

L'ANALYSE DE LA DISTANCE

DANS LA DIFFÉRENCIATION SPATIALE DES VALEURS FONCIÈRES

INTRODUCTION DES NOTIONS DE COVARIATION ET RÉGRESSION SPATIALES

Cécile HELLE

Laboratoire Structures et Dynamiques Spatiales
Université d'Avignon et des Pays de Vaucluse
Avignon

Résumé

L'hypothèse fondamentale de toutes les recherches sur la formation des valeurs du sol en milieu urbain postule la dépendance des prix fonciers de la distance au centre-ville. L'existence d'une telle relation est confirmée dans l'espace périurbain d'Avignon, mais la recherche de variables explicatives secondaires soulève un certain nombre de problèmes. L'ensemble des régresseurs du prix du mètre carré (superficie moyenne d'une transaction, rôle des professionnels de l'immobilier sur le marché, indice de congestion territoriale...) s'avère, en effet, lié plus ou moins intensément à la distance. Cette identité de structure spatiale (gradient aréolaire) entre la variable expliquée et les variables explicatives trouble l'interprétation : la nature des diverses covariations n'est pas explicite et les corrélations entre les régresseurs sont nombreuses.

Mots-clés

Avignon - Covariation spatiale - Distance - Gradient - Prix du terrain à bâtir - Régression spatiale

De nombreuses tentatives ont été menées, aux Etats-Unis [8] comme en Europe [5] [11], pour comprendre la formation des prix fonciers en milieu urbain : elles ont en commun de s'intéresser plus au phénomène qu'à sa répartition spatiale. Certes, par l'étude de la conjonction des fluctuations du coût du sol avec une ou plusieurs variables, les principales propriétés de la logique foncière dans les organismes urbains sont saisies, mais les incidences spatiales de cette logique demeurent obscures. A travers l'exemple du Grand Avignon, on se propose d'avoir recours aux modèles « explicatifs » les plus simples de la statistique classique (régressions simple et multiple, analyse de variance...) pour découvrir conjointement les structures spatiales et fonctionnelles qui régissent le jeu foncier dans cette périphérie urbaine.

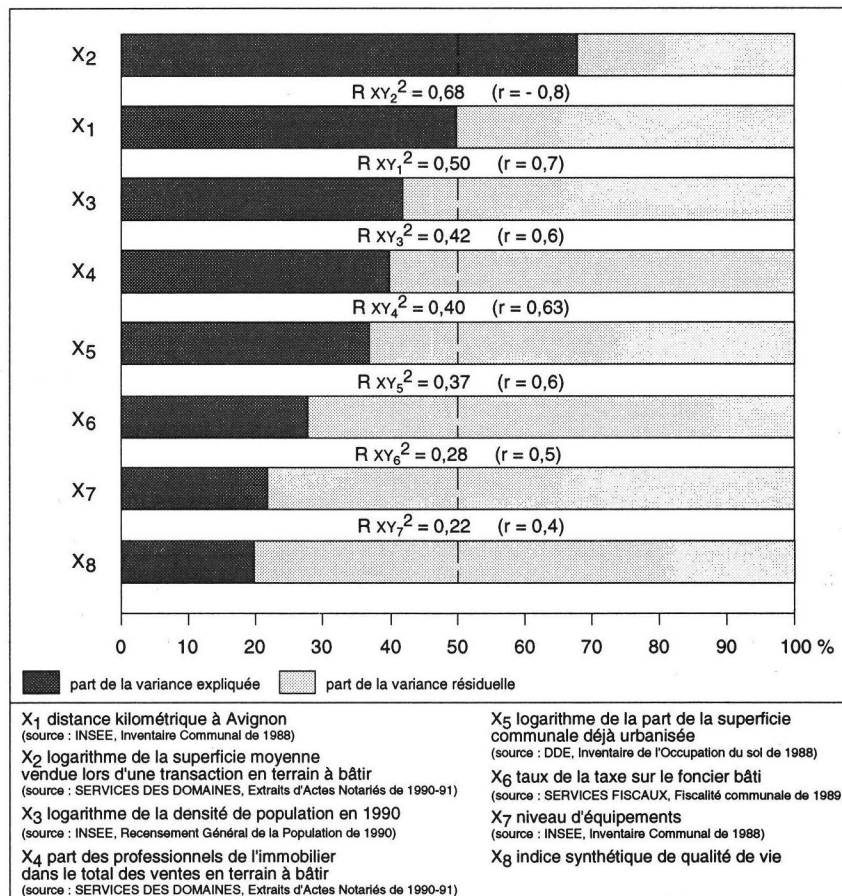
L'introduction de la modélisation statistique dans une optique d'analyse spatiale oriente les travaux et les résultats qui en découlent. Elle privilégie l'analyse des groupes spatiaux sur les groupes sociaux et fournit une vision globale de la géographie des valeurs foncières sans connaître réellement les processus et les motivations individuelles qui lui sont sous-jacents. Mais ce positionnement oblige surtout à se confronter à un certain nombre de difficultés. En effet, il implique la prise en compte des localisations par des techniques d'analyse universelles, transdisciplinaires, particulièrement performantes dans les sciences expérimentales, donc a priori peu appropriées aux spécificités de la géographie. Deux cas de figure sont en fait possibles : soit l'intégration de la dimension spatiale des objets analysés est explicite (les variables explicatives sont des expressions formalisées de l'espace) et les problèmes restent minimes ; soit, on estime que l'espace est indirectement contenu dans l'information (l'espace n'est pas mis dans le modèle, il est dans l'interprétation) et les difficultés sont plus nombreuses.

