Treatment of Spatial Autocorrelation in Geocoded Crime Data

Krista Collins, Colin Babyak, Joanne Moloney
Household Survey Methods Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6

Abstract

In modelling spatial data, when measurements at one location are influenced by the measurements at neighbouring or nearby locations, we say that spatial autocorrelation is present. This violates the assumption of statistically independent observations commonly applied in standard regression analysis. We first examine the basic theory and methods used to analyse spatial lattice data (i.e., data aggregated to regions as opposed to observations at discrete points). Next, we focus on the two basic forms of simultaneous autoregressive models, termed the spatial lag and spatial error models. Emphasis is placed on how the spatial effects are incorporated into the model. Finally, we implement the spatial models in a study of urban neighbourhood crime rates in the Canadian city of Montreal. We examine the differences in the models obtained with and without modelling the spatial effects.
Keywords: Spatial autocorrelation, Simultaneous autoregressive, Spatial lag model, Spatial error model, Crime data

1. Introduction

When examining cross-sectional data there is often reason to suspect that observations from nearby locations will be more similar (spatial attraction) or less similar (spatial repulsion) than observations that are further apart. This may occur because observations are related to the characteristics of the location and nearby observations are affected by the same factors or because the variable of interest may have a direct impact on its value at nearby locations.

For example, the distribution of crime in city neighbourhoods can be affected by the properties of the location and by social interactions. The crime rate tends to be higher in regions with a higher rate of unemployment and a lower median household income (Morenoff et al. 2001; Fitzgerald et al. 2004). These socio-economic conditions are usually spatially clustered within cities, which may cause a spatial dependence in the crime rate. Social interactions, which are not confined to neighbourhood boundaries, may also cause the crime rate in adjacent regions to be related. This may take the form of spatial diffusion with the spreading of crime or copy-cat behaviour (Puech 2004). Copy-cat behaviour occurs when a person decides to commit a crime because their perceived probability of punishment is reduced by the high volume of crime in surrounding locations and the limitation on the possible number of arrests by the police.

If the spatial dependence in the data is not completely accounted for in the regression model it will lead to spatial autocorrelation in the residuals. Spatial autocorrelation, identified by the non-zero covariance between a pair of observations that are related in space, can cause inefficient estimation of the standard regression model parameters, and inaccuracy of the sample variance and significance tests. It may be caused by spatial dependence that is not adequately explained by the explanatory variables, systematic measurement errors or a mismatch between the spatial scale used to measure the variable and the scale at which it actually occurs (Anselin and Bera 1998).

There are a variety of spatial analysis techniques for the different forms of spatial data and the desired type of analysis. This paper focuses on the spatial models used to analyse data that has been aggregated to regions, commonly referred to as lattice or regional data. We first present a brief overview of the types of spatial models for lattice data, with attention on the spatial lag and spatial error models. The use of these two types of models is then illustrated by comparing the results of a standard linear regression and a spatial regression that accounts for autocorrelation of the neighbourhood crime rates on the Island of Montreal.

2. Spatial Analysis Models for Lattice Data

Lattice data consists of observations that are aggregated to regions or geographic areas as opposed to representing individual data points. The observed data in a given region is considered representative of that entire geographic area. The regions span the entire study area such that there is no possibility of interpolation between two adjacent regions; rather the analysis focuses on modelling data from the observed regions as they are defined.

The two principal approaches for modelling lattice data are referred to as simultaneous autoregressive (SAR) and conditional autoregressive (CAR) models. Both approaches relate the data at a given location to a linear combination of neighbouring values, which represents the autoregressive structure. In the simultaneous approach the autoregressive terms are based on the average value from all neighbouring locations. The value at a given location is specified in conjunction with the values at all other locations, thus indicating the autoregression occurs simultaneously for each region. In the conditional approach the value at a given location is specified conditionally on the values from neighbouring regions. Thus the values from neighbouring locations are assumed to be known and the conditional distribution of the variable is modelled (Anselin 2002). This paper addresses only the simultaneous approach, which is the method preferred in the literature for analysing crime data (Anselin et al. 2000, Baller et al. 2001, Anselin 2002).

