{(i)}(Y)$$

Où $IG(a, b)$ dénote la distribution inverse gamma avec les paramètres a, b .

On rappelle que $\tau(\rho) + v_2$ représente la somme des carré résiduels pour un paramètre donné ρ (voir IV-5-1) .

La loi a posteriori conditionnelle pour chaque valeur $v_{ij}, i, j=1, \dots, n$ s'exprime comme :

$$E_{ij}^2 + r/v_{ij} \setminus \rho, \delta, \sigma^2 \sim \chi^2(r+1) \quad (7)$$

Où E_{ij} fait référence à i, j élément de la matrice E .

Dans ce model on doit échantillonner chacun des deux paramètres $\rho_i, i = 1, 2$ en conditionnant par rapport à l'autre, que l'on dénote ρ_{-i} .

Connaissant les paramètres (δ, σ^2, V) le log de la loi a posteriori conditionnelle prend la forme suivante :

$$P(\rho_i / \rho_{-i}, \delta, \sigma^2, V) \propto |I_N - \rho_1 W_o - \rho_2 W_d| \cdot \exp\left(-\frac{1}{2\sigma^2} \tau(\rho)' Q \tau(\rho)\right) \quad (8)$$

$$\tau(\rho) = (1 - \rho_o - \rho_d)'$$

Comme ρ est aussi présent dans le déterminant, il est difficile de faire un échantillonnage à partir de cette distribution a posteriori conditionnelle. Nous allons donc adopter une démarche d'approximation en utilisant l'algorithme de Métropolis Hastings (II-3-2-1-1-b).

Annexe C

Opération matricielle : Produit d'HADAMARD et de KRONECKER

1) Le produit de HADAMARD, c'est in produit terme à terme. On dénote le produit de Hadamard avec le signe Θ ainsi, le produit de Hadamard de deux matrices de même dimension de A et B se note $A \Theta B$.

$A = a_{i,j}$ (de dimension L lignes et J colonnes)

$B = b_{i,j}$ (de dimension J lignes et K colonnes)

$A \Theta B = [a_{ij} \times b_{ij}]$

$$A \Theta B = \begin{pmatrix} a_{1,1} \times b_{1,1} & a_{1,2} \times b_{1,2} & \cdots & a_{1,J} \times b_{1,J} \\ a_{2,1} \times b_{2,1} & a_{2,2} \times b_{2,2} & \cdots & a_{2,J} \times b_{2,J} \\ \vdots & \vdots & \vdots & \vdots \\ a_{L,1} \times b_{L,1} & a_{L,2} \times b_{L,2} & \cdots & a_{L,J} \times b_{L,J} \end{pmatrix}$$

2) Produit de KRONECKER

Il existe encore une autre manière de définir la notion de produit d'une matrice (autre que le produit de Hadamard). Il s'agit du produit de Kronecker.

On dénote le produit de Kronecker avec le signe \otimes ainsi le produit de Kronecker de A et

B se note $A \otimes B$, est défini comme la matrice avec $(I \times K)$ lignes et $(J \times L)$ colonnes notée.

$$A \otimes B = \begin{pmatrix} a_{1,1} \times B & a_{1,2} \times B & \cdots & a_{1,J} \times B \\ a_{2,1} \times B & a_{2,2} \times B & \cdots & a_{2,J} \times B \\ \vdots & \vdots & \vdots & \vdots \\ a_{L,1} \times B & a_{L,2} \times B & \cdots & a_{L,J} \times B \end{pmatrix}$$