La diversité des facteurs explicatifs du jeu foncier amène à se trouver aussi bien dans l'une que dans l'autre de ces situations. Il suffit de considérer la distance à la ville, variable placée au cœur des théories foncières, comme une contrainte majeure de l'espace géographique pour que le modèle de régression acquière une dimension spatiale (on parle alors de "régression spatiale", [3]). De même, il suffit de considérer que dans une covariation, on cherche à étudier la répartition spatiale de deux phénomènes pour que cette analyse s'insère, elle aussi, dans une perspective spatiale (on parle alors de "covariation spatiale" [3]).

1. Des régressions simples pour identifier les principales variables explicatives du jeu foncier d'une périphérie urbaine, le grand Avignon

Dans l'espace périurbain d'Avignon, le déchiffrement de la géographie des prix du mètre carré de terrain à bâtir repose sur l'introduction de huit variables explicatives ($X_1 \dots X_8$). Outre la distance kilométrique à Avignon, sont également sélectionnés deux indicateurs de fonctionnement du marché des terrains à bâtir (la superficie moyenne d'une transaction et le rôle des professionnels de l'immobilier dans les procédures de vente), deux caractéristiques de démographie générale (la densité de population et l'indice de congestion spatiale selon le stade d'urbanisation), deux critères sur le cadre de vie communal (la desserte en équipement commercial et service de proximité et la pression fiscale), enfin une variable qualitative sur l'agrément différencié des sites. Ces divers paramètres couvrent de façon presque exhaustive le champ explicatif traditionnel de la cherté du sol urbain. Seul l'environnement social est omis en raison des problèmes rencontrés pour saisir le profil socio-professionnel des populations communales. Plusieurs variables sont exprimées sous forme logarithmique, pour respecter le principe de linéarité implicite au modèle de régression (par exemple, taille des parcelles vendues) ou pour homogénéiser des distributions statistiques caractérisées par des valeurs étendues et contrastées (par exemple, densité de population). Cette transformation a, en effet, pour propriété d'anamorphoser une courbe en droite et de diminuer les écarts entre des valeurs fortes.

