

Documents de travail

du Laboratoire d'Economie et de
Gestion
Working Papers

**Économétrie spatiale et données spatiales empilées dans le temps :
Proposition d'une modélisation adaptée**

Jean Dubé†, Diègo LEGROS*

† Université du Québec à Rimouski (UQAR),
300, allée des Ursulines, Rimouski, Canada, G5L 3A1, jean_dube@uqar.qc.ca

*Université de Bourgogne & CNRS
Fre3496 Laboratoire d'Economie et de Gestion

[e2013-01]

Equipe Analyse et Modélisation des Interactions Economiques (AMIE)

Laboratoire d'Economie et de Gestion
Université de Bourgogne & CNRS Fre3496

Pôle d'Economie et de Gestion - 2 boulevard Gabriel BP 26611 - F21066 DIJON cedex
Tel. +33 (0)3 80 39 54 30 Fax +33 (0)3 80 39 54 43
secretariat.leg@u-bourgogne.fr - www.u-bourgogne.fr/leg



Économétrie spatiale et données spatiales empilées dans le temps : Proposition d'une modélisation adaptée

RÉSUMÉ –

L'article s'attache à présenter la particularité des données spatiales empilées et montre pourquoi il peut être contre-indiqué d'utiliser les approches spatiales développées pour les données en coupe transversale empilées dans le temps. La dimension temps implique des relations unidirectionnelles, par opposition aux relations multidirectionnelles spatiales. La construction d'une matrice de pondérations spatio-temporelles unique, à partir de matrices de pondérations spatiales et temporelles, permet d'utiliser les modèles et les tests développés pour les données spatiales tout en tenant compte des deux dimensions simultanément. Une série d'applications empiriques montre que la non prise en compte de la dimension temporelle dans les analyses a pour conséquence de surévaluer les mesures de la dépendance spatiale en plus de surévaluer les coefficients autorégressifs spatiaux estimés. Finalement, la prise en compte des deux dimensions, spatiales et temporelles, permet de générer de nouvelles variables explicatives dynamiques, comparables à des effets de pairs, qui s'avèrent significatives dans l'explication des prix de vente immobiliers.

Spatial Econometric and Spatial Data Pooled over Time: Towards an adapted modelling approach

ABSTRACT –

This paper presents the characteristics of spatial data pooled over time and show why these data bases cannot be considered as the same way as spatial panel data or strictly spatial data. The temporal dimension implies a unidirectionality of relations, while spatial relations are multidirectional. The construction of spatio-temporal weights matrix, lying on spatial and temporal weights matrices, allow to use usual statistic models and tests developed for spatial analysis while accounting simultaneously for temporal and spatial dimensions. Empirical examples established the impact of neglecting the temporal dimension in spatial analysis and show how such approach overestimate the pattern of spatial dependence as well as overestimate the spatial autoregressive coefficient estimated. Finally, accounting for both dimensions, spatial and temporal, allow to generate additional independent variables, considering dynamic effect, that appear to play a significant role on determination of real estate prices.

INTRODUCTION

Si le développement de l'économétrie spatiale date d'une quarantaine d'années¹ (Goodchild, 2009; Haining, 2009), les progrès en matière de modélisation spatio-temporelle sont relativement récents (Elhorst, 2003; Baltagi et Li, 2004; Baltagi et al., 2003, 2007; Kapoor et al., 2007; Yu et Lee, 2008; Lee et Yu, 2010; Parent et Lesage, 2011). Les avancées en matière de problématique spatio-temporelles portent principalement sur les modèles appliqués à des données de type panel. Cependant, les données spatio-temporelles ne prennent pas systématiquement la forme d'un panel, cylindré ou pas. Dans plusieurs cas, les données spatiales sont récoltées dans le temps sans que les unités soient observées plus d'une fois sur l'horizon temporel. Le cas des données de transactions immobilières est particulièrement frappant : les données possèdent de réelles caractéristiques spatiales et elles sont récoltées au fil du temps sans que la majorité des observations soient répétées (Case et Shiller, 1989; Abraham et Schauman, 1991; Clapp *et al.*, 1991; Dubé *et al.*, 2011a).

Traditionnellement, lorsque l'on mobilise des données individuelles spatiales empilées dans le temps, on suppose que l'effet du temps est négligeable dans le processus générateur des données, ce qui justifie l'utilisation des méthodes d'analyses spatiales usuelles. Or, certains travaux récents ont mis en lumière l'importance de tenir compte de la dimension temps dans les études spatiales (LeSage et Pace, 2009; Dubé et Legros, sous presse). D'autres travaux ont proposé de tenir compte de la particularité des données spatiales empilées dans le temps en construisant une matrice de pondérations spatio-temporelles pour tenir compte simultanément des deux dimensions (Smith et Wu, 2009; Huang et al., 2010; Nappi-Choulet et Maury, 2011; Dubé et Legros, 2011; Dubé *et al.*, 2011b). La forme unique de la matrice permet de conserver les propriétés des modèles spatiaux et des statistiques spatiales existants.

La construction d'une matrice de pondérations spatio-temporelles unique repose sur deux éléments distincts : une matrice de pondérations spatiales et une matrice de pondérations temporelles. La matrice de pondérations temporelles tient compte de l'unidirectionnalité des relations : les observations passées peuvent avoir une influence sur les observations actuelles, alors que l'inverse est plutôt improbable², voire impossible (selon le pas de temps retenu). Cette matrice permet d'identifier les observations qui sont enregistrées durant la même période, mais également les observations qui précèdent ou qui suivent. La matrice de pondérations spatiales tient compte de la multidirectionnalité des relations : une observation influence le comportement des voisins, alors que les voisins influencent également son propre comportement. La matrice de pondérations spatio-temporelles permet de conjuguer les relations multidirectionnelles spatialisées pour une même période en plus de cerner les relations unidirectionnelles liant les observations spatialisées passées aux observations spatialisées présentes et futures.

L'objectif de cet article est de montrer comment le développement de matrices de pondérations spatio-temporelles et leur utilisation dans les approches économétriques

¹ Anselin (2010) suggère plutôt que l'économétrie spatiale serait vieille d'une trentaine d'années.

² En fait, le phénomène d'anticipation constitue une figure de cas où le futur influence en partie les réalisations d'aujourd'hui.

spatiales³ peuvent aider à contrôler pour la réalité temporelle de la collecte de données spatiales individuelles et non répétées. Plus particulièrement, l'article cherche à vérifier si les conclusions quant à une possible surestimation de l'effet spatial dans un contexte spatio-temporel sont vérifiées empiriquement et comment cette surestimation peut se refléter sur l'ampleur des coefficients autorégressifs estimés. L'article montre également comment la construction de matrices de pondérations spatio-temporelles peut aider à développer de nouvelles variables qui peuvent ensuite être introduites dans le modèle afin de capter les effets de dynamiques temporelles spatialisées. Trois exemples empiriques sont présentés à partir des données de transactions immobilières de Québec (1990-1996), de Paris (1990-2001) et de Lucas County, en Ohio aux États-Unis (1993-1998).

L'article est divisé en quatre sections. La première section présente le processus de structuration des données spatiales empilées dans le temps. Elle montre comment la dimension temporelle (unidirectionnelle) peut jouer un rôle important dans la détermination des relations spatiales (multidirectionnelles) et montre l'importance de restructurer les relations dans une forme hybride. La seconde section montre comment considérer les deux types de relations, multidirectionnelles spatiales et unidirectionnelles temporelles, à partir d'une matrice de pondérations spatio-temporelles qui tient compte à la fois de l'éloignement spatial et de l'éloignement temporel des observations. La troisième section présente quelques applications empiriques effectuées sur des données de transactions immobilières au Canada, en France et aux États-Unis afin de tirer quelques constats sur l'influence que peut avoir le développement d'une matrice spatio-temporelle sur le possible biais lié à la seule considération de la dimension spatiale lorsque les données sont empilées dans le temps. Finalement, une brève conclusion, effectuant un retour sur les principaux constats, clôture le papier.

STRUCTURATION DES DONNÉES SPATIALES EMPILÉES DANS LE TEMPS

La représentation de "l'effet de l'éloignement spatial" entre les observations provient de la première loi de la géographie (Tobler, 1970) qui stipule que tous les phénomènes sont reliés entre eux, mais que les plus près le sont plus fortement. Cette vision est inspirée d'une structure en deux dimensions (latitude (Y), longitude (X)) permettant d'établir la distance séparant deux observations (Figure 1).

INSÉRER FIGURE 1 ICI

Or, la configuration de certaines bases de données spatiales fait en sorte que cette représentation en deux dimensions néglige l'aspect temporel de la collecte des observations spatiales individuelles (Figure 2). En fait, la structuration des données s'apparente plutôt à une collection de couches spatiales (une par période) empilée au fil du temps. Ainsi, la métrique "distance spatiale" séparant les observations cache une réalité fondamentale : l'espace n'est pas la seule dimension dans le processus de collecte

³ Les exemples empiriques reposent sur l'exploitation des programmes d'estimation existants sous le logiciel MatLab (LeSage, 1999) en modifiant la forme de la matrice de pondérations afin d'exploiter les informations des observations sous trois dimensions : latitude, longitude et temps.

des données. À la distance spatiale, il est nécessaire d'ajouter la distance temporelle tout en tenant compte de l'effet unidirectionnel de cette dimension.

