

Isolating Spatial from A-spatial Components of Housing Attributes Using Kriging Techniques

by

François Des Rosiers¹, Marius Thériault²,

Paul-Y Villeneuve³ and Yan Kestens⁴

**Extended abstract of a paper submitted for presentation
at the European Real Estate Society 8th Annual Conference
Alicante, Spain, June 26-29, 2001**

1. Faculty of Business Administration, Planning and Research Centre, Laval University, Québec, Canada, G1K 7P4
Phone: 418-656-2131 ext. 5012,
e-mail: Francois.Desrosiers@fsa.ulaval.ca
2. Department of Planning, Director, Planning and Research Centre, Laval University, Québec, Canada, G1K 7P4
Phone: 418-656-2131 ext. 5899, Fax: 418-524-6701
e-mail: Marius.Theriault@crad.ulaval.ca
3. Department of Planning, Planning and Research Centre, Laval University, Québec, Canada, G1K 7P4
Phone: 418-656-2131 ext. 3791,
e-mail: Paul.Villeneuve@crad.ulaval.ca
4. Ph.D candidate, Planning and Research Centre, Laval University, Québec, Canada, G1K 7P4
Phone: 418-656-2131 ext. 8878
e-mail: Yan.Kestens@crad.ulaval.ca

Key Words: GIS, Hedonic modelling, Accessibility, Neighbourhood factors, Urban externalities, Housing markets, Spatial analysis, Factor analysis, Kriging.

SUMMARY OF PAPER

1. Introduction : Context and Objective of Research

The current piece of research is an attempt to isolate spatial from a-spatial components of housing attributes, using kriging techniques. The hedonic approach, on which is based this investigation, is applied to the residential market of the Quebec City region for two points in time. As put by several authors, property values are a reflexion of complex overlapping effects which combine externalities and location factors (Krantz et al. 1982; Hickman et al. 1984; Shefer 1986; Yinger et al. 1987; Strange 1992; Can 1993; Dubin 1998; Hoch and Waddell 1993; Des Rosiers et al. 1996). While hedonic models have long proved their usefulness as an analytical device, previous research has shown that substantial portion of price variability remains unexplained (Anselin and Can 1986; Dubin and Sung 1987; Can 1993; Dubin 1998). The great deal of neighbourhood factors needed to adequately account for submarket specifics (Adair, Berry and McGreal 1996 & 1998) raises methodological issues linked to the presence of excessive multicollinearity between model attributes, as well as to structural heteroskedasticity and spatial autocorrelation among residuals; all of these are detrimental to the stability of regression coefficients (Dubin 1988; Anselin and Rey 1991; Can and Megbolugbe 1997; Basu and Thibodeau 1998; Pace, Barry and Sirmans 1998; Des Rosiers and Thériault 1999).

Among the issues that need be addressed, spatial dependence - a current feature of property markets - deserves substantial research efforts. Indeed, several situations involve the presence of spatial autocorrelation, thereby violating the basic assumptions underlying hedonic models. For instance, while liveable area will partly mirror size features that are exclusive to a given piece of property, it also reflects neighbourhood characteristics which are common to all houses in the same market segment. Consequently, implicit prices of housing attributes may prove to be both biased and unstable (Can 1990 & 1993, Dubin 1998). In this paper, kriging techniques are used to address the problem. Once spatial dependence has been identified for each housing descriptor, relevant variables are then split into their spatial and a-spatial components; this results in two sets of unbiased coefficients being calibrated. Model residuals are also processed

via kriging. In so doing, space-related influences on prices can be isolated for each attribute and their consistency assessed over space and time as well.

Resorting to kriging requires an extensive use of Geographic Information Systems (GIS) and, in particular, of spatial statistics methods (Anselin and Getis 1992; Griffith 1993a; Zhang and Griffith 1993; Thériault and Des Rosiers 1995; Levine 1996) such as variography and spatial interpolation techniques (Dubin 1992; Panatier 1996).