Figure 1 : Les principales variables explicatives des prix fonciers dans les communes périurbaines du Grand Avignon



D'un point de vue purement statistique (fig. 1), toutes ces régressions simples sont pertinentes et satisfaisantes puisqu'elles permettent de préciser le sens et l'intensité des relations entre les huit variables explicatives ($X_1 \dots X_8$) et le prix du mètre carré de terrain à bâtir (Y). Mais elles ne clarifient pas, pour autant, les propriétés qui président à la configuration spatiale des valeurs foncières en périphérie urbaine. De toutes les analyses effectuées, seule l'interprétation de la régression spatiale ne prête à aucune discussion : c'est bien la distance à la ville qui explique les valeurs des

prix du mètre carré de terrain à bâtir. Une première structure spatiale de l'espace foncier est ainsi révélée : il s'agit du gradient aréolaire des prix. Au contraire, l'interprétation des autres covariations, troublée par « l'intrusion » de la distance, reste à faire. Prix fonciers et variables explicatives des prix semblent s'inscrire dans une même logique d'organisation de type gravitaire. On se demande alors si les covariations mises à jour ne reposent pas essentiellement sur cette identité de structure spatiale. Ce problème est comparable à celui qui se pose lors de l'analyse de la liaison entre deux chroniques affectées d'une même tendance. Est-ce que les promoteurs sont des déterminants directs de la cherté des prix du sol ou est-ce que leur influence est seulement liée au fait qu'ils tendent à se concentrer sur les marchés les plus saturés, donc les plus proches du centre urbain ?

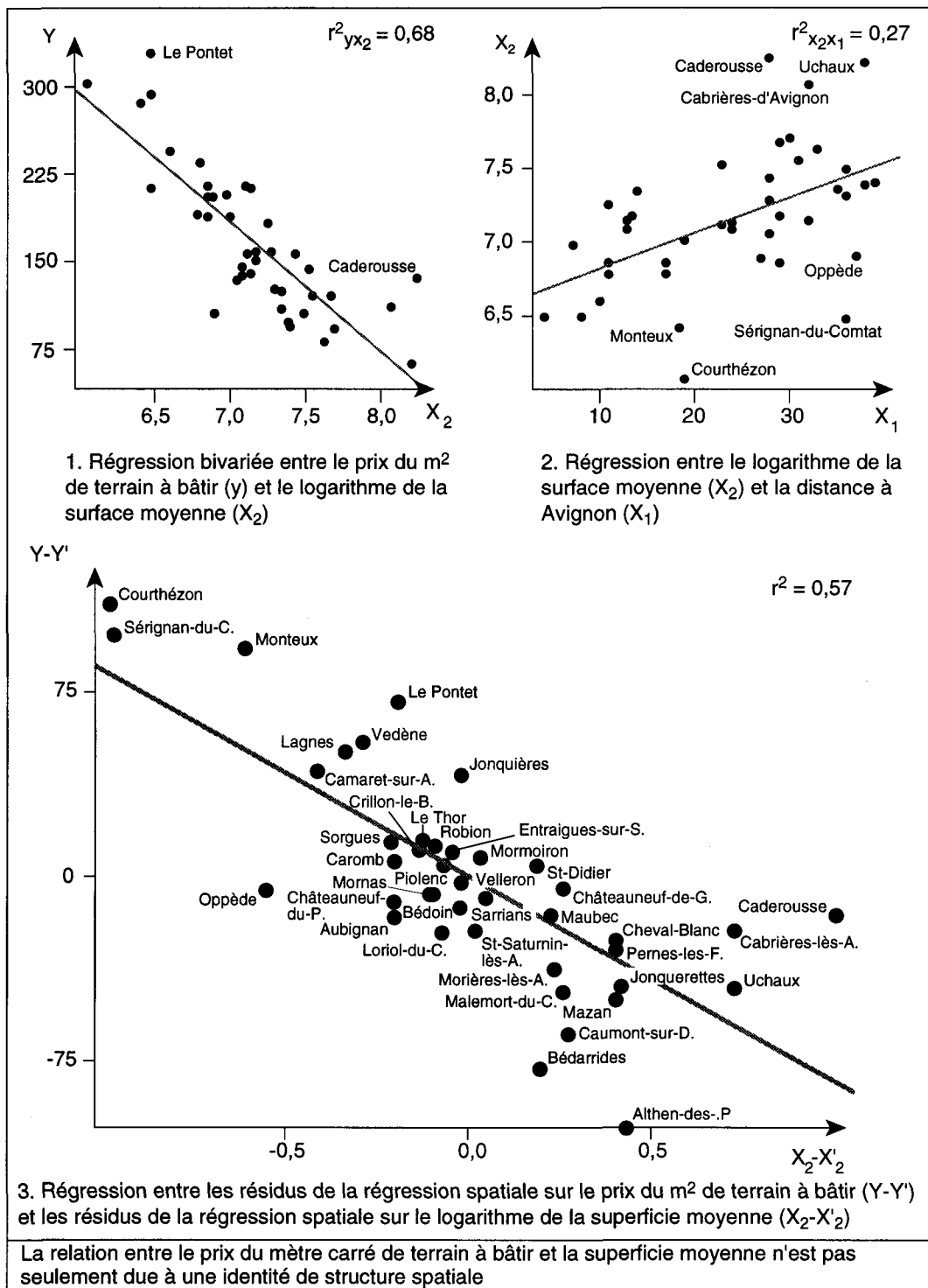
2. Des covariations spatiales pour établir ce que cache la distance