2.1 Taxonomy of Simultaneous Autoregressive Models

The autoregressive structure of SAR models requires a definition of what constitutes neighbouring locations for lattice data. There are three main ways of defining neighbouring regions: contiguity structure, which is based on the configuration of the neighbourhoods; distance band, which includes all neighbourhoods within a specified distance; or k-nearest neighbours which include the number of specified regions which are the closest (Dubin 1992). Figure 1 illustrates some of the different possible neighbourhood structures.

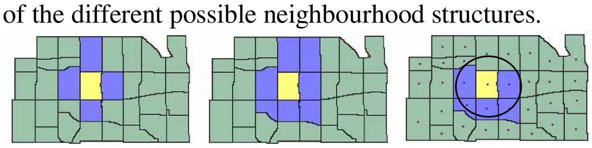

Figure 1: Different methods for defining neighbouring regions for the selected region in yellow. A) Rook contiguity – includes locations with a common border, B) Queen contiguity – includes regions with a common border or vertex, and C) Distance – includes all regions whose centroid is within the specified distance band.

The neighbourhood structure is represented mathematically by a spatial weights matrix, W. This is a binary $N \times N$ matrix, where the off-diagonal elements, W_{ij} for $i \neq j$, equal one if location j is a neighbour of location i and zero otherwise. The spatial weights matrix thus reflects the potential interaction of neighbouring locations and rules out spatial dependence for non-neighbouring locations. By convention, zeros are placed on the diagonal elements of the matrix, W_{ii}, indicating that a location cannot be a neighbour of itself. In the final weights matrix that is used in the model the elements in each row are standardized so that they sum to one.

The general form of the simultaneous autoregressive model is:

$$\begin{aligned} y &= X\beta + \rho W_1 y + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \end{aligned},$$

where y is the $N \times 1$ vector of the dependent variable, X is the $N \times k$ matrix of independent variables, β is the $k \times 1$ vector of regression coefficients, ρ and λ are scalar spatial autoregressive parameters, W_1 and W_2 are spatial weights matrices and ε and u are the $N \times 1$ vectors of error terms (Anselin 1988). For this paper the residuals, u, are assumed to be normal with a homoskedastic variance, although it is possible to introduce heterogeneity in the model by relaxing this assumption.

There are two spatial terms in the general SAR model: an autoregressive term $\rho W_1 y$ and the spatial dependence in the error terms, $\lambda W_2 \varepsilon$. The two spatial weights matrices, W_1 and W_2, allow for a different spatial structure in these two processes (Anselin 1988). With the appropriate row standardization of the spatial weights matrices, the terms $W_1 y$ and $W_2 \varepsilon$ represent the average value of the dependent variable and the residuals from neighbouring locations, respectively.

From this general equation the four forms of SAR models can be obtained by setting either or both of the parameters ρ and λ equal to zero. In the simplest case both ρ=0 and λ=0 the SAR model reduces to the standard regression model

$$y = X\beta + \varepsilon$$

(Anselin 1998). Setting λ=0 yields a model with a spatial autoregressive term, $\rho W_1 y$, called the spatial lag:

$$y = X\beta + \rho W_1 y + \varepsilon$$

By setting ρ=0 the spatial dependence is contained in the model residuals, giving the spatial error model:

$$y = X\beta + \varepsilon, \quad \varepsilon = \lambda W_2 \varepsilon + u .$$

Finally if neither ρ nor λ are zero, the model contains both a spatial lag term and spatial disturbance in the error terms. This higher order spatial model is rarely used in practice and is not considered further in this paper.

2.2 Spatial Lag Model

The spatial lag model, as shown in Section 2.1, introduces spatial dependence by directly adding the spatial lag of the dependent variable as a covariate in the model. The spatial lag model derives from a theoretical application where there is a reason to suspect that the dependent variable has a direct effect from neighbouring locations. This may be caused by spatial diffusion, copy-cat behaviour or other social interactions. The spatial autoregressive parameter, ρ, is then a measure of the magnitude of this interaction.

The interpretation of the model coefficients for the explanatory variables is the same as that of a standard regression model. However, the interpretation of the spatial lag coefficient depends on the context of the data. If there is a direct match between the spatial units used to measure the phenomenon and the scale of the underlying process that controls it, the spatial lag coefficient illustrates the direct effect of neighbouring locations. In this case the variable of interest is distributed homogeneously in each neighbourhood, as shown in Figure 2A. Alternatively, if the spatial unit used to measure the phenomenon does not conform to the spatial scale at which it occurs, as shown in Figure 2B, then the spatial lag coefficient represents the mismatch between the spatial scales and thus has no direct interpretation in the model. This pattern is typical for crime data, if the crime rate is measured on predefined administrative units such as census tracts.