INSÉRER FIGURE 2 ICI

Malgré cette particularité, la plupart des applications empiriques utilisant ce type de données continuent de supposer, implicitement, que la dimension temps n'a aucun impact puisque les analyses reposent sur les méthodes spatiales classiques malgré le fait que cette hypothèse, plutôt forte (LeSage et Pace, 2009; Dubé *et al.*, 2011b; Dubé et Legros, sous presse), n'est que rarement vérifiée et validée. Une première tentative de formaliser les liens spatiaux et temporels entre les observations et de tenir compte de la réalité temporelle des données spatiales empilées dans le temps date de plus de dix ans (Pace *et al.*, 1998, 2000). Depuis, certaines applications ont repris ce cadre d'analyse pour certaines applications empiriques particulières (Tu *et al.*, 2004; Sun *et al.*, 2005) ou encore pour servir de point de départ pour la modélisation spatio-temporelle (Nappi-Choulet et Maury, 2009). Cependant, cette approche a le désavantage de ne pas croiser les informations spatiales et temporelles, mais plutôt d'utiliser de façon indépendante les matrices de pondérations spatiales et temporelles. En fait, le produit matriciel des deux matrices de pondérations capture de façon indirecte seulement les effets spatio-temporels (Smith et Wu, 2009).

Une autre façon d'incorporer la métrique "distance temporelle" dans la modélisation spatio-temporelle est de spécifier un processus autorégressif temporel des termes d'erreurs (Gelfand *et al.*, 1998). Cette spécification fut reprise afin de contrôler pour l'effet dynamique de la détermination des valeurs immobilières tout en tenant compte des effets de débordements spatiaux (Smith et Wu, 2009). Bien que ce processus a l'avantage de tenir compte de l'unidirectionnalité de l'effet temps, il néglige néanmoins la situation géographique des observations et suppose ainsi que l'effet dynamique est identique pour des observations spatialement éloignées, mais temporelles rapprochées (Dubé et Legros, 2011).

Seule une spécification permettant de contrôler à la fois pour la distance spatiale et temporelle peut tenir compte du processus réel de structuration des observations spatiales individuelles collectées au fil du temps (Dubé *et al.*, 2011b). Une telle structure doit donc prendre en compte à la fois les distances spatiales et les distances temporelles (Figure 3).

INSÉRER FIGURE 3 ICI

Le développement d'une matrice de pondérations spatio-temporelles qui tiennent compte simultanément des réalités spatiales et temporelles s'avère une solution intéressante. Cette approche permet ainsi de contrôler pour l'effet de rétroaction multidirectionnel spatial pour une période de temps délimité. De plus, la création d'un ensemble de matrices de pondérations permet également de développer de nouvelles variables explicatives qui peuvent être ajoutées aux modèles dans le but de capter différents patrons de relations spatio-temporelles.

CONSTRUCTION D'UNE MATRICE DE PONDÉRATIONS SPATIO-TEMPORELLES

La construction d'une matrice de pondérations spatio-temporelles repose sur le développement de deux matrices : une matrice de pondérations spatiales et une matrice de pondérations temporelles. Chacune de ces matrices possède des propriétés particulières (Dubé *et al.*, 2011b, à venir) qui leur confèrent une structure différente. Néanmoins, elles permettent, de façon générale, de relier l'ensemble des N_T observations⁴ entre elles.

MATRICE DE PONDÉRATIONS SPATIALES

Une matrice de pondérations spatiales, de dimension $N_T \times N_T$, permet d'établir les liens spatiaux entre deux observations, i et j , à partir d'un élément particulier, noté s_{ij} . L'élément particulier est habituellement défini en fonction de l'inverse de la distance séparant les observations i et j , d_{ij} , tout en introduisant un paramètre de friction, α , pénalisant plus ou moins fortement pour l'éloignement⁵ (équation 1).

$$s_{ij} = \begin{cases} \bar{d}_{ij}^{-\alpha} & \text{si } d_{ij} \leq \bar{d} \\ 1 & \text{si } d_{ij} = 0 \\ 0 & \text{sinON} \end{cases} \quad \forall \begin{matrix} i \neq j \\ i \neq j \end{matrix} \quad (1)$$

où \bar{d} représente une distance seuil au-delà de laquelle l'influence spatiale entre deux observations est supposée nulle. Au final, l'ensemble des éléments individuels permettant de relier spatialement les observations entre elles est utilisé pour générer la matrice de pondérations spatiales, non standardisée, notée \mathbf{S} (équation 2).

$$\mathbf{S} = \begin{bmatrix} 0 & s_{12} & \dots & s_{1j} & \dots & s_{1N_T} \\ s_{21} & 0 & \dots & s_{2j} & \dots & s_{2N_T} \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ s_{ji} & s_{j2} & \dots & s_{jj} & \dots & s_{jN_T} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{N_T1} & s_{N_T2} & \dots & s_{N_Tj} & \dots & 0 \end{bmatrix} \quad (2)$$

La construction de la matrice se fait habituellement, mais non exclusivement, de manière exogène (Chasco et Lopez, 2008; Fingleton, 2009). Nonobstant les débats autour de la manière de conceptualiser la matrice de pondérations (Giacomini et Granger, 2004; Bhattacharjee et Holly, 2011), un des défis consiste à fixer le paramètre la distance seuil optimale, \bar{d} . Certains proposent d'adopter des valeurs qui reflètent, au meilleur des connaissances, la réalité analysée par le modélisateur (Getis, 2009). D'autres suggèrent d'adopter une valeur qui fait en sorte de maximiser l'estimation de l'autocorrélation spatiale (Boots et Dufournaud, 1994). Dans tous les cas, la

⁴ Le nombre total d'observations est défini par la somme des observations sur chacune des T périodes, N_t : $N_T = \sum_{t=1}^T N_t \neq NT$.

⁵ Le paramètre de friction prend habituellement une valeur de zéro (voisinage), de un (inverse de la distance) ou de deux (inverse de la distance au carré).

détermination de la structure optimale des éléments de la matrice demeure un défi (Getis et Aldstadt, 2004).

Une fois la matrice de pondérations spatiales construite, il est de pratique courante de standardiser celle-ci. La standardisation est effectuée afin d'assurer que les résultats obtenus peuvent être comparables entre eux et facilite l'interprétation des variables spatialement décalées. La standardisation consiste à prendre chacun des éléments d'une ligne et de les diviser par la somme des éléments de la ligne (équation 3).

$$s_{ij}^* = \frac{s_{ij}}{\sum_j s_{ij}} \quad (3)$$

La nouvelle forme de la matrice standardisée, notée S^* , est ensuite utilisée pour calculer les statistiques de tests permettant de détecter la présence d'autocorrélation spatiale ou encore pour estimer un paramètre autorégressif.

MATRICE DE PONDÉRATIONS TEMPORELLES

Une matrice de pondérations temporelles, de dimension $N_T \times N_T$, permet aussi d'établir les liens, temporels ceux-là, entre deux observations, i et j , à partir d'un élément particulier, noté t_{ij} . Un élément particulier est défini en fonction de la distance temps séparant l'observation i , dont la valeur temporelle est établie à v_i , et l'observation j , dont la valeur temporelle est établie à v_j . L'inverse de la distance temporelle séparant les observations i et j , $v_i - v_j$, est ainsi considéré tout en introduisant un paramètre de friction, γ , permettant de pondérer plus ou moins fortement pour l'éloignement⁶ (équation 4).

$$t_{ij} = \begin{cases} c(v_i - v_j)^{-\gamma} & \text{si } v_i - v_j \leq \bar{v}_p; \forall i \neq j; v_i \neq v_j \text{ et } v_i > v_j \\ \kappa |v_i - v_j|^{-\gamma} & \text{si } v_i - v_j \leq \bar{v}_f; \forall i \neq j; v_i \neq v_j \text{ et } v_i < v_j \\ 1 & \text{si } v_i = v_j \forall i \neq j \\ 0 & \text{sinon} \end{cases} \quad (4)$$

où \bar{v}_p représente une autre distance seuil au-delà de laquelle l'influence temporelle du passé est supposée nulle, \bar{v}_f représente une distance seuil au-delà de laquelle l'influence temporelle liée à l'anticipation est nulle et κ un scalaire permettant de mettre une moins grande importance sur les données passées et futures que sur celles du présent ($\kappa \leq 1$). La fonction temporelle, v_i , permet d'établir, en fonction du pas de temps établi, la chronologie des observations pour des trimestres⁷ (équation 5a) ou des mois⁸ (équations 5b).

$$v_i = 4 \times (\text{année} - \text{année de base}) + \text{trimestre} \quad (5a)$$

⁶ Voir la note de bas de page numéro 5.

⁷ Une observation i est ainsi récoltée à l'année $yyyy_i$ et au trimestre qq_i .