La distance, donnée essentielle de la spatialité avec laquelle les sociétés humaines ont à composer lors de la réalisation de leurs projets d'aménagement de l'espace, a tendance à occulter les autres formes possibles d'explication. Pour y voir plus clair, pour repérer et caractériser les composantes secondaires (spatiales ou structurelles) du jeu foncier, il convient de procéder comme dans une analyse de chroniques, à un filtrage de l'information. Il faut extraire la structure spatiale pour découvrir ce qu'elle cache. Ce prolongement de l'analyse des relations, indépendamment du rôle de l'espace, s'articule autour de trois étapes successives : la première consiste à s'assurer que le modèle récuratif d'organisation est bien de type aréolaire (elle nécessite simplement d'entreprendre des régressions spatiales sur les variables explicatives des prix), la deuxième neutralise la composante spatiale (elle repose sur le calcul des résidus spatiaux des variables indépendantes), enfin la troisième cherche à établir le sens réel des covariations initialement observées par l'examen des relations des résidus spatiaux. Si ces résidus sont en relation, la dépendance entre Y et X est d'une double nature, fonctionnelle et spatiale ; à l'inverse, si ces résidus sont sans relation, la covariation initiale ne reflète qu'une identité de structure spatiale.

Sur les représentations graphiques des régressions spatiales, l'allongement des nuages de points se vérifie systématiquement tandis que leur étroitesse est plus fluctuante : le premier caractère assure de l'action perturbatrice de la distance et le second de son intensité variable. La densité de population et l'indice de congestion territoriale se démarquent par des coefficients de détermination élevés, respectivement égaux à 0,63 et 0,57 : au fur et à mesure qu'on s'éloigne du pôle régional, ces deux variables enregistrent des valeurs de plus en plus faibles. Le logarithme de la superficie moyenne des transactions est le seul régresseur à être lié positivement à la distance. L'éloignement à Avignon s'accompagne d'une augmentation de la taille des parcelles constructibles. La distance a un effet amplificateur sur la surface des terrains à bâtir. Tous les facteurs explicatifs des valeurs foncières sont donc corrélés à la distance : comme les prix, ils se répartissent en auroles concentriques à partir du centre-ville. Mais alors que, pour certains, cette identité de structure spatiale explique partiellement les relations initialement observées, pour d'autres, elle épuise totalement l'explication.