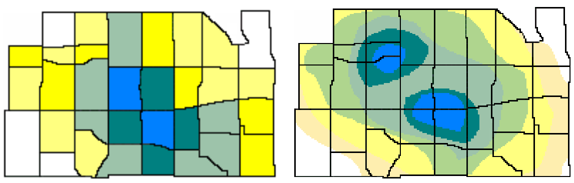

Figure 2: Interpretation of the spatial lag coefficient: A) illustrates the spatial units matching and B) illustrates a different spatial scale of measurement and occurrence.

An alternate form of the spatial lag model:

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

illustrates that the dependent variable at a given location is determined not only by the independent variables at that location but also by the independent variables at all other locations by means of the spatial multiplier $(I - \rho W_1)^{-1}$ (Baller et al. 2001). Similarly, the dependent variable at a given location is correlated with the error terms at all other locations. Due to these correlations the model cannot be efficiently estimated by a least squares technique.

Instead, a two-stage approach is preferred for estimation of the spatial lag model. First an estimate of ρ is obtained by the non-linear optimization of a log-likelihood equation. The resulting value of $\hat{\rho}$ is used to estimate the regression coefficients and the standard deviation by maximum likelihood. The estimate $\hat{\rho}$ is obtained by maximizing the concentrated log-likelihood:

$$L = -\frac{N}{2} \ln\left[\frac{(e_O - \rho e_L)'(e_O - \rho e_L)}{N}\right] + \sum_i \ln(1 - \rho w_i),$$

where the variables w_i represent the eigenvalues of the spatial weights matrix, e_o represents the residuals from a regression of y on X, and e_L represents the residuals from the regression of $W_1 y$ on X (Anselin and Bera 1998). The concentrated log-likelihood equation is obtained by substituting the maximum likelihood estimators of β and σ^2 into the likelihood equation of the spatial model and rearranging to express the likelihood as a function of only one parameter, ρ. This likelihood equation makes use of the simplification of the Jacobian term, $|I - \rho W_1|$, which can be computed as

$$|I - \rho W_1| = \prod_{i=1}^{N} (I - \rho w_i)$$

(Ord 1975).

The resulting estimates have the usual asymptotic properties of consistency, normality and asymptotic efficiency (Anselin and Bera 1998). For hypothesis testing the asymptotic variance is estimated as

$$\mathrm{AsyVar}\left[\hat{\rho}, \hat{\beta}, \hat{\sigma}\right] =$$

$$\begin{bmatrix} tr[W_A]^2 + tr[W_A' W_A] + \frac{[W_A X\hat{\beta}]'[W_A X\hat{\beta}]}{\hat{\sigma}^2} & \frac{(X' W_A X\hat{\beta})'}{\hat{\sigma}^2} & \frac{tr(W_A)}{\hat{\sigma}^2} \\ \frac{X' W_A X\hat{\beta}}{\hat{\sigma}^2} & \frac{X'X}{\hat{\sigma}^2} & 0 \\ \frac{tr(W_A)}{\hat{\sigma}^2} & 0 & \frac{N}{2\hat{\sigma}^4} \end{bmatrix}^{-1}$$

where $W_A = W_1 (I - \rho W_1)^{-1}$. From this matrix it can be shown that the covariance between the error term and the regression coefficients is zero, as it is in the standard regression model, but this is not true for the covariance between the spatial coefficient and the error, or between the spatial and regression coefficients.

There are two types of observed error terms that can be defined for the spatial lag model, prediction errors and residual terms (Anselin 2005). Prediction errors are defined as the difference between the observed and fitted values, $p.e. = y - \hat{y}$, where $\hat{y} = (I - \hat{\rho} W_1)^{-1} X\hat{\beta}$. The prediction errors are expected to be spatially correlated according to the model. The residuals, defined as $e = (I - \hat{\rho} W_1) y - X\hat{\beta}$, will be uncorrelated if the model adequately accounts for spatial autocorrelation. Therefore, the residuals and not the prediction errors are used to test for any remaining spatial autocorrelation in the model.