⁸ Une observation i est ainsi récoltée à l'année $yyyy_i$ et au mois mm_i .

$$v_i = 12 \times (\dots) \quad (5b)$$

où $yyyy_{min}$ représente la première année à laquelle les observations sont disponibles. La forme des éléments généraux permet une généralisation de plusieurs applications proposées jusqu'à maintenant, dont la structure de Pace *et al.* (1998, 2000) et de Smith et Wu (2009).

En supposant que les observations sont préalablement ordonnées de façon chronologique, la forme de la matrice de pondérations temporelles peut être décomposable selon les parties triangulaires. Certains éléments de la partie triangulaire supérieure peuvent être négatifs lorsque l'unité i est observée avant l'unité j . Pour cette raison l'opérateur de valeur absolue est utilisé dans la définition des éléments généraux de la matrice (équation 4). Les éléments non nuls et non unitaires (différents de 1) relient les observations futures aux observations présentes. La partie triangulaire inférieure permet d'obtenir les liens unissant les observations passées aux observations présentes. Finalement, les éléments unitaires des deux parties triangulaires marquent les relations de simultanéité temporelle (ligne 3 - équation 4)⁹. Au final, l'ensemble des éléments généraux permettant de relier l'ensemble des observations entre elles définit la spécification de la matrice de pondérations spatiales, notée \mathbf{T} (équation 6).

$$\mathbf{T} = \begin{bmatrix} 0 & t_{12} & \dots & t_{1j} & \dots & t_{1N_T} \\ t_{21} & 0 & \dots & t_{2j} & \dots & t_{2N_T} \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ t_{i1} & t_{i2} & \dots & t_{ij} & \dots & t_{iN_T} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{N_T1} & t_{N_T2} & \dots & t_{N_Tj} & \dots & 0 \end{bmatrix} \quad (6)$$

La construction de la matrice est déterminée de manière exogène. Évidemment, un des défis consiste à établir les paramètres liés à la distance seuil optimale des anticipations, \tilde{v}_f , et à la distance seuil optimale pour les observations passées, \tilde{v}_p . Ces valeurs peuvent simplement être fixées par le modélisateur en fonction de certains *a priori*.

MATRICE DE PONDÉRATIONS SPATIO-TEMPORELLES

La matrice de pondérations spatio-temporelles, notée \mathbf{W} , est obtenue en multipliant les éléments, termes à termes, de la matrice de pondérations spatiales à la matrice de pondérations temporelles (équation 7).

⁹ En supposant que la constante multiplicative, κ , soit inférieure à 1.

$$\mathbf{W} = \begin{bmatrix} 0 & s_{12} \times t_{12} & \cdots & s_{1j} \times t_{1j} & \cdots & s_{1N_T} \times t_{1N_T} \\ s_{21} \times t_{21} & 0 & \cdots & s_{2j} \times t_{2j} & \cdots & s_{2N_T} \times t_{2N_T} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ s_{i1} \times t_{i1} & s_{i2} \times t_{i2} & \cdots & s_{ij} \times t_{ij} & \cdots & s_{iN_T} \times t_{iN_T} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{N_T1} \times t_{N_T1} & s_{N_T2} \times t_{N_T2} & \cdots & s_{N_Tj} \times t_{N_Tj} & \cdots & 0 \end{bmatrix} \quad (7)$$

Cette opération peut être simplement résumée avec l'aide d'un opérateur matriciel : le produit Hadamard (équation 8).

$$\mathbf{W}_{(N_T \times N_T)} = \mathbf{S}_{(N_T \times N_T)} \odot \mathbf{T}_{(N_T \times N_T)} \quad (8)$$

Cette forme de matrice permet de s'assurer que les relations spatiales sont retenues seulement lorsqu'elle prennent forme à la même période. Les relations spatiales multidirectionnelles se répartissent chaque côté de la diagonale principale, composée d'éléments nuls. La forme de la matrice permet également de tenir compte d'une unidirectionnalité des effets spatiaux qui surviennent à des périodes de temps différentes. La partie triangulaire inférieure de la matrice permet de tenir compte des relations unidirectionnelles spatiales. Évidemment, les composantes multidirectionnelles peuvent aussi tenir compte d'un certain écart temporel pour lequel la proximité temporelle est assez près pour supposer une quasi-simultanéité. Finalement, la matrice spatio-temporelle permet de tenir compte de possibles relations d'anticipations. La partie triangulaire supérieure marque ces relations. Au final, la structure de la matrice de pondérations spatio-temporelles permet de respecter les contraintes spatiales et les contraintes temporelles à l'intérieur d'une structure unique.

La matrice spatio-temporelle peut, comme c'est habituellement le cas avec les matrices de pondérations strictement spatiales, être standardisée en ligne et utilisée pour détecter les patrons d'autocorrélation spatiale dans le temps (Dubé et Legros, sous presse), pour estimer les coefficients autorégressifs spatiaux tout en contrôlant pour la dimension temps (Dubé *et al.*, à venir) ou encore pour estimer des effets dynamiques spatialement localisés (Dubé et Legros, 2011, à venir). La standardisation fonctionne de la même façon que pour la matrice spatiale.

APPLICATIONS EMPIRIQUES

Les applications empiriques cherchent à déterminer si l'utilisation de matrices de pondérations spatiales, par opposition aux matrices de pondérations spatio-temporelles, surestime l'effet spatial dans un contexte où les observations spatiales individuelles sont collectées dans le temps. Plus particulièrement, l'article s'attarde à tester trois hypothèses :

i) les tests de détection de l'autocorrélation spatiale sont systématiquement biaisés vers le rejet de l'hypothèse nulle d'absence d'autocorrélation spatiale, dans un contexte spatio-temporel, lorsque les matrices de pondérations spatiales sont utilisées;

ii) les coefficients autorégressifs estimés dans les modèles économétriques spatiaux sont plus importants lorsque les matrices de pondérations spatiales sont utilisées

par rapport aux matrices de pondérations spatio-temporelles du fait que la dépendance spatiale mesurée est plus forte; et

iii) les matrices de pondérations spatio-temporelles permettent de générer de nouvelles variables captant un effet dynamique qui s'avère significatif.

La dernière hypothèse permet d'améliorer et de perfectionner la modélisation afin de tenir compte de la notion des "comparables", qui s'apparente à la notion d'"effets de paires", largement reconnue en économie de l'éducation et en économie du travail.

Les applications consistent en des travaux antérieurs et font une synthèse des résultats obtenus jusqu'à présent dans le but de tracer certaines tendances communes. La comparaison des résultats permet aussi d'évaluer, à la lumière des résultats et comparaisons, l'impact d'utiliser une approche de modélisation strictement spatiale dans un contexte où les données consistent en une collection de couches d'observations spatiales individuelles empilées dans le temps. Le premier exemple provient du Québec (Canada), le second de la France et le dernier des États-Unis.

Le cas de Québec (Canada)

Les données pour les applications portant sur la ville de Québec proviennent d'une application empirique cherchant à déterminer l'impact des comparables dans la détermination des valeurs marchandes immobilières (Des Rosiers *et al.*, 2011). Dans cette recherche, les auteurs s'attardent aux transactions résidentielles unifamiliales recensées sur le territoire de la ville de Québec (incluant les villes défusionnées) entre avril 1990 et décembre 1996¹⁰. La base de données contient 15 729 transactions et chaque observation fait état du prix nominal de vente (en dollars canadiens)¹¹ et d'une multitude de caractéristiques physiques¹² et environnementales¹³ des résidences telles que l'accessibilité aux services les plus proches (Des Rosiers *et al.*, 2000). De plus, l'espace est structuré en sept (7) sous-marchés (Voisin *et al.*, 2010) afin de contrôler pour certains effets fixes liés à la localisation qui influence la détermination des valeurs, mais dont l'identification de la source de l'effet est probablement historique. Ces variables d'effets fixes ont pour conséquence de contrôler en partie pour le problème d'autocorrélation spatiale (Dubé *et al.*, 2011c).

Cette première analyse cherche à vérifier la première hypothèse, à savoir si l'omission de la dimension temps à un impact sur la sur-évaluation du patron de dépendance spatial entre les résidus de l'équation de prix hédonique. Les résultats suggèrent que l'utilisation de la matrice de pondérations spatiales dans un contexte spatio-temporel surestime le patron de dépendance spatial entre les observations, et ce peu importe le type de matrice retenu et la spécification du modèle de prix hédonique (Tableau 1). Malgré la surestimation de l'effet de dépendance spatiale, les résultats

¹⁰ À l'exception de l'année 1992 pour laquelle il n'y a aucune transaction formellement disponible.

¹¹ Les données retenues visent les propriétés dont les prix de transaction se situaient entre 50 000\$ et 250 000\$ canadien.

¹² Ce sont les caractéristiques intrinsèques telles que la surface habitable, la surface du terrain, l'âge de la construction, le nombre de salles de bains, la qualité des matériaux, les caractéristiques architecturales, la présence de cheminées, d'une piscine, de terrasses, le type de garage, etc.