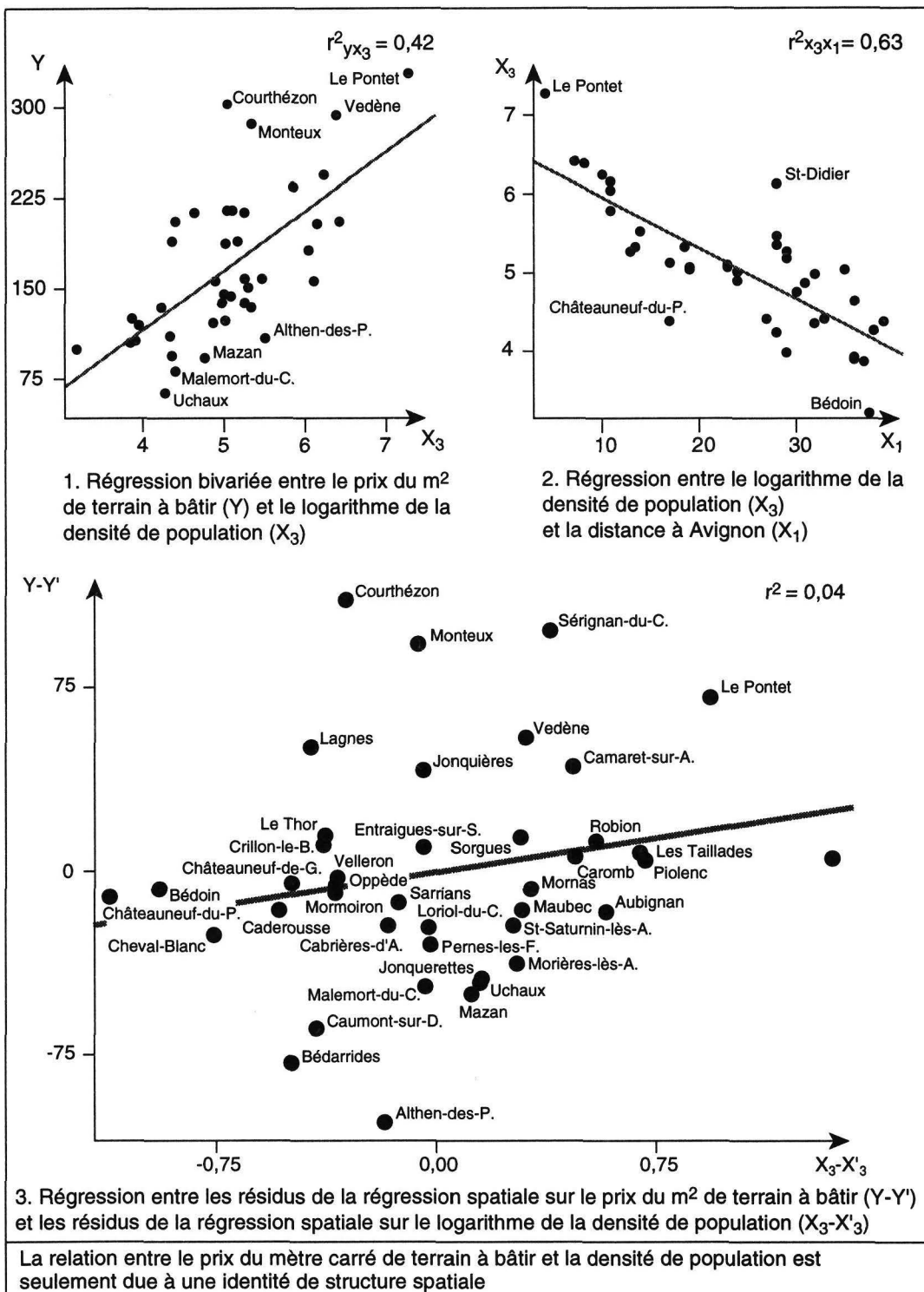
Après l'extraction de leur structure aréolaire, les deux caractéristiques du marché des terrains à bâtir continuent d'être proportionnelles au prix du mètre carré. Ainsi, les résidus spatiaux de la surface des terrains vendus expliquent-ils près de 58 % de ceux sur les prix (fig. 2). La pente négative de la droite de régression signale que les communes où le prix du m² est anormalement coûteux, étant donné leur localisation dans le Grand Avignon, se particularisent par le maillage très fin des zones récemment édifiées. Inversement, les localités anormalement « bon marché » s'illustrent par des lots constructibles plutôt vastes. L'extrême cherté de certaines communes situées dans les périphéries proches des centres urbains du Grand Avignon (Le Pontet, Monteux, Courthézon, Vedène, Sorgues) devient plus intelligible, tout comme le faible prix du mètre carré sur les entités des marges de cette région urbaine (Cheval-Blanc, Jonquières, Caumont-sur-Durance). Le maintien de relations dans les résidus démontre que les professionnels de l'immobilier et la taille des parcelles vendues fournissent une explication complémentaire de la logique foncière des périphéries urbaines. Principaux indicateurs de fonctionnement des terrains à bâtir, celle-ci est de nature fonctionnelle plus que spatiale (puisque les liaisons sont observées pour une distance fixée). Outre leur fréquente proximité avec le pôle urbain principal, les communes les plus chères se distinguent par une urbanisation organisée sous forme de lotissements, souvent promu par les professionnels de l'immobilier. S'y opposent des localités, en règle générale, éloignées de la ville centre, caractérisées par une extension du bâti plus lâche.