2.3 Spatial Error Model

In the spatial error model the spatial disturbance is in the model residuals, as shown by the following alternate form of the model:

$$y = X\beta + (I - \lambda W_2)^{-1} u \,.$$

The model derives from a practical application where spatial autocorrelation is detected in the residuals of a typical linear regression model and the effect is modelled to obtain unbiased and efficient estimates of the regression parameters. The spatial autocorrelation may result from unmeasured variables or measurement errors that have a systematic spatial pattern. The spatial coefficient, λ, is therefore treated as a nuisance parameter in the model and has no direct interpretation.

An alternate way to express the spatial error model is by spatially filtering the dependent and independent variables by the spatial multiplier $(\mathrm{I} - \lambda W_2)$. This gives the following model

$$(I - \lambda W_2)y = (I - \lambda W_2)X\beta + u$$

where, as previously noted, the error terms, u, are normally distributed with a homoskedastic variance. This form of the model provides an alternate interpretation of how the spatial effects are modelled in the spatial error model with an autoregressive error structure.

Since the spatial structure of this model is in the residuals the parameters for this model can be estimated using generalised least squares techniques that are conditional on the value of λ. As in the case of the spatial lag model, the value of λ is estimated first using maximum likelihood and this value is used in the step that estimates β and σ^2 by generalized least squares.

The concentrated log-likelihood equation used to estimate λ is obtained by substituting the generalised least squares estimators of β and σ^2 as functions of λ into the likelihood equation of the spatial error model. The likelihood equation also uses the Ord simplification for the Jacobian term, $|I - \lambda W_2|$ as shown for the spatial lag model (1975). The concentrated log-likelihood equation is then:

$$L = -\frac{N}{2}\ln\left(\frac{u'u}{N}\right) + \sum_i \ln(1 - \lambda w_i),$$

where $u'u = y_L' y_L - y_L' X_L [X_L' X_L]^{-1} X_L' y_L$, and y_L and X_L are spatially filtered variables: $y_L = y - \lambda W_2 y$ and $X_L = X - \lambda W_2 X$, and the variables w_i represent the eigenvalues from the spatial weights matrix.

The asymptotic variance matrix of the parameter estimates is block diagonal between $\hat{\beta}$ and the other two parameters, $\hat{\sigma}^2$ and $\hat{\lambda}$. It is estimated by

$$\mathrm{AsyVar}\left[\hat{\beta}\right] = \hat{\sigma}^2 [X_L' X_L]^{-1}$$

and

$$\mathrm{AsyVar}\left[\hat{\sigma}^2, \hat{\lambda}\right] = \begin{bmatrix} N/2\hat{\sigma}^4 & \dfrac{tr(W_B)}{\hat{\sigma}^2} \\ \dfrac{tr(W_B)}{\hat{\sigma}^2} & tr(W_B)^2 + tr(W_B' W_B) \end{bmatrix}^{-1}$$

where $W_B = W_2 (I - \hat{\lambda} W_2)^{-1}$ (Anselin and Bera 1998).

Alternate forms of the spatial error model can be obtained by specifying a different structure of the error process, contrary to the autoregressive process illustrated here. For example, a spatial moving average process is also possible where the error term is dependent on a random uncorrelated error term and the average neighbouring value of the uncorrelated errors (Anselin and Bera 1998).

3. Analysis of Crime in the Neighbourhoods of Montreal

3.1 Description of the Study Area

The census metropolitan area (CMA) of Montreal is located in the province of Quebec in eastern Canada. It is the second largest CMA in Canada, with a population of over 3.4 million people in 2001. The Island of Montreal, which falls within the larger CMA boundary, is the focus of this study (Figure 3). The residential population of the Island of Montreal is 1.8 million and spans an area of approximately 500 square kilometres.

The neighbourhood boundaries used in this study are defined by the 521 census tracts on the island. Fifteen of the neighbourhoods were removed from the analysis because the resident population was less than 250 inhabitants and Statistics Canada suppresses information on such regions due to issues of confidentiality and data quality. Neighbouring locations were defined by the queen contiguity structure, which includes regions that have either a border or vertex in common.

Figure 3: The Island of Montreal, with neighbourhoods defined by census tracts.