¹³ Ce sont les caractéristiques extrinsèques telles que le revenu médian du quartier, la proportion de familles monoparentales, de diplômés de l'enseignement supérieur, l'accessibilité aux services régionaux ou locaux, etc.

montrent néanmoins qu'un patron de dépendance spatiale est significatif même en tenant compte de la dimension temporelle. Un premier constat est donc qu'il ne suffit pas de tenir compte de la dimension temps pour contrôler en totalité le problème d'autocorrélation spatiale des termes d'erreurs. Ces résultats supportent les conclusions précédentes dressées par Dubé et Legros (sous presse). Néanmoins, ils montrent l'importance de tenir compte de la dimension temps dans le calcul de la dépendance spatiale.

INSÉRER TABLEAU 1 ICI

Évidemment, une surévaluation du patron de dépendance spatiale peut également cacher un autre problème : celui de la surestimation des coefficients autorégressifs estimés dans les modèles spatiaux autorégressifs. Cette hypothèse s'avère impossible à rejeter dans le cas des données de Québec (Tableau 2). Ainsi, le coefficient autorégressif estimé en utilisant l'approche spatio-temporelle est plus faible, du quart, que celui obtenu en utilisant l'approche spatiale. De plus, les variables liées aux caractéristiques environnementales sont plus fortement significatives, suggérant qu'une approche strictement spatiale diminue l'effet des variables sur la détermination du prix de vente.

INSÉRER TABLEAU 2 ICI

En bref, les applications sur les données de Québec font ressortir le danger potentiel lié à une surévaluation de l'effet spatial lorsque les données consistent en des unités spatiales récoltées en continu au fil du temps. Ce qui est difficile à établir c'est le rôle que pourrait avoir l'introduction de nouvelles variables explicatives basées sur la création de matrices de pondérations spatio-temporelles et susceptibles de capter un effet dynamique temporel spatialisé dans la détermination des valeurs marchandes immobilières. Cette hypothèse est abordée plus formellement avec les deux applications empiriques suivantes.

Le cas de Paris (France)

Une application sur la ville de Paris, incluant les départements de Paris, de Seine-Saint-Denis, de Val-de-Marne et de Hauts-de-Seine, afin de vérifier l'impact de l'ajout de variables dynamiques spatialisées sur la détermination des prix de vente des appartements est effectuée par Dubé et Legros (2011). Les données de transactions proviennent de la "Base d'Informations Économiques Notariales - BIEN" et couvrent la période 1990-2001. La base de données contient les informations sur le prix nominal de vente des appartements (en euros) et certaines informations sur les caractéristiques physiques des appartements. Par contre, aucune information sur les caractéristiques environnementales des résidences n'est disponible. Au total, ce sont 127 787 transactions qui sont disponibles.

Pour des raisons de temps de calcul, les analyses sont effectuées en prenant trois échantillons aléatoires de 10 000 observations afin de vérifier la robustesse des résultats obtenus. Cette approche permet également d'effectuer les analyses sans alourdir inutilement les calculs. Le but ultime de cette analyse n'étant pas de développer un

nouveau modèle complet de prévision des valeurs immobilières, mais plutôt de voir si l'ajout de variables dynamiques, mesurant la moyenne des prix de vente ayant eu cours dans un certain rayon avant une transaction donnée, a un effet significatif sur la détermination des prix de vente. Les variables dynamiques peuvent être vues comme des effets de pairs ou encore l'effet des comparables largement utilisées dans la pratique professionnelle pour établir le prix de vente demandé (Des Rosiers *et al.*, 2011).

L'analyse cherche également à déterminer si l'ajout de ces variables a un effet important sur la diminution de l'autocorrélation spatiale des résidus puisque les nouvelles variables possèdent indubitablement une dimension spatiale. Le premier constat de l'analyse est que l'ajout des variables dynamiques (effets de pairs) réduit sensiblement la mesure du degré de dépendance spatiale entre les termes d'erreurs (Tableau 3). Par contre, le seul ajout des variables dynamiques spatialisées ne suffit pas à éliminer le patron global de dépendance. Néanmoins, l'ajout de deux variables dynamiques a pour effet de diminuer de moitié, ou presque, le patron.

INSÉRER TABLEAU 3 ICI

Il est donc nécessaire d'introduire une spécification permettant de contrôler la corrélation spatiale résiduelle. La correction classique via une spécification autorégressive du terme d'erreur permet aussi de générer un modèle spatio-temporel autorégressif : introduction de retards dynamiques de la variable dépendante et contrôle de l'effet multidirectionnel spatial via une spécification latente. Les résultats suggèrent que les transactions ayant cours deux trimestres avant expliquent environ 15% du prix de vente, montrant ainsi l'importance marquée de cette variable dans la détermination des valeurs immobilières (Tableau 4). Par contre, l'effet temporel est relativement concentré puisque l'ajout d'une variable prenant en compte les transactions survenant deux à quatre trimestres avant est significatif, mais l'ampleur du coefficient est nettement plus faible, expliquant moins d'un pourcent du prix de vente.

INSÉRER TABLEAU 4 ICI

En bref, l'effet des comparables est statistiquement significatif, mais ne permet pas, à lui seul, d'occulter l'importance des caractéristiques individuelles des biens immobiliers. Par contre, il est légitime de vérifier si cette conclusion peut se généraliser à d'autres cas particuliers. Une application sur les données américaines permet de vérifier, en partie, la robustesse des résultats.

Le cas de Lucas County (Ohio, USA)

Les données de transactions résidentielles unifamiliales du comté de Lucas en Ohio (États-Unis) proviennent d'une application empirique menée par LeSage et Pace (2004) et sont reprises pour certains exemples dans le récent manuel d'économétrie de LeSage et Pace (2009). La base de données est accessible dans la librairie de MatLab et procure des informations sur le prix de vente nominal (en dollars américains), la superficie du terrain (en pieds carrés), la superficie habitable (en pieds carrés) et plusieurs caractéristiques physiques telles que le nombre de pièces, le nombre de chambres, le

nombre de salles de bains, le type de résidence et la présence d'un garage. L'information sur l'année de construction permet également d'obtenir l'âge de la résidence en années. Au total, c'est plus de 25 357 transactions qui sont disponibles pour effectuer l'analyse empirique.

Les résultats suggèrent que le développement d'une matrice de pondérations spatio-temporelles pour évaluer le degré de dépendance spatiale entre les résidus des modèles diminue l'importance et la significativité de la détection si l'on compare avec une spécification basée uniquement sur la matrice de pondérations spatiales (Tableau 5). En fait, ce problème montre que le fait de négliger la dimension temporelle des données peut mener à une surestimation du problème de dépendance spatiale et confirme à nouveau la première hypothèse.

INSÉRER TABLEAU 5 ICI

Les résultats d'estimation montrent que les variables résumant les prix de vente moyens du voisinage sont significatives sur la détermination des valeurs résidentielles, autant avec l'approche par moindres carrés ordinaire (MCO) que celle corrigeant pour le problème d'autocorrélation spatiale des résidus dans un contexte spatio-temporel (Tableau 6). Les coefficients et leur significativité confirment ainsi la troisième hypothèse et les résultats obtenus pour Paris. Or, malgré la significativité des effets dynamiques spatialisés, ces variables seules n'arrivent pas à contrôler pour la totalité du processus de dépendance spatiale noté entre les résidus du modèle (modèle MCO). Il est donc nécessaire de corriger le modèle, un cas particulier du modèle spatial autorégressif (SAR), mais dont le retard spatial de la variable dépendante est également situé dans le temps. Une spécification autorégressive des termes d'erreurs est donc nécessaire pour tenir compte des dimensions spatiales et temporelles.

L'ampleur des coefficients corrigés pour la dépendance spatiale entre les résidus suggère une certaine concentration temporelle de l'effet spatial (effet dynamique direct) et ce, peu importe l'éloignement spatial (colonne 2). Tout comme précédemment, la spécification strictement spatiale de la matrice de pondérations montre que l'effet autorégressif est surévalué, confirmant ainsi la seconde hypothèse. En fait, dans le contexte actuel, l'ampleur du coefficient autorégressif spatial estimé avec une matrice de pondérations strictement spatiales fait état de la présence d'une racine unitaire, introduisant ainsi la possibilité de résultats fallacieux (colonne 3). S'ajoute à ce problème, la fait que la surévaluation de l'effet spatial diminue la significativité des effets dynamiques liés aux prix de vente moyens des résidences (comparables) lors du mois précédent.

INSÉRER TABLEAU 6 ICI

En bref, les résultats confirment, encore une fois, les hypothèses soulevées, à savoir que la prise en compte des effets spatiaux et temporels diminue l'ampleur de l'autocorrélation spatiale estimée, autant sur les tests de détections que sur l'ampleur des coefficients autorégressifs, et que l'ajout de variables dynamiques spatialisées joue un rôle significatif sur le processus de détermination des valeurs marchandes immobilières résidentielles. Ainsi, il est important de tenir compte simultanément des deux dimensions

et de leurs effets lorsque les données utilisées font état d'accumulation d'observations spatiales au fil du temps. Le développement de matrices de pondérations spatio-temporelles s'avère donc une solution intéressante dans ces situations.