Annexe D : Soldes migratoires

Wilaya	Sorties	Entrées	Solde migratoire	Wilaya	Sorties	Entrées	Solde migratoire
MEDEA	72 865	13 357	-59508	TEBESSA	11 780	13 233	1453
ADRAR	30 092	9 194	-20898	TLEMCEN	15 453	17 594	2141
JIJEL	28 656	9 532	-19124	BEJAIA	19 190	22 109	2919
TISSEMSSILT	17 775	6 150	-11625	TIZI OUZOU	23 934	28 696	4762
RELIZANE	27 668	12 359	-15309	ALGER	123 151	148 111	24960
AIN DEFLA	23 691	11 956	-11735	MILA	14 701	17 898	3197
CHLEF	25 069	14 264	-10805	ANNABA	18 915	24 493	5578
BOUIRA	27 151	16 899	-10252	S,B,ABBES	14 690	19 644	4954
BECHAR	13 312	8 375	-4937	KHENCHELA	6 263	8 910	2647
O, E, BOUAGHI	17 365	12 165	-5200	BLIDA	41 416	59 668	18252
BATNA	24 000	16 855	-7145	MOSTAGANEM	10 214	15 453	5239
SAIDA	8 700	6 828	-1872	TAMANRASSET	5 092	8 376	3284
TIARET	22 776	18 282	-4494	BISKRA	13 560	22 747	9187
SOUK AHRAS	10 200	8 266	-1934	NAAMA	4 295	7 226	2931
M'SILA	21 423	17 406	-4017	EL BAYADH	4 781	8 066	3285
MASCARA	18 428	15 325	-3103	DJELFA	14 660	26 441	11781
B,B,ARRERIDJ	14 623	12 560	-2063	AIN TEMOUCHENT	7 580	13 891	6311
EL OUED	10 594	9 697	-897	TINDOUF	1 900	3 796	1896
LAGHOUAT	11 572	10731	-841	OUARGLA	10 715	22 941	12226
SKIKDA	12 819	11 916	-903	ORAN	31 263	72 222	40959
GUELMA	10 515	10 255	-260	TIPAZA	11 038	26 124	15086
SETIF	28 625	29 658	1033	BOUMERDES	18 420	47 476	29056
GHARDAIA	7 986	8 368	382	EL TARF	6 119	19 097	12978
CONSTANTINE	22 016	24 148	2132	ILLIZI	1 424	4 977	3553

Annexe E:

L'échantillonnage par MCMC

L'échantillonnage par MCMC peut commencer par des valeurs initiales arbitraires, que l'on note $\delta^0, (\sigma^2)^0, V^0$ et ρ_o^0, ρ_d^0 . on va procéder à un tirage séquentiel, tout en suivant les distributions a posteriori des paramètres (voir annexe B).

1. $p(\delta \mid \rho_o^0, \rho_d^0, (\sigma^2)^0, V^0)$, on va procéder à un tirage par rapport à la distribution multi normale avec comme moyenne $\bar{\delta}$ et variance $\sigma^2 \bar{D}$. On appellera ces nouveaux paramètres tirés δ^1 .

2. $p(\sigma^2 \mid \rho_o^0, \rho_d^0, \delta^1, V^0) \propto \text{IG}[(N+v_1)/2, (\tau(\rho) \text{Q}\tau(\rho) + v_2)/2]$, on appelle ce nouveau paramètre tiré $(\sigma^2)^1$.

3. $p(V \mid \delta^1, (\sigma^2)^1, \rho_o^0, \rho_d^0)$, qui suit une Qui Deux avec $r+1$ degrés de liberté, on déterminera ainsi, les éléments de la matrice R^2 . (Voir annexe B).

4. On suit les étapes de Metropolis pour échantillonner ρ_o .

$$\rho_o = \rho_o^0 + c\theta, \theta \sim N(0,1).$$

On appelle le paramètre c le paramètre de réglage, ensuite on évalue la probabilité d'acceptation :

$$\alpha = (\rho_o^0, \rho_o) = \min\left(\frac{p(\rho_o)}{p(\rho_o^0)}, 1\right)$$

Et finalement on attribue à $\rho_o^1 = \rho_o$ avec la probabilité (ρ_o^0, ρ_o) , sinon $\rho_o^1 = \rho_o^0$. toutefois nous allons adopter l'algorithme de rejet de Metropolis, pour contraindre ρ à être dans l'intervalle désiré. (Voir annexe B).

Même procédure que 4. Pour ρ_d .

On retourne à l'étape 1, en employant les paramètres mis à jour en remplacement des paramètres initiaux $\delta^0, (\sigma^2)^0, V^0$ et ρ_o^0, ρ_d^0 . et ainsi de suite.

A chaque itération on collecte un paramètre tiré, qui sera utilisé pour la construction des distributions a posteriori pour les paramètres du modèle.