3.2 Description of Variables

3.2.1 Crime Data

The crime data for Montreal in 2001 was obtained from the incident-based Uniform Crime Reporting survey (UCR), which collects detailed information on criminal incidents. The UCR is a census of administrative data from all police services across Canada. Thus the crime data reflect only the incidents known to the police, giving a specific picture of the nature and extent of crime (Savoie et al. 2006). In the UCR survey, individual offences of the criminal incident are classified by their level of seriousness, which is based on the maximum possible sentence under the *Criminal Code*.

The criminal incidents were grouped into the two broad categories of violent and property crime for the analysis. Violent offences are defined as crimes against the person and include: homicide, attempted murder, sexual assault, assault, robbery and extortion. Property offences, i.e. crimes against property, include: arson, breaking and entering, theft, possession of stolen goods, fraud and mischief. The number of incidents recorded in each category is based on only the most serious offence reported in each criminal incident. Thus the less serious offences may be undercounted if they occur in conjunction with more serious crimes.

For the analysis the crime categories were represented by the rate within each neighbourhood, which was computed as the number of incidents per population at risk. A count of the number of incidents that occurred in each neighbourhood was obtained by geocoding the street address of each criminal incident reported to the UCR survey. The population at risk was defined as the combined resident population and the population who work in the area. It is used to more accurately represent the risk of crime in each neighbourhood, which would otherwise be inflated for downtown areas where there is a relatively small residential population but a high concentration of people either working or engaged in other activities (Savoie et al. 2006).

3.2.2 Explanatory Variables

The neighbourhood characteristics included variables describing the demographics and socio-economic status of the resident population, the condition of dwellings and zoning for city land use. The population and dwelling characteristics were obtained from the Census of Population conducted on May 15, 2001 by Statistics Canada. Demographic characteristics of the resident population included the percentage of males in the neighbourhood aged 15 to 24, the percentage of the population aged 15 and over who are single and have never been married and the percentage of visible minority residents in the neighbourhood. Socio-economic status and dwelling characteristics included variables such the percentage of the population in private households with a low income in 2000, the percentage of the residents over 20 years of age who have obtained a bachelor's degree, the median household income, the percentage of dwellings in need of major repair and the percentage of owner occupied dwellings in the neighbourhood.

The variables on city land use, including the proportion of the area in the neighbourhood that was zoned for commercial land use and for residential single and multiple family dwellings, were obtained from land use data in the 2005 geomatics department database at the Communauté métropolitaine de Montréal, and zoning data was obtained from the Montreal planning department. Finally, the Business Register Division of Statistics Canada provided the address of all drinking places, including bars, taverns and other drinking establishments on the Island of Montreal in 2001. This was used to create a variable on the bar density, representing the number of drinking places over the geographic area of each neighbourhood.

For further information on the variables used in this analysis and information on their potential influence on the crime rate refer to Savoie et al. (2006).

3.3 Analysis

Prior to fitting the model, a log transform was applied to the crime rate to obtain an approximately normal distribution for the dependent variable. The original explanatory variables were standardised to have a mean of zero and unit variance. At this point some of the variables with a high pairwise correlation were removed from the analysis to avoid a potential problem with multicollinearity. A standard linear regression model, estimated by the least squares technique was then fit to the transformed data. The standard regression model was selected from the initial set of explanatory variables using a stepwise analysis in the SAS PROC REG procedure. Variance Inflation Factors (VIFs) were used to confirm that there was no problem of multicollinearity, as the VIFs for all variables in the models were less than 3. VIFs measure the linear association between each variable and all of the other variables in the model by means of a coefficient of determination (Montgomery, Peck and Vining 2002). Inference for the standard regression model is based on the assumptions of normality and independence of the model error terms.

The residuals of the standard regression were then tested for the presence of spatial autocorrelation using Moran's I statistic. This is a global measure of the linear association between residuals and their average value from neighbouring locations (Anselin and Bera 1998). Moran's I statistic was 0.23 for property crime ($p<0.001$) and 0.15 for violent crime ($p<0.001$), indicating significant spatial autocorrelation in each model.