Figure 2 : L'analyse du rôle de l'espace dans une covariation, l'exemple du logarithme de la superficie moyenne d'une transaction



Les graphiques cartésiens des trois autres facteurs quantitatifs –indice de congestion territoriale, densité de population (fig. 3) et pression fiscale¹– se caractérisent par des nuages de points de forme arrondie et trapue, typique de l'indépendance statistique. La correspondance des fortes valeurs de prix, de densité, de saturation spatiale et de pression fiscale dans les communes de la proche banlieue avignonnaise ne légitime pas l'établissement de liaisons directes entre ces caractères. Ces fluctuations simultanées sont dues avant tout à la proximité du pôle régional, Avignon.

Figure 3 : L'analyse du rôle de l'espace dans une covariation, l'exemple de la densité de population communale



3. Utilité des régressions pour l'analyse spatiale

A l'issue de cette exploration statistique du jeu foncier du Grand Avignon, on se rend compte des précieux outils d'analyse spatiale que constituent les régressions simples et du rôle singulier joué par la distance dans la structuration des périphéries urbaines.

Les analyses de régression s'avèrent particulièrement performantes pour découvrir les régularités spatiales et fonctionnelles des configurations territoriales. La distance à la ville est un procédé commode pour « mettre en variable » l'espace. La régression spatiale est un moyen facile de modéliser la répartition aréolaire des valeurs foncières. Enfin, les covariations spatiales sont des méthodes simples pour trouver d'autres sources d'explication, à condition toutefois de bien préciser lors d'une étape initiale, le rôle de l'espace.

La logique foncière offre par ailleurs un aperçu des amples gradients que suivent les phénomènes économiques, sociaux et culturels dans les organismes urbains. Aux gradients des prix, des densités, des superficies constructibles vendues s'ajoutent ceux des déplacements pendulaires, des revenus ou du degré de motorisation des ménages... Tous rendent compte de l'interaction entre les décisions des individus ou des sociétés et la loi de la distance, que cette interaction soit directe (comme dans le cas des déplacements) ou indirecte (comme dans le cas des revenus, liés aux prix eux-mêmes dépendants de la distance). Derrière ces répartitions se dissimulent en effet les processus d'appropriation, de consommation, de gestion, de mobilité... qui président à la formation et à la transformation des périphéries urbaines et au cours desquels les hommes cherchent avant tout à minimiser les contraintes de cette donnée essentielle de la spatialité. Même si les auréoles peuvent présenter des contenus très divers, la structure demeure. On ne doit donc pas être étonné d'en retrouver la trace dès qu'on s'intéresse à l'organisation de régions urbaines polarisées. En revanche, on doit impérativement s'en défaire pour prétendre analyser convenablement ce qui lui échappe et qui peut relever d'autres propriétés de l'espace et/ou de règles fonctionnelles.