CONCLUSION

L'article propose une présentation de la caractérisation particulière des données spatiales individuelles empilées dans le temps et suggère une approche simple permettant de tenir compte des dimensions spatiales et temporelles dans la construction d'une matrice de pondérations appropriées pour ce type de données. Puisque la structure des données est clairement différente de la structure des données de panel et que les approches spatiales développées en données de panel sont difficilement transposables aux données spatiales empilées dans le temps, il est nécessaire d'établir une façon de traiter ces données. L'application des techniques et méthodes strictement spatiales dans un contexte où les données spatiales comportent également une dimension temporelle s'avère non-appropriée. En fait, l'utilisation des approches spatiales dans un contexte spatio-temporel a pour effet de surévaluer la mesure du degré de dépendance spatiale entre les observations en plus de générer une surévaluation des effets spatiaux dans les modèles statistiques.

À partir d'applications empiriques pour Québec, Paris et Lucas County (USA), l'article propose de tester trois hypothèses. La première est que le test de détection d'autocorrélation spatiale est systématiquement biaisé vers le rejet de l'hypothèse nulle (absence d'autocorrélation spatiale) lorsque la matrice de pondérations est strictement spécifiée dans un contexte spatial et ignore, par le fait même, la dimension temps. La seconde hypothèse est que l'utilisation de matrices de pondérations strictement spatiales dans un contexte spatio-temporel surestime l'ampleur des coefficients autorégressifs. Finalement, la troisième hypothèse est que l'ajout de variables spatialement et temporellement décalées permet de tenir compte d'effets de paires (effets de comparables) qui s'avèrent jouer un rôle significatif dans la détermination des valeurs immobilières résidentielles.

Les résultats suggèrent que ces hypothèses ne peuvent pas être rejetées, et ce, pour les trois différentes applications. Non seulement la non-prise en compte de la dimension temporelle biaise à la hausse la mesure de détection des patrons de dépendances spatiales entre les résidus, mais elle influence également l'ampleur des coefficients autorégressifs. Les coefficients estimés avec une matrice de pondérations strictement spatiales procurent toujours des coefficients supérieurs à ceux obtenus utilisant une matrice de pondérations spatio-temporelles. De plus, la création de matrices de pondérations spatio-temporelles permet de définir un ensemble de nouvelles variables qui jouent un rôle significatif dans la détermination du prix de vente immobilier. La création de variables d'effets de paires ou de comparables s'avère potentiellement important dans l'explication du prix de vente, malgré qu'elles ne réussissent pas à contrôler en totalité pour l'effet de dépendance spatiale dans les exemples retenus.

Au final, la création de matrices de pondérations spatio-temporelles, issues d'une combinaison de matrices de pondérations spatiales et temporelles, fait en sorte d'ajuster les méthodes et tests économétriques spatiaux existants sans pour autant avoir à retravailler les propriétés de ceux-ci. L'ajustement des matrices de pondérations dans un

contexte spatio-temporel s'avère donc important et essentiel, surtout si on tient compte de la possibilité d'obtenir des coefficients autorégressifs spatiaux montrant un comportement de racine unitaire.

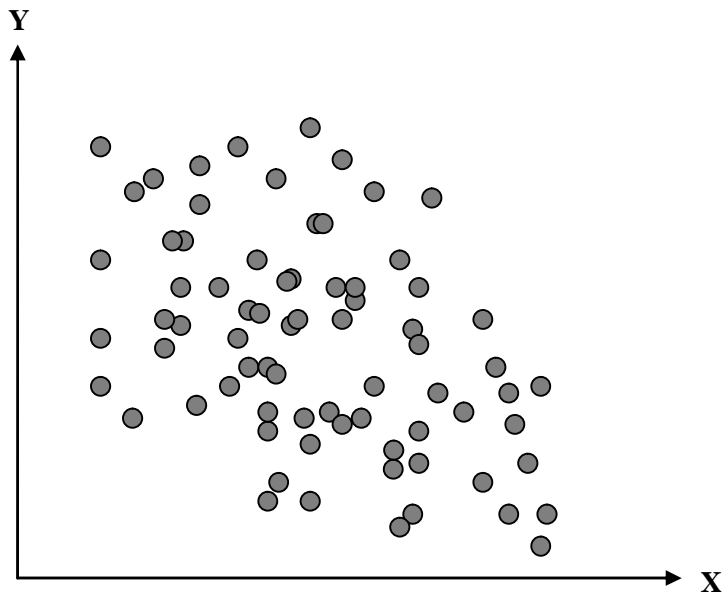
BIBLIOGRAPHIE

- ABRAHAM, J.M. ET SCHAUMAN, W.S. (1991). New Evidence on Home Prices from Freddie Mac Repeat Sales, *AREUEA Journal*, 19: 333-352.
- ANSELIN, L. (2010). Thirty Years of Spatial Econometrics, *Papers in Regional Science*, 89(1): 3-25.
- BALTAGI, B.H. ET LI, D. (2004). Prediction in the Panel Data Model with Spatial Correlation, dans ANSELIN, L., FLORAX, R. ET REY, S.J. (eds). *Advances in Spatial Econometrics*, Springer: 283-296.
- BALTAGI, B.H., SONG, S.H., JUNG, B.C. ET KOH, W. (2007). Testing for Serial Correlation, Spatial Autocorrelation and Random Effects Using Panel Data. *Journal of Econometrics*, 140: 5-51.
- BALTAGI, B.H., SONG, S.H. ET KOH, W. (2003). Testing Panel Data Regression Models with Spatial Error Correlation. *Journal of Econometrics*, 117:123-150.
- BHATTACHARJEE, A. ET HOLLY, S. (2011). Understanding Interactions in Social Networks and Committees, *Proceedings of the Spatial Econometric Association Annual Meetings*, 43 p.
- BOOTS, B. ET DUFOURNAUD, C. (1994). A Programming Approach to Minimizing and Maximizing Spatial Autocorrelation, *Geographical Analysis*, 26(1): 54-66.
- CASE, K.E. ET SHILLER, R.J. (1989). The Efficiency of the Market for Single-Family Homes, *American Economic Review*, 79: 125-137.
- CHASCO, C.Y. ET LOPEZ, F.A.H. (2008). Is Spatial Dependence an Instantaneous Effect? Some Evidence in Economic Series of Spanish Provinces, *Estadística Española*, 50(167): 100-118.
- CLAPP, J.M., GIACOTTO, C. ET TIRTIROGLU, D. (1991). Housing Price Indices Based on All Transactions Compared to Repeat Subsamples, *AREUEA Journal*, 19: 270-285.
- DES ROSIERS, F., DUBÉ, J. ET THÉRIAULT, M. (2011). Do Peer Effects Shape Property Values?, *Journal of Property Investment and Finance*, 29(4/5): 510-528.
- DES ROSIERS, F., DUBÉ, J. ET THÉRIAULT, M. (2000). Sorting out Access and Neighborhood Factors in Hedonic Price Modelling, *Journal of Property Investment and Finance*, 18(3): 291-315.
- DUBÉ, J. ET LEGROS, D. (SOUS PRESSE). A Spatio-temporal Measure of Spatial Dependence: An Example Using Real Estate Data, *Papers in Regional Science*, disponible en ligne.
- DUBÉ, J. ET LEGROS, D. (SOUS PRESSE). Dealing with Spatial Data Pooled over Time in Statistical Models, *Letters in Spatial and Resource Sciences*, disponible en ligne.
- DUBÉ, J. ET LEGROS, D. (2011). Using a Spatio-temporal Weights Matrix in the Development of a Spatio-temporal Autoregressive (STAR) Model, *Proceedings of the Spatial Econometric Association Annual Meetings*, 31 p.
- DUBÉ, J., BAUMONT, C. ET LEGROS, D. (à venir). Matrice de pondérations et contexte spatio-temporel en économétrie spatiale, *Revue canadienne de science régionale*, soumis.
- DUBÉ, J., BAUMONT, C. ET LEGROS, D. (2011b). *Utilisation des matrices de pondérations en économétrie spatiale : Proposition dans un contexte spatio-temporel*,