The standard and robust Lagrange Multiplier (LM) tests described by Anselin et al. (1996) were used to determine the form of the spatial dependence in the data. For both models the standard LM tests, which are powerful but not robust to local misspecification of the model, did not distinguish which spatial model was more appropriate for the data (Table 1). The robust LM tests indicated that the spatial lag model was more appropriate for the property crime data. For violent crime the spatial error model is slightly more appropriate; however the distinction between the two spatial models is not as clear as the results for property crime.

Table 1: Lagrange multiplier tests to determine the form of spatial dependence in the data.

	Property Crime		**Violent Crime**	
	Statistic	p-value	Statistic	p-value
Standard LM-lag	105.90	<0.001	27.37	<0.001
Standard LM-error	74.88	<0.001	29.52	<0.001
Robust LM-lag	32.47	<0.001	3.13	0.07
Robust LM-error	1.36	0.243	5.29	0.02

A spatial lag model was then fit to the property crime data and a spatial error model was fit to the violent crime data set using programs written in the PROC IML procedure of SAS. The residuals from the spatial lag model on property crime were tested for any remaining spatial autocorrelation in the error terms using a specific LM test defined by Anselin et al. (1996). The value of this statistic was 1.72 ($p=0.190$), indicating that there was no remaining spatial autocorrelation in the spatial lag residuals. Similarly, the residuals from the spatial error model were tested for model misspecification in the form of an omitted spatial lag variable using an analogous LM test. The value of this statistic was 1.77 ($p=0.183$) indicating that the spatial structure of the data was adequately specified by the spatial error model.

Model diagnostics from the standard regression and the spatial models were used to confirm that the spatial model was more appropriate for the data and that the addition of the spatial term did not result in any reduction in the goodness-of-fit. The Akaike Information Criterion (AIC) was significantly lower for the spatial model than the standard regression model in both the property and violent crime regressions (Table 2). The squared correlation between the observed and predicted values was slightly higher for the spatial lag model than the standard regression of property crime, indicating a slight improvement of the fit. For violent crime, the squared correlation was the same for the two models, indicating there was no loss of model fit with the spatial error model.

Table 2: Comparison of the model diagnostics for the standard regression and spatial models of property and violent crime.

	Property Crime		**Violent Crime**	
	Stand Reg	Spatial Lag	Stand Reg	Spatial Error
AIC	304.1	220.5	581.3	558.6
Sq-Corr (obs, pred)	0.59	0.62	0.60	0.60

3.4 Results

Table 3 compares the results of a standard regression and a spatial lag model for the property crime data. There are minor differences in the estimated regression coefficients from the standard regression and spatial models. However, of greater interest is the difference in the significance level of certain variables between the standard and spatial models. The variable representing the percentage of dwellings in need of major repair appears to be significant in the standard model but is clearly not significant in the spatial model. There is also a minor change in the p-value for the variables representing the percentage of visible minority residents and the bar density. These differences in the significance level can be attributed to the positive spatial autocorrelation present in the standard regression residuals, as indicated by Moran's

I statistic. Positive autocorrelation causes an underestimation of the sample variance which makes factors appear more significant then they actually are (Legendre 1993). The large difference in the significance of the major repairs variable may also be caused by the proportionately large difference in the estimated coefficient of the standard and spatial models (-0.039 to -0.013, respectively). Thus the final model for the property crime data should eliminate the major repairs variable.

Table 3: Comparison of a standard regression and a spatial lag model for the rate of property crime.

	Standard Reg		**Spatial Lag**	
Variable	Coeff	p-value	Coeff	p-value
Intercept	3.47	<.0001	1.98	<.0001
Low Inc. Private	0.17	<.0001	0.11	<.0001
Visible Minority	-0.08	<.0001	-0.05	0.0040
Single	0.19	<.0001	0.11	<.0001
Commercial Area	0.14	<.0001	0.12	<.0001
Bar Density	0.07	<.0001	0.05	0.0011
Major Repairs	-0.04	0.0107	-0.01	0.3405
Spatial Lag			0.43	<.0001

Comparing the results of a standard regression and spatial error model for the violent crime rate we see very small differences in the regression coefficients and the p-values (Table 4). The similarity of the regression coefficients reflects that the standard regression estimates are unbiased although inefficient, due to the non-spherical structure of the error covariance in the data (Anselin and Bera 1998). The spatial model, however, allows for more efficient estimation of the parameters by modelling the spatial process in the error terms. The significance of the spatial autoregressive coefficient confirms that the spatial error model is more appropriate for this data than a standard regression model. Similar regression coefficients are also obtained from a spatial lag model fit to the violent crime data. The only noticeable difference in the model parameters from a spatial lag and spatial error model is the p-value for the bar density variable, which is 0.1287 for the spatial lag model and 0.0577 for the spatial error model. This difference reflects the importance of examining alternative forms of spatial models for a given dataset.