2. Previous Research Findings: Controlling for Collinearity

Previous research papers have addressed both the multicollinearity and spatial autocorrelation issues. Following other examples in the recent hedonic literature (Bourassa et al. 1999), Des Rosiers, Thériault and Villeneuve (2000a) resort to factor analysis (Thurstone 1947; Rummel 1970; Can 1992) in order to generate independent complex variables used as substitutes for initial attributes, thereby reducing collinearity to a minimum. The study is based on a sample of some 2 400 town cottages sold in the Quebec region from January 1993 to January 1997. Accessibility is computed at the level of individual properties (Nijkamp, Van Wissen and Rima 1993) and based on travel times, using some 19 250 street intersections (Thériault, Des Rosiers and Vandersmissen 1999). Census data used to define neighbourhood attributes are aggregated at the level of 604 enumeration areas.

Using the principal component method (PCA) with a Varimax rotation, some 49 initial variables are thus being reduced to only six significant factors - four neighbourhood and two access principal components - which are in turn integrated in the hedonic equation as complex independent variables. Findings clearly suggest that factor analysis is highly efficient at sorting out access and neighbourhood dimensions. In particular, the educational profile of local residents as well as access to regional and local services emerge as powerful price determinants.

As an extension to the above study, Des Rosiers, Thériault and Villeneuve (2000b) look at house price determinants from a dynamic perspective by investigating the influence of socio-demographic, as well as socio-economic, changes over the 1986-1996 period upon property values in the Quebec Urban Community (QUC - roughly 675 000 in population). The general data base includes some 14 400 single-family properties transacted between

January 1993 and January 1997. Two market segments are analyzed, namely town cottages (1 708 sales) and condominium units (940 sales), with sales in the \$100 000 - \$250 000 price range.

Here, enumeration areas are recomputed using spatial smoothing techniques in order to produce some 416 homogeneous delimitations that allow for comparison over time. Calculation of standardized, or time de-trended, socio-demographic changes between censuses (1986-91; 1991-96; 1986-96) are then performed using linear regression adjustments, while a series of five PCAs are computed on socio-demographic attributes: while state components are derived from each census (1986, 1991 and 1996), change components are computed on the 1986-91 and 1991-96 periods. A sixth PCA is also performed on access variables (1988 road network, extended to the whole Quebec Metropolitan Area). Finally, significant principal components are substituted for individual variables in the hedonic equations.

Findings indicate that both socio-demographic structures (state variables) and their evolution (change variables) over time do affect house prices significantly, with most principal components retained in the analysis displaying strong *t* values. As expected then, neighbourhood profiles as well as changes are part of the price setting process and should not be overlooked in hedonic modelling.

3. Dealing With Spatial Autocorrelation Using Kriging

Spatial autocorrelation, or dependence, refers to the relationships between neighbouring values in space. It measures the degree of resemblance (positive) or dissimilarity (negative) between places or individuals as a function of the distance which separates them. In real estate, it stems from the influence that geographical structures exert upon urban rent through differentials in access to services, socio-demographic structures and environmental quality. It also mirrors the space-related similarities among building attributes within homogeneous neighbourhoods. There are a number of statistical indexes for measuring the degree of spatial autocorrelation (Anselin 1995; Cressie 1993; Getis and Ord 1992; Griffith 1993a; Odland 1988; Ord and Getis 1995; Ripley 1981). The one most commonly used - also considered the most robust from a mathematical standpoint - was developed by the statistician Moran (1948 and 1950) after whom it was named (Moran's I).

The Moran's I expresses the intensity of the relationship between any value located at a specific place (census tract or property) and the similar for its neighbours. The role it plays with respect to the analysis of spatial structures is similar to that of the correlation coefficient in conventional statistics. While there is no bounds for the index, in practice, its values range between +1 and -1. A positive value expresses an association by resemblance between neighbours; a negative value corresponds to a dissimilarity; and a zero value indicates the absence of any defined spatial structure. The Moran's I thus provides a parametric test for assessing whether, and to what degree, observed spatial structures are random. Since multiple regression analysis requires that observations be independent from one another in order for hypothesis testing to be reliable (Anselin and Rey 1991; Griffith 1993b), testing for the presence of spatial autocorrelation is a prerequisite for good hedonic modelling.

