VII－16

Analysis of Urbanization Related Environmental Impacts from their Spatial Perspective

Tohoku University, Student Member, Antonio Paez
Tohoku University, Regular Member, Takashi Uchida
Tohoku University, Fellow, Kazuaki Miyamoto

1. Background.

It has been argued that there is a clear and present necessity of urban-environmental models, with which to conduct analysis of the complex interactions occurring between the city and the environment. In particular, there is a perceived lack of models to evaluate and analyze the environmental impacts caused by urbanization (Miyamoto and Sathyaprasad, 1995).

Given the complexity of the respective systems, causal, deterministic relationships between human activities and environment are difficult to establish and analyze. On the other hand, statistical relationships can be frequently found that easily incorporate a variety of data available for analysis, and allow statistical model development. However, difficulties arise when conventional statistic methods are applied to the kind of data commonly used in the environmental and social sciences, which is quite frequently spatial in nature.

The purpose of this study is to demonstrate that explicit consideration of space is important or even necessary, given the limitations imposed by non-spatial modeling specifications. This is illustrated by taking the results of an urban heat island model, and testing them, by means of an underlying alternative spatial model, to find possible sources of misspecification.

2. The Model.

Here, a modeling approach to the description of the urban heat island effect is presented. The urban heat island is an atmospheric effect that consists of a temperature differential between the city and its surrounding countryside. As the result of land cover changes, it is a clear manifestation of urban-environment interactions and thus represents a unique opportunity to study the impacts that urbanization has on its immediate environment.

Data was used for Sendai City, including temperatures during a heat island episode, and a number of land use, city size, and activity intensity variables handled in a GIS environment. Relevant variables were found to be land prices as an index of urban activity; percentage of land use by use by zone (i.e. commercial, green or open space, etc.); physical measures of city's size, as distance from station; and other measures of size, like population density. The model was then specified as a regression of the temperature on the rest of the variables, using the ordinary least squares (OLS) model:

$$Y=X\beta+\varepsilon, \qquad (1)$$

where Y and X are the usual vector and matrix of observations, β is a vector of parameters, and ε a vector of independent and identically distributed error terms.

The regression is satisfactory in terms of the commonly reported statistics, as they appear in table 1. Analysis of the regression error terms, however, reveals that the usual assumption of independence among the residuals can not be sustained. Moran's I autocorrelation statistic was calculated using a binary weight matrix W and a row standardized weight matrix Wst that formally impose a spatial structure to the problem.

Table 1. Urban Heat Island Model. OLS Regression

VARIABLE	VALUE	T
CONST	299.6173	571.62
Log Land Price	1.140737	5.86
Dist from Station	−0.00012	−4.08
Total Use Area	−0.05007	−1.77
Commercial	0.050202	1.76
Industrial	0.046269	1.60
Residential	0.061732	2.07
Population Density	7.79E−05	3.00
Population	−0.00039	−6.63
Green	0.026608	3.68

ST DEVIATION= 0.9841 R2= 0.5174
Adjusted R2= 0.5084 N= 493

Table 2. Spatial Autocorrelation Statistics.

Binary Weight Matrix W		St Weight Matrix Wst	
I	0.533	I	0.546
E(I)	−0.0090	E(I)	−0.0092
VAR(I)	6.76E−04	VAR(I)	6.99E−04
Norm Z	20.85	Norm Z	20.99

Keywords: Modeling, Spatial Autocorrelation, Spatial Dependence
Laboratory for Regional Planning and Applications, Department of Civil Engineering, Tohoku University, Aoba, Aoba-ku, Sendai 980-8579. Tel. 022-217-7478. Email: antonio@rs.civil.tohoku.ac.jp

It is clear from the autocorrelation statistics shown in table 2 and appearing as standardized normal (Z) that the model does not comply with the independence condition assumed for the regression residuals, and is thus incorrect.

To find possible sources of error, the following specification corresponding to a spatial model was considered:

$$Y=\rho W Y+X\beta+\varepsilon \qquad (2)$$

$$\varepsilon=\lambda W\varepsilon+\mu \qquad (3)$$

where ρ and λ are autocorrelation parameters, W is the aforementioned weight matrix, μ is a vector of stochastic errors, and the rest is as defined before. This specification explicitly considers space, by means of the weight matrix, and several forms of dependence.

A testing framework based on Maximum Likelihood consists of diagnosing whether any or all parameters are statistically different from zero. If this is the case, it is taken as evidence that the above OLS model is underspecified and certain parameters are missing. For the case of missing ρ, an omitted lag is indicated, while for λ the indication is for spatial error autocorrelation. The test is based on a Lagrange Multipliers approach, and the results are shown in table 3 for the case of a binary weight matrix *W*. Similar values result for the standardized weight matrix *Wst*. The statistics are χ^2 distributed with q degrees of freedom and it is clear that all the tests are statistically significant, with probabilities of over 99.5%, except for one test in the W matrix set of diagnostics.

Table 3. LM Diagnostics for Spatial Effects, Using a Binary Weight Matrix.

Lagrange Multiplier Diagnostics for Spatial Effects in OLS regression	q	chi 2
One Directional Tests		
Spatial Error Autocorrelation SEA	1	401.4097
Omitted Spatial Lag	1	2.6367
Random Coefficient Variation	9	145.1095
Multidirectional Tests		
Spatial Dependence	2	404.622
SEA and Heteroskedasticity	10	546.5192
All Effects	11	549.7315

4. Discussion.

The model for the urban heat island originally presented was specified using a standard statistical/econometric methodology. However, considering the spatial nature of the data under study, two shortcomings to this kind of approach were suspected. First, standard regression on the data is believed to be wasteful of information, since it ignores that additionally provided by the known relative locations of the observations. And second, a model fitted for spatial data that does not incorporate space in an explicit way, will most certainly be ill specified and fail to represent the spatial dimension of the phenomena.

For this particular model, these suspicions were corroborated in the first place by statistical testing of the OLS regression residuals, using Moran's *I* autocorrelation statistic. The test pointed out that the model was incorrect or at least incomplete, by showing violations to underlying modeling assumptions (most importantly the independence condition). An intrinsic characteristic of the *I* statistic is, however, that it does not identify the sources of misspecification, which thus remained uncovered.

A more in depth, albeit somewhat more complex search, involved additional testing of an underlying fully specified spatial model. From among three testing procedures, namely the Likelihood Ratio Test, Wald Tests, and the Lagrange Multiplier Tests, selection was made of the latter because it does not necessitate the evaluation of the full model, and therefore could be more easily implemented. Several sources of misspecification were detected in this way, including omitted spatial dependencies and heteroskedasticity, thus alerting of the necessity of including the spatial effects by means of a model expanded with a spatially lagged dependent variable, or with autocorrelated error terms.

A difficulty is that OLS parameters for this latter kind of models (spatial effects) do not retain their BLUE properties, and Maximum Likelihood techniques have to be used for estimation. However, it is believed that the computational difficulties associated with ML estimation are offset by a richer gamma of modeling possibilities, and the potential for incorporating spatial data (i.e. GIS handled) into the analysis without losing its spatial context.

Acknowledgments.

Thanks are given to Prof. Saito for kindly granting permission to use the heat island data simulated by the Department of Aeronautics and Space Engineering at Tohoku University.

References.

1. Miyamoto, K., Sathyaprasad, S., An Estimation System of Urban Land-Use and the Environment for Metropolitan Areas in Developing Countries. *Studies in Regional Science*, 25-1, 1995.