Table 4: Comparison of a standard least squares and a spatial error model for the rate of violent crime.

	Standard Reg		**Spatial Error**	
Variable	Coeff	p-value	Coeff	p-value
Intercept	2.09	<.0001	2.08	<.0001
Low Inc. Private	0.22	<.0001	0.20	<.0001
Bachelor's Degree	-0.27	<.0001	-0.26	<.0001
Single	0.21	<.0001	0.20	<.0001
Commercial Area	0.07	0.0006	0.07	0.0008
Single Family Zone	0.11	<.0001	0.09	0.0005
Multi Family Zone	0.11	<.0001	0.12	<.0001
Bar Density	0.04	0.0689	0.04	0.0577
Spatial Error			0.32	<.0001

4. Conclusions

The comparison of the standard regression and spatial models for the crime rate on the Island of Montreal illustrate the importance of considering the spatial nature of the data. The spatial models not only provide a better fit to the data, but they allow for more accurate inference on the model parameters, in turn yielding a more accurate final model. This may have an implication for crime reduction programs in cases where the decisions on how to allocate funds for reducing crime are based on results from analyses such as the one performed here. In the case of property crime in Montreal, the use of a standard regression model may have erroneously resulted in the recommendation of a policy to repair housing in at-risk neighbourhoods as a possible means of reducing crime.

Acknowledgements

Funding for this project was provided by the National Crime Prevention Centre at Public Safety and Emergency Preparedness Canada and by the Canadian Centre for Justice Statistics at Statistics Canada.

The authors wish to thank Josée Savoie and Frédéric Bédard from Statistics Canada for their work in preparing the dataset and in the analysis.

References

Anselin, L. 1988. *Spatial econometrics: Methods and Models*. Kluwer, Dordrecht.

Anselin, L. 2002. Under the hood: Issues in the specification and interpretation of spatial regression analysis. *Agricultural economics*. 27:247-267.

Anselin, L, A. Bera, R. Florax, M. Yoon. 1996. Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*. 26: 77-104.

Anselin, L. and A. Bera. 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. *IN*: A. Ullah and D. Giles (eds.) *Handbook of applied economic statistics*. New York: Marcel Dekker.

Anselin, L., J. Cohen, D. Cook, W. Gorr and G. Tita. 2000. Spatial analysis of crime. *In:* D. Duffee (ed.) *Criminal Justice: Vol 4. Measurement and analysis of crime and justice*. Washington, D.C. National Institute of Justice. pp. 213-262.

Baller, R., L. Anselin, S. Messner, G. Deane and D. Hawkins. 2001. Structural covariates of U.S. county homicide rates: incorporating spatial effects. *Criminology*. 39(3): 561-590.

Fitzgerald, R., M. Wisener, J. Savoie. 2004. *Neighbourhood characteristics and the distribution of crime in Winnipeg*. Crime and Justice Research Paper Series. Statistics Canada Catalogue no. 85-561-MIE2004004, Ottawa. Canadian Centre for Justice Statistics.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology*. 74(6): 1659-1673.

Morenoff, J., R. Sampson and S. Raudenbush. 2001. *Neighbourhood inequality, collective efficacy and the spatial dynamics of urban violence*. Research Report, no. 00-451. Population Studies Cenre. Institute for Social Research. University of Michigan.

Montgomery, D., E. Peck and G. Vining. 2001. *Introduction to linear regression analysis.* 3rd edition. Wiley, New York.

Puech, F. 2004. How do criminals locate? Crime and spatial dependence in Minas Gerais. *Urban/ Regional 0509018 Economics Working Paper Archive Econ WPA*.

Savoie, J., F. Bédard and K. Collins. 2006. *Neighbourhood characteristics and the distribution of crime on the Island of Montreal*. Crime and Justice Research Paper Series. Statistics Canada Catalogue no. 85-561-MIE2006007, Ottawa. Canadian Centre for Justice Statistics.