Enfin, même si elles tendent à schématiser exagérément la réalité en ne se préoccupant pas de la combinaison des effets des variables explicatives, les techniques de régression et covariation spatiales représentent une étape indispensable vers le déchiffrement de la complexité. En permettant de se forger une première opinion sur les propriétés spatiales et fonctionnelles à l'œuvre dans le jeu foncier, elles clarifient les desseins d'une exploration multivariée. D'une part, elles excluent les déviations classiques des régressions multiples vers une recherche exclusive de causalités ; d'autre part, elles fondent une stratégie originale d'explication multiple reposant sur la neutralisation préalable de la distance. Outre qu'elle fournit un résultat global totalement comparable au cheminement classique (part de variance expliquée égale à 83 %), la régression multiple sur les résidus spatiaux des prix et des diverses variables explicatives éclaire l'interprétation par la diminution du nombre de corrélations entre régresseurs.

Bibliographie

- [1] BELSLEY D.A., KUH E., WELSCH R.E. : *Regression diagnostics : Identifying influential data and sources of collinearity*, New York, John Wiley, 1980, 292 pages
- [2] BRIGHAM E.F. : « The determinants of residential land values », *Land Economics*, n° 4, 1965, pp. 325-334
- [3] CHARRE J. : *Statistique et territoire*, Montpellier, GIP RECLUS, 1995, 120 pages
- [4] CLARK V.A., DUNN O.J. : *Applied Statistics : Analysis of Variance and Regression*, New-York, John Wiley, 381 pages
- [5] DUTAÏLLY J.-C. : « Les valeurs foncières en région parisienne. Recherche d'un modèle », *Cahiers de l'I.A.U.R.I.F.*, vol. 25, 1971, 31 pages
- [6] GRANELLE J.J. : *La valeur du sol et ses facteurs : sources statistiques et données bibliographiques*, Paris, Centre de recherche et de documentation sur la consommation, 1965, 117 pages
- [7] GRANELLE J.J. : *Espace urbain et prix du sol*, Paris, Edition Sirey, Collection recherches économiques et financières, 1970, 292 pages

- [8] KNOS D.S. : « The distribution of land values in Topeka, Kansas », *Spatial Analysis. A reader in statistical geography*, 1962, pp. 269-289
- [9] JOHNSTON R.J. : *Multivariate statistical analysis in geography*, London, Longman, 1978, 280 pages
- [10] MOSTELLER F., TUKEY J.W. : *Data Analysis and regression. A second course in statistics*, Reading, Addison - Wesley Publishing Company, 1977, 588 pages
- [11] RUFFY V. : *Structure et dynamique économiques et géographiques d'un espace foncier*, Lausanne, Thèse Université de Lausanne, 1989, 320 pages
- [12] SCHEFFE H. : *The analysis of variance*, New York, John Wiley, Inc, 1959, 477 pages
- [13] SILK J. : « Analysis of covariance and comparison of regression lines », *Concepts and techniques in modern Geography*, n° 20, 1979, pp. 3-45
- [14] TODD D. : « An introduction to the use of simultaneous regression analysis in geography », *Concepts and techniques in modern Geography*, n° 21, 1979, pp. 3-42
- [15] WRIGLEY N. : « An introduction to the use of logit model in geography », *Concepts and techniques in modern Geography*, n° 10, 1976, pp. 3-32
- [16] YEATES M.H. : « Some factors affecting the spatial distribution of Chicago land values, 1910 - 1960 », *Economic Geography*, n° 41, 1965, pp. 57-70

Notes

¹ - L'INSEE ne publie cette information que pour les communes comptant plus de 2 000 habitants, seuil au-dessous duquel la fiabilité des données n'est plus garantie. Par ailleurs, les tentatives de régressions entre le prix du mètre carré et certaines catégories d'acheteurs (cadres, ouvriers...) se sont avérées peu concluantes

² - La mesure synthétique de la desserte en équipements est écartée de l'analyse en raison du faible coefficient de détermination obtenu à l'issue de la régression spatiale sur ce régresseur ($r^2 = 0,17$). Par ailleurs, la neutralisation de la distance dans la relation entre les prix fonciers et la variable sur l'agrément des sites entre dans une autre logique (celle de l'analyse de covariance) du fait de la nature qualitative de cette variable