- Documents de travail du Laboratoire d'Économie et de Gestion, Université de Bourgogne, e2011-01.
- DUBÉ, J., DES ROSIERS, F. ET THÉRIAULT, M. (2011c). Impact de la segmentation spatiale sur le choix de la forme fonctionnelle pour la modélisation hédonique, *Revue D'Économie Régionale et Urbaine*, 1 : 9-37.
- DUBÉ, J., DES ROSIERS, F., THÉRIAULT, M. ET DIB, P. (2011a). Economic Impact of a Supply Change in Mass Transit in Urban Areas: A Canadian Example, *Transportation Research Part A*, 45(1): 46-62.
- ELHORST, J.P. (2003). Specification and Estimation of Spatial Panel Data Models, *International Regional Science Review*, 26: 244-268.
- FINGLETON, B. (2009). Spatial Autocorrelation, *Geographical Analysis*, 41: 385-391.
- GELFLAND, A.E., GHOSH, S.K., KNIGHT, J.R. ET SIRMANS, C.F. (1998). Spatio-Temporal Modeling of Residential Sales Data, *Journal of Business and Economics Statistics*, 16(3): 312-321.
- GETIS, A. (2009). Spatial Weights matrices, *Geographical Analysis*, 41 (4): 404-410.
- GETIS, A. ET ALDSTADT, J. (2004). Constructing the Spatial Weights Matrix Using a Local Statistics, *Geographical Analysis*, 36: 90-104.
- GIACOMINI, R. ET GRANGER, C.W.J. (2004). Aggregation of Space-Time Processes, *Journal of Econometrics*, 118: 7-26.
- GOODCHILD, M.F. (2009). What Problem? Spatial Autocorrelation and Geographic Information Science, *Geographical Analysis*, 41: 411-417.
- HAINING, R.P. (2009). Spatial Autocorrelation and the Quantitative Revolution, *Geographical Analysis*, 41: 364-374.
- HUANG, B., WU, B. ET BARRY, M. (2010). Geographically and Temporally Weighted Regression for Modeling Spatio-temporal Variation in House Prices, *International Journal of Geographical Information Science*, 24(3): 383-401.
- KAPOOR, M., KELEJIAN, H.H. ET PRUCHA, I.R. (2007). Panel Data Models with Spatially Correlated Error Components, *Journal of Econometrics*, 140: 97-130.
- LEE, L.-F. ET YU, J. (2010). Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects, *Journal of Econometrics*, 154 : 165-185.
- LESAGE, J. (1999). *Applied Econometrics using MATLAB*, www.spatial-econometrics.com.
- LESAGE, J. ET PACE, R.K. (2009). *Introduction to Spatial Econometrics*, Taylor & Francis Group.
- LESAGE, J. ET PACE, R.K. (2004). Models for Spatially Dependent Missing Data, *Journal of Real Estate Finance and Economics*, 29(2): 233-254.
- NAPPI-CHOLET, I. ET MAURY, T.-P. (2011). A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market, *Journal of Regional Science*, 51(4): 732-750.
- NAPPI-CHOLET, I. ET MAURY, T.-P. (2009). A Spatiotemporal Autoregressive Price Index for the Paris Office Property Market, *Real Estate Economics*, 37(2): 305-340.
- PACE, R.K., BARRY, R., CLAPP, J.M. ET RODRIGUEZ, M. (1998). Spatiotemporal Autoregressive Models of Neighborhood Effects, *Journal of Real Estate Finance and Economics*, 17(1): 14-33.

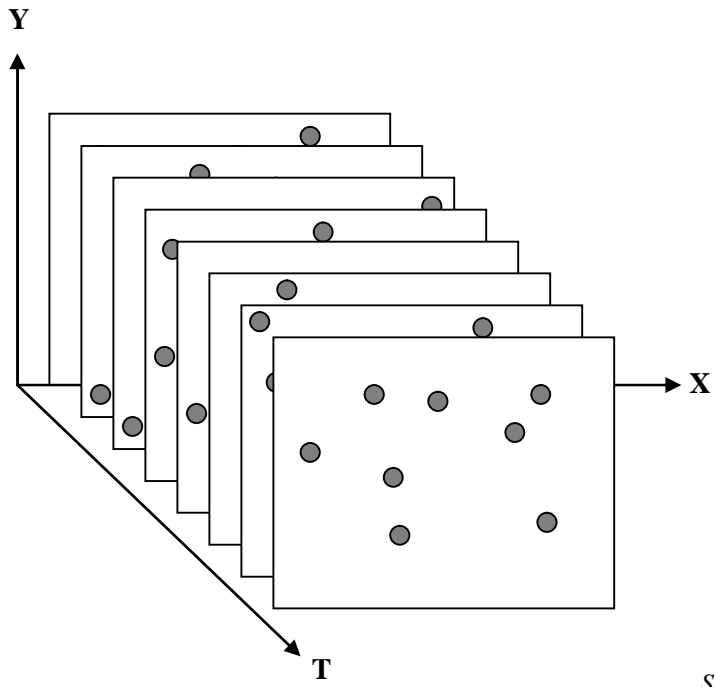
- PACE, R.K., BARRY, R., GILLEY, W. ET SIRMANS, C.F. (2000). A Method for Spatio-Temporal Forecasting with an Application to Real Estate Prices, *International Journal of Forecasting*, 16(2): 229-246.
- PARENT, O. ET LESAGE, J. (2011). A Space-time Filter for Panel Data Models Containing Random Effects, *Computational Statistics & Data Analysis*, 55: 475-490.
- SMITH, T.E. ET WU, P. (2009). A Spatio-temporal Model of Housing Prices Based on Individual Sales Transactions Over Time, *Journal of Geographical Systems*, 11(4) : 333-355.
- SUN, H., TU, Y. ET YU, S.-M. (2005). A Spatio-Temporal Autoregressive Model for Multi-Unit Residential Market Analysis, *Journal of Real Estate Finance and Economics*, 31(2): 155-187.
- TOBLER, W. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography*, 46: 234-240.
- TU, Y., YU, S.-M. ET SUN, H. (2004). Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach, *Real Estate Economics*, 32(2): 297-328.
- VOISIN, M., DUBÉ, J., DES ROSIERS, F. ET THÉRIAULT, M. (2010). Les découpages administratifs sont-ils pertinents en analyse immobilière? *Le cas de Québec, Cahiers de géographie du Québec*, 54(2) : 249-274.
- YU, J., DE JONG, R. ET LEE, J.-F. (2008). Quasi-maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T are Large, *Journal of Econometrics*, 146: 118-134.

Figure 1
Structuration observations permettant de calculer les distances et d'établir les relations



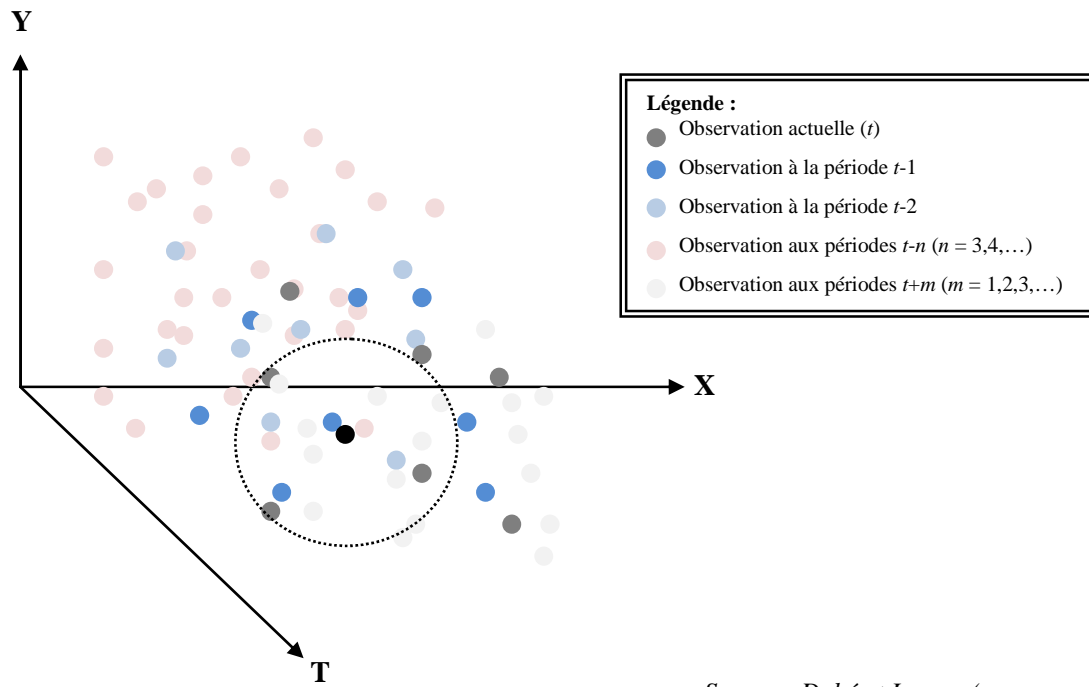
Source : Dubé et Legros (sous presse)

Figure 2
Structuration temporelle des observations spatiales individuelles



Source : Dubé et Legros (sous presse)

Figure 3
Relations spatio-temporelles pour une observation donnée



Source : Dubé et Legros (sous presse)

Tableau 1

Synthèse des tests de détection de l'autocorrélation spatiale selon le type de matrice de pondérations spatiales utilisé - Québec, 1990-1996.

| | Caractéristiques intrinsèques | |
|-------------------------------------------------------------------------------------|-----------------------------------------------|----------------------|
| | Indice <i>I</i> de Moran | Statistique <i>t</i> |
| Matrice spatiale 1 500 m. ($\mathbf{W}=\mathbf{S}_1$) | 0,3044 | 21,34 |
| Matrice spatio-temporelle 1 500 m. ($\mathbf{W}=\mathbf{S}_1 \odot \mathbf{T}_1$) | 0,1706 | 9,88 |
| Matrice spatiale 3 000 m. ($\mathbf{W}=\mathbf{S}_2$) | 0,2865 | 21,21 |
| Matrice spatio-temporelle 3 000 m. ($\mathbf{W}=\mathbf{S}_2 \odot \mathbf{T}_1$) | 0,1708 | 9,98 |
| | Caractéristiques intrinsèques et extrinsèques | |
| | Indice <i>I</i> de Moran | Statistique <i>t</i> |
| Matrice spatiale 1 500 m. ($\mathbf{W}=\mathbf{S}_1$) | 0,1736 | 12,90 |
| Matrice spatio-temporelle 1 500 m. ($\mathbf{W}=\mathbf{S}_1 \odot \mathbf{T}_1$) | 0,0809 | 5,09 |
| Matrice spatiale 3 000 m. ($\mathbf{W}=\mathbf{S}_2$) | 0,1609 | 12,78 |
| Matrice spatio-temporelle 3 000 m. ($\mathbf{W}=\mathbf{S}_2 \odot \mathbf{T}_1$) | 0,0819 | 4,79 |

Note : Calculs effectués sur 10% de la base de données (1 569 observations).