First developed by Krige (1951) to model geological structures, kriging is now considered as one of the most powerful spatial interpolation methods (Panatier 1996). It is based on an analysis of the spatial variance of space-distributed values. The latter are used to build experimental variograms whereby the variation in the observed means differentials between values - in this case, property sale prices - is expressed in terms of the distance which separates them in space (Cressie 1993). Variograms are then applied theoretical functions so as to obtain the best adjustment for value variations resulting from proximity, which is accounted for in the spatial interpolation of the phenomenon.

Kriging is to the study of spatial distributions what regression analysis is to conventional statistics. While independently developed, spatial autocorrelation and kriging techniques rest upon similar spatial variance and covariance concepts (Griffith and Csillag 1993). Kriging makes it possible to model proximity relationships between the values of socio-demographic and economic attributes of adjoining neighbourhood units in order to compute space de-trended residuals which are representative of a local variation, and therefore independent from the regional trend. In so doing, kriging can contribute to properly identify areas with socio-demographic and economic profiles that differ from those of adjacent areas while bringing out flaws in the urban fabric. Among other things, it should help understand the detrimental effect of an inadequate division of space for census purposes (Wong and Amrhein 1996). Finally, kriging adjusts continuous functions to

principal components, thereby smoothing expressions of spatial structures for individual properties included in hedonic models.

In a recent paper, Des Rosiers et al. (2001) resort to kriging in order to measure the effect of micro-level, neighbourhood profiles on house values and market differentiation in the presence of spatial autocorrelation. While principal component analysis (PCA) derived from former research serves as a sorting device to identify space-structuring factors, kriging is used to deal with spatial dependence among model residuals. A global sample of 2 405 town cottages sold in the QUC from January 1993 to January 1997 is used for the modelling process. Two randomly chosen, equal-size sub-samples are also defined for cross-validation purposes. Isotropic exponential and spherical variogram functions are calibrated for interpolation of model residuals.

Once reinserted in the hedonic equations as independent vectors, kriged variables are shown to substantially improve both explanatory and predictive performances. Thus, the global model adjusted R-Square is raised to 95.8% from 88.8% previously while the standard prediction error drops from 11.3% to 6.9% of mean sale price. In spite of methodological limitations to the method which translate into lower performances for cross-validated sub-samples, findings suggest that most spatially structured variations are being captured by the adjusted model, leaving the remaining spatial autocorrelation statistically non significant while regression coefficients are unbiased. Kriging therefore emerges as a promising device to complement the traditional modelling approach. However, it could lead to further improvement if performed individually on every housing attribute so as to sort out in each case the spatial and a-spatial components of price variations. This is precisely the object of the current piece of research.

4. Spatial vs. A-spatial Components of House Price Variations

As mentioned earlier, space-related housing attributes will cause spatial dependence to alter the reliability of hedonic prices. While applying kriging on model residuals allows for an overall adjustment to be made on house values, it does not help isolate the spatial from the a-spatial components of implicit prices. These may differ substantially among housing attributes and will vary according to neighbourhoods as well as over time. The object of this research is to identify such components and to measure their consistence

over both space and time. In order to achieve that, kriging is performed on the QUC bungalow (single-family, one story, detached houses) segment of the residential market. Two points in time are considered for comparative purposes, namely the 1990-91 and the 1993-97 periods. For each sample, a sub-sample is defined at random for cross-validation purposes. Over twenty descriptors are used in the initial modelling stages. Considering that kriging is applied to each vector as well as to the ensuing residuals, the final model will include twice that number of parameters, although several of them may eventually prove statistically non significant.

Based on recent research findings, it is expected that such a modelling approach will bring useful insights into the house price determination process by sorting out micro-level, space-related specifics operating at the neighbourhood level and truly individual housing characteristics which part from local averages. While explicitly accounting for spatial autocorrelation, variable-designed kriging is also expected to improve the robustness of derived hedonic coefficients.