Tableau 2

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques et extrinsèques), Québec, 1990-1996.

| | Modèle MCO | | Modèle SEM ($W = S$) | | Modèle SEM ($W = S \odot T$) | |
|-----------------------------------------------|-------------|-----------------|------------------------|-----------------|--------------------------------|-----------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| Constante | 8,9648 | 222,33 | 8,9956 | 1456,70 | 8,9700 | 1472,12 |
| Superficie habitable (log - m. ²) | 0,4339 | 58,57 | 0,4261 | 64,02 | 0,4290 | 68,58 |
| Taille du terrain (log - m. ²) | 0,0752 | 19,97 | 0,0802 | 22,63 | 0,0830 | 25,17 |
| Âge du bâtiment (log) | -0,1033 | -66,29 | -0,1033 | -57,71 | -0,1079 | -64,45 |
| Cottage | -0,0549 | -12,56 | -0,0538 | -12,23 | -0,0545 | -13,15 |
| Atachée | -0,1548 | -23,34 | -0,1421 | -17,91 | -0,1444 | -21,04 |
| Indice de qualité | 0,1148 | 22,67 | 0,1165 | 22,64 | 0,1175 | 23,72 |
| Nb. salles de bains | 0,0416 | 14,45 | 0,0390 | 13,75 | 0,0405 | 14,69 |
| Sous-sol aménagé | 0,0435 | 14,75 | 0,0413 | 14,43 | 0,0435 | 15,36 |
| Façade en brique (51% et +) | 0,0180 | 5,72 | 0,0145 | 4,61 | 0,0195 | 6,37 |
| Nb. de foyers | 0,0433 | 14,16 | 0,0399 | 13,45 | 0,0428 | 14,76 |
| Planchers de qualité supérieure | 0,0202 | 6,98 | 0,0229 | 8,04 | 0,0208 | 7,44 |
| Escalier en bois massif | 0,0401 | 10,20 | 0,0381 | 9,71 | 0,0391 | 10,30 |
| Comptoirs de cuisine de qualité supérieure | 0,0190 | 0,92 | 0,0208 | 1,05 | 0,0221 | 1,12 |
| Luminosité inférieure | -0,0193 | -4,78 | -0,0145 | -3,56 | -0,0166 | -4,23 |
| Toit cathédrale | 0,0331 | 9,31 | 0,0297 | 8,23 | 0,0285 | 8,25 |
| Balayeuse centrale | 0,0431 | 9,61 | 0,0379 | 8,65 | 0,0424 | 9,89 |
| Garage simple attaché | 0,1123 | 17,86 | 0,0948 | 14,78 | 0,1010 | 16,58 |
| Garage double attaché | 0,0920 | 10,09 | 0,0879 | 10,04 | 0,0873 | 10,03 |
| Garage simple détaché | 0,0320 | 6,94 | 0,0293 | 6,42 | 0,0309 | 6,90 |
| Garage double détaché | 0,0586 | 9,30 | 0,0637 | 10,38 | 0,0579 | 9,66 |

Tableau 2 (suite)

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques et extrinsèques), Québec, 1990-1996.

| | Modèle MCO | | Modèle SEM ($\mathbf{W} = \mathbf{S}$) | | Modèle SEM ($\mathbf{W} = \mathbf{S} \odot \mathbf{T}$) | |
|-----------------------------------------------|------------------|-----------------|------------------------------------------|-----------------|-----------------------------------------------------------|-----------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| Terrasse | 0,0329 | 3,33 | 0,0207 | 2,11 | 0,0260 | 2,75 |
| Piscine creusée | 0,0908 | 14,12 | 0,0916 | 14,54 | 0,0885 | 14,41 |
| Desservi par l'acqueduc | 0,1238 | 10,94 | 0,0954 | 8,64 | 0,1180 | 11,46 |
| Taux de taxe local | -0,0756 | -16,52 | -0,0737 | -13,71 | -0,0784 | -15,11 |
| Année 1990 | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Année 1991 | 0,0232 | 5,65 | 0,0228 | 4,33 | 0,0240 | 4,33 |
| Année 1992 | -- | -- | -- | -- | -- | -- |
| Année 1993 | 0,0632 | 15,21 | 0,0680 | 18,06 | 0,0640 | 11,51 |
| Année 1994 | 0,0604 | 14,41 | 0,0682 | 17,88 | 0,0612 | 10,95 |
| Année 1995 | 0,0430 | 9,58 | 0,0513 | 12,57 | 0,0442 | 7,39 |
| Année 1996 | 0,0438 | 8,65 | 0,0491 | 10,52 | 0,0439 | 6,49 |
| Indice d'accessibilité régionale | 0,0535 | 25,15 | 0,0563 | 19,39 | 0,0563 | 21,82 |
| Indice d'accessibilité locale | 0,0335 | 17,95 | 0,0336 | 13,08 | 0,0345 | 15,17 |
| Nb familles monoparentales | -0,0059 | -4,42 | -0,0049 | -2,90 | -0,0047 | -3,23 |
| Revenu médian des ménages | 0,0080 | 4,70 | 0,0076 | 3,68 | 0,0065 | 3,62 |
| Pourcentage détenant un diplôme universitaire | 0,0058 | 37,33 | 0,0057 | 28,67 | 0,0056 | 32,22 |
| | -- | -- | 0,3880 | 87,49 | 0,2880 | 48,82 |
| Nombre d'observations | 15 729 | | 15 729 | | 15 729 | |
| R ² | 0,7666 | | 0,8000 | | 0,7833 | |
| R ² ajusté | 0,7661 | | 0,7996 | | 0,7829 | |
| Log-vraisemblance | -- | | 13 179,96 | | 12 738,71 | |

Note : Matrice de pondérations spatiales avec seuil fixé à 3 000 mètres et matrices de pondérations temporelles pour le trimestre en cours

Tableau 3

Synthèse des tests de détection de l'autocorrélation selon la spécification du modèle avec une matrice de pondérations spatio-temporelles ($\mathbf{W} = \mathbf{S} \odot \mathbf{T}$), Paris, 1991-2000.

| Modèle MCO | | |
|--------------------------------------|-------------------------------------|-----------------------------|
| | Indice <i>I</i> de Moran | Statistique <i>t</i> |
| Échantillon 1 | 0,1911 | 67,74 |
| Échantillon 2 | 0,1911 | 69,96 |
| Échantillon 3 | 0,1949 | 70,90 |
| Modèle MCO avec effets dynamiques | | |
| | Indice <i>I</i> de Moran | Statistique <i>t</i> |
| Échantillon 1 | 0,1036 | 36,79 |
| Échantillon 2 | 0,1136 | 40,35 |
| Échantillon 3 | 0,0926 | 32,78 |

Tableau 4

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques seulement), Paris, 1991-2000.

| | Échantillon 1 | | Échantillon 2 | | Échantillon 3 | |
|-----------------------------------------------------------------|------------------|---------------|------------------|---------------|------------------|---------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| Comparables - deux trimestres précédents (y_{t-1}) | 0,1521 | 56,73 | 0,1439 | 56,32 | 0,1967 | 56,98 |
| Comparables - deux à quatre trimestres précédents (y_{t-2}) | 0,0065 | 5,50 | 0,0079 | 6,82 | 0,0085 | 7,62 |
| Constant | 5,1098 | 295,99 | 5,239 | 310,87 | 4,6417 | 158,82 |
| Superficie habitable (en m ² - log) | 1,0542 | 133,41 | 1,0524 | 131,47 | 1,0579 | 138,30 |
| Présence d'un ascenseur | 0,0747 | 10,20 | 0,0706 | 9,57 | 0,0624 | 8,85 |
| Nombre de salle de bains (log) | 0,2285 | 15,00 | 0,2193 | 14,52 | 0,2331 | 15,81 |
| Présence d'une terrasse | 0,0953 | 6,46 | 0,1432 | 9,82 | 0,1092 | 7,73 |
| Présence d'un garage | 0,0501 | 6,78 | 0,0432 | 5,92 | 0,0421 | 5,92 |
| Chauffage collectif | 0,0289 | 2,36 | -0,0032 | -0,26 | 0,0053 | 0,47 |
| Construction avant 1850 | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Construction entre 1850 et 1913 | -0,0449 | -2,96 | -0,1025 | -6,43 | -0,0920 | -6,21 |
| Construction entre 1914 et 1947 | -0,0545 | -3,36 | -0,1043 | -6,18 | -0,0908 | -5,80 |
| Construction entre 1948 et 1969 | -0,0869 | -5,35 | -0,1356 | -8,07 | -0,1139 | -7,29 |
| Construction entre 1970 et 1980 | -0,0733 | -4,31 | -0,1168 | -6,60 | -0,1067 | -6,50 |
| Construction entre 1981 et 1991 | 0,0541 | 2,72 | 0,0043 | 0,21 | 0,0178 | 0,94 |
| Construction entre 1992 et 2000 | 0,2393 | 12,4 | 0,2097 | 10,56 | 0,2089 | 11,26 |
| Situé au rez-de-chaussée | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Situé au premier étage | 0,0573 | 4,75 | 0,0753 | 6,23 | 0,0606 | 5,23 |
| Situé au deuxième étage | 0,0812 | 6,71 | 0,0967 | 8,10 | 0,0905 | 7,83 |
| Situé au troisième étage | 0,0894 | 7,29 | 0,1146 | 9,42 | 0,0871 | 7,48 |
| Situé au quatrième étage | 0,0906 | 7,18 | 0,1165 | 9,15 | 0,0934 | 7,77 |
| Situé au cinquième étage ou plus | 0,0690 | 5,81 | 0,0903 | 7,67 | 0,0698 | 6,16 |