REFERENCES

- Adair. A.S., J.N. Berry and S.W. McGreal, 1998. Assessing Influences upon the housing Market in Northern Ireland. *Journal of Property Research*, 15 (2): 121-134.
- Adair. A.S., J.N. Berry and S.W. McGreal, 1996. Hedonic Modelling, Housing Submarkets and Residential Valuation. *Journal of Property Research*, 13 (1): 67-83.
- Anselin, L., 1995. Local indicators of spatial association - LISA. *Geographical Analysis*, 27(2): 93-115.
- Anselin, L., and A. Can, 1986. Model comparison and model validation issues in empirical work on urban density functions. *Geographical Analysis*, 18: 179-197.
- Anselin, L., and A. Getis, 1992. Spatial statistical analysis and geographic information systems. *The Annals of Regional Science*, 26: 19-33.
- Anselin, L., and S. Rey, 1991. Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23(2): 112-131.
- Basu, S., and T. G. Thibodeau, 1998. Analysis of spatial autocorrelation in house prices. *The Journal of Real Estate Finance and Economics*, 17(1): 61-86.

- Bourassa, S. C., V.S. Peng, M. Hoesli and P. Forer, 1999. Defining Housing Submarkets Using Transaction Data. Forthcoming in the *Journal of Housing Economics*.
- Can, A., 1990. The measurement of neighbourhood dynamics in urban house prices. *Economic Geography*, 66: 254-272.
- Can, A., 1992. Residential quality assessment: Alternative approaches using GIS. *The Annals of Regional Science*, 26: 97-110.
- Can, A., 1993. Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22: 453-474.
- Can, A., and I. Megbolugbe, 1997. Spatial dependence and house price index construction. *Journal of Real Estate Finance and Economics*, 14: 203-222.
- Cressie, N. A. C., 1993. *Statistics for Spatial Data*, New York, Wiley, 900 p.
- Des Rosiers, François, Marius Thériault, Yan Kestens and Paul-Y. Villeneuve, 2001. Neighbourhood Profiles and House Values: Dealing with Spatial Autocorrelation Using Kriging Techniques. Working Paper # 2001-001, Faculty of Business Administration, Laval University, January, 27 pages.
- Des Rosiers, F., M. Thériault and P-Y. Villeneuve, 2000a. Sorting out Access and Neighbourhood Factors in Hedonic Price Modelling. *Journal of Property Valuation and Investment*, Vol.18 No. 3, pp. 291-315. .
- Des Rosiers, F., M. Thériault and P-Y. Villeneuve, 2000b. Assessing the Effects of Socio-Demographic Evolution On Housing Market Differentiation. Working Paper # 2000-020, Faculty of Business Administration, Laval University, August, 27 pages.
- Des Rosiers, F. and M. Thériault, 1999. " House Prices and Spatial Dependence : Towards an Integrated Procedure to Model Neighborhood Dynamics ", Working Paper # 1999-002, Faculty of Business Administration, Laval University, January, 23 p.
- Des Rosiers, F. A. Lagana, M. Thériault and M. Beaudoin, 1996. Shopping centers and house values: An empirical investigation. *Journal of Property Valuation and Investment*, 14(4): 41-63.
- Dubin, R. A., 1988. Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics*, 70: 466-474.
- Dubin, R. A., 1992. Spatial autocorrelation and neighbourhood quality. *Regional Science and Urban Economics*, 22: 433-452.
- Dubin, R. A., 1998. Predicting house prices using multiple listings data. *The Journal of Real Estate Finance and Economics*, 17(1): 35-60.
- Dubin, R. A., and C. H. Sung, 1987. Spatial variation in the price of housing: Rent gradients in non-monocentric cities. *Urban Studies*, 24: 193-204.
- Getis, A., and J. K. Ord, 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3): 189-206.
- Griffith, D. A., 1993a. Advanced spatial statistics for analyzing and visualizing geo-references data. *International Journal of Geographical Information Systems*. 7(2): 107-124.