Tableau 4 (suite)

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques seulement), Paris, 1991-2000.

| | Échantillon 1 | | Échantillon 2 | | Échantillon 3 | |
|-----------------------|------------------|---------------|------------------|---------------|------------------|---------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| Seine-Saint-Denis | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Paris | 0,5829 | 36,34 | 0,5946 | 39,33 | 0,5593 | 37,90 |
| Hauts-de-Seine | 0,3896 | 23,23 | 0,4041 | 25,23 | 0,3505 | 22,68 |
| Val-de-Marne | 0,2061 | 11,64 | 0,1847 | 10,84 | 0,1849 | 11,28 |
| Vendu en 1990 | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Vendu en 1991 | -0,0198 | -1,00 | -0,0440 | -2,25 | -0,0418 | -2,19 |
| Vendu en 1992 | -0,0890 | -4,73 | -0,0947 | -5,20 | -0,0993 | -5,58 |
| Vendu en 1993 | -0,1687 | -9,05 | -0,1624 | -8,84 | -0,1785 | -9,82 |
| Vendu en 1994 | -0,1675 | -9,20 | -0,1828 | -10,23 | -0,1931 | -10,97 |
| Vendu en 1995 | -0,2159 | -11,52 | -0,2462 | -13,42 | -0,2505 | -13,70 |
| Vendu en 1996 | -0,3102 | -17,32 | -0,3052 | -17,23 | -0,3133 | -17,98 |
| Vendu en 1997 | -0,3193 | -17,56 | -0,3202 | -17,92 | -0,3485 | -19,62 |
| Vendu en 1998 | -0,3129 | -17,27 | -0,3134 | -17,53 | -0,3436 | -19,71 |
| Vendu en 1999 | -0,2533 | -14,24 | -0,2660 | -15,14 | -0,2831 | -16,47 |
| Vendu en 2000 | -0,1864 | -10,18 | -0,1857 | -10,3 | -0,1978 | -11,19 |
| Vendu en 2001 | -0,1281 | -6,64 | -0,1205 | -6,41 | -0,1132 | -6,09 |
| | 0,6290 | 54,59 | 0,6360 | 55,25 | 0,6070 | 45,16 |
| R ² | 0,8220 | | 0,8249 | | 0,8396 | |
| R ² ajusté | 0,8214 | | 0,8244 | | 0,8391 | |

Note : Matrice de pondérations spatiales utilisant les 15 plus proches voisins et matrices de pondérations temporelles pour le trimestre en cours et l'année précédente.

Tableau 5

Synthèse des tests de détection de l'autocorrélation spatiale selon la spécification du modèle et la matrice de pondérations utilisée, Lucas County, Ohio (USA), 1993-1998.

| | Modèles MCO | |
|----------------------------------------|------------------------------------------------------------------------------------------|-----------------------------|
| | Indice <i>I</i> de Moran (matrice $W=$ $S \odot T$) | Statistique <i>t</i> |
| Caractéristiques intrinsèques | 0,2234 | 77,50 |
| Effets dynamiques directs | 0,2180 | 75,70 |
| Effets dynamiques directs et indirects | 0,2152 | 74,72 |

| | Modèles MCO | |
|----------------------------------------|--------------------------------------------------------|-----------------------------|
| | Indice <i>I</i> de Moran (matrice <i>S</i>) | Statistique <i>t</i> |
| Caractéristiques intrinsèques | 0,3240 | 36,79 |
| Effets dynamiques directs | 0,3084 | 321,56 |
| Effets dynamiques directs et indirects | 0,3052 | 318,44 |

Note : la matrice de pondérations spatiales est basée sur un rayon de 500 mètres et la matrice de pondérations temporelles sur les ventes ayant lieu durant le même mois.

Tableau 6

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques seulement), Lucas County, Ohio (USA), 1993-1998.

| | Modèle MCO | | Modèle SEM ($\mathbf{W} = \mathbf{S} \odot \mathbf{T}$) | | Modèle SEM ($\mathbf{W} = \mathbf{S}$) | |
|--------------------------------------------------|------------------|-----------------|-----------------------------------------------------------|-----------------|------------------------------------------|-----------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| Constante | 3,1131 | 41,29 | 3,6054 | 49,94 | 3,7151 | 5,86 |
| Log de l'âge (en années) | 0,4603 | 60,12 | 0,0341 | 43,33 | 0,0327 | 40,92 |
| Âge (en années) | -0,0219 | -103,91 | -0,0164 | -74,33 | -0,0133 | -59,92 |
| Superficie habitable (en pi^2 - log) | 0,6191 | 61,96 | 0,6473 | 95,01 | 0,6306 | 77,64 |
| Superficie du terrain (en pi^2 - log) | 0,2138 | 54,68 | 0,1714 | 72,72 | 0,1140 | 29,51 |
| Nombre de salles de bain | 0,1006 | 14,19 | 0,0939 | 16,40 | 0,0716 | 13,22 |
| Présence d'un garage | 0,3356 | 44,51 | 0,2521 | 38,36 | 0,1815 | 31,67 |
| Maison à deux étages ou plus | 0,1930 | 31,94 | 0,1358 | 26,23 | 0,0848 | 17,65 |
| Vente en 1993 | <i>référence</i> | | <i>référence</i> | | <i>référence</i> | |
| Vente en 1994 | 0,0049 | 0,51 | -0,0066 | -0,28 | 0,0303 | 4,30 |
| Vente en 1995 | 0,0458 | 4,88 | 0,0113 | 0,49 | 0,0749 | 10,80 |
| Vente en 1996 | 0,0790 | 8,63 | 0,0358 | 1,59 | 0,0991 | 14,64 |
| Vente en 1997 | 0,1132 | 12,45 | 0,0780 | 3,49 | 0,1415 | 21,00 |
| Vente en 1998 | 0,1603 | 17,21 | 0,1314 | 5,72 | 0,1922 | 27,91 |
| Effet dynamique un mois avant* (1 000 mètres) | 0,0225 | 12,09 | 0,0088 | 4,57 | -0,0002 | -0,15 |
| Effet dynamique deux mois avant* (1 000 mètres) | 0,0190 | 11,16 | 0,0157 | 7,99 | 0,0027 | 2,11 |
| Effet dynamique trois mois avant* (1 000 mètres) | 0,0116 | 7,61 | 0,0106 | 5,52 | 0,0012 | 1,10 |
| Effet dynamique un mois avant (2 000 mètres) | -0,0164 | -7,19 | -0,0017 | -0,71 | 0,0035 | 1,99 |
| | -- | -- | 0,6830 | 171,78 | 0,9900 | 133,91 |

Tableau 5 (suite)

Résultats d'estimation du modèle de prix hédonique (caractéristiques intrinsèques seulement), Lucas County, Ohio (USA), 1993-1998.

| | Modèle MCO | | Modèle SEM ($\mathbf{W} = \mathbf{S} \odot \mathbf{T}$) | | Modèle SEM ($\mathbf{W} = \mathbf{S}$) | |
|-----------------------|-------------|-----------------|-----------------------------------------------------------|-----------------|------------------------------------------|-----------------|
| | Coefficient | Statistique t | Coefficient | Statistique t | Coefficient | Statistique t |
| R ² | 0,7392 | | 0,7960 | | 0,8584 | |
| R ² ajusté | 0,7390 | | 0,7959 | | 0,8583 | |
| Log-vraisemblance | --- | | -1 055,01 | | 4 124,72 | |
| Nombre d'observations | 25 357 | | 25 357 | | 25 357 | |

Légende :

* : Transactions ayant cours dans un rayon de 1 000 mètres

: transactions ayant cours dans un rayon de 2 000 mètres

Note : Matrice de pondérations spatiales basée sur un rayon de 500 mètres et matrices de pondérations temporelles pour le mois en cours pour l'effet autorégressif spatial .