- Griffith, D. A., 1993b. Exploring relationships between semi-variogram and spatial autoregressive models. *Papers in Regional Science*, 72(3): 283-296.
- Griffith, D. and Csillag, F., 1993. Exploring relationships between semi-variogram and spatial autoregressive models. *Papers in Regional Science*, 72(3): 283-296.
- Hickman, E. P., J. P. Gaines and F. J. Ingram, 1984. The influence of neighbourhood quality on residential values. *The Real Estate Appraiser and Analyst*, 50(2): 36-42.
- Hoch, I., and P. Waddell, 1993. Apartment rents: Another challenge to the monocentric model. *Geographical Analysis*, 25(1): 20-34.
- Krantz, D. P., R. D. Waever and T. R. Alter, 1982. Residential tax capitalization: Consistent estimates using micro-level data. *Land Economics*, 58(4): 488-496.
- Krige, D. G., 1951. A statistical approach to some basic mine valuation problems on the Wiatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52: 119-139.
- Levine, N., 1996. Spatial statistics and GIS. Software tools to quantify spatial patterns. *Journal of the American Planning Association*, 62(3): 381-390.
- Moran, P. A. P., 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10: 243-251.
- Moran, P. A. P., 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37: 17-23.
- Nijkamp, P., L. Van Wissen and A. M. Rima, 1993. A household life cycle model for residential relocation behaviour. *Socio-Economic Planning Science*, 27(1): 35-53.
- Odland, J., 1988. *Spatial Autocorrelation*. *Scientific Geography*, vol. 9, Newbury Park, Sage, 85 p.
- Ord, J. K., and A. Getis, 1995. Spatial autocorrelation statistics: Distribution issues and an application. *Geographical Analysis*, 27(4): 286-306.
- Pace, R. K., R. Barry and C. F. Sirmans, 1998. Spatial statistics and real estate. *The Journal of Real Estate Finance and Economics*, 17(1): 5-14.
- Panatier, Y., 1996. *Variowin: Software for Spatial Data Analysis in 2D*. *Statistics and Computing*, Springer-Verlag, New York, 91 p.
- Ripley, B. D., 1981. *Spatial Statistics*. New York, Wiley.
- Rummel, R.J., 1970. *Applied Factor Analysis*. Evanston, Northwestern University Press.
- Shefer, D., 1986. Utility changes in housing and neighbourhood services for households moving into and out of distressed neighbourhoods, *Journal of Urban Economics*, 19: 107-124.
- Strange, W., 1992. Overlapping neighbourhoods and housing externalities. *Journal of Urban Economics*, 32: 17-39.
- Thériault, M., and F. Des Rosiers, 1995. Combining hedonic modelling, GIS and spatial statistics to analyze residential markets in the Quebec Urban Community. In *Proceedings of the Joints European Conference on Geographical Information*, EGIS Foundation, March 26-31, The Hague, The Netherlands, 2: 131-136.

- Thérault, M., F. Des Rosiers and H. Vandermissen, 1999. GIS-Based Simulation of Accessibility to Enhance Hedonic Modelling and Property Value Appraisal. Joint IAAO-URISA 1999 Conference, New Orleans, April 11-14, 10 pages.
- Thurstone, L.L., 1947. Multiple Factor Analysis. Chicago, University of Chicago Press.
- Wong, D. and Amrhein, C., 1996. The Modifiable Areal Unit Problem. Special issue of *Geographical Systems*, vol. 3, no. 2-3.
- Yinger, J. H., H. S. Bloom, A. Borsch-Supan and H. F. Ladd, 1987. Property Taxes and House Values: The Theory and Estimation of Intra-jurisdictional Property Tax Capitalization. *Studies in Urban Economics*, Academic Press, Princeton Univ., 218 p.
- Zhang, Z., and D. Griffith, 1993. Developing user-friendly spatial statistical analysis modules for GIS: An example using ArcView. *Computer, Environment and Urban Systems*, 21(1): 5-29.
-

Authors not quoted:

- Golledge, R.G., M. P. Kwan and T. Gärling, 1994. Computational process modelling of household travel decision using a geographical information system. *Papers in Regional Science*, 73(2): 99-117.
- Goodchild, M. F., 1998. GIS and disaggregate transportation modelling. *Geographical Systems*, 5(1): 19-44.
- Goodman, A. C., and T. G. Thibodeau, 1995. Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, 6: 25-42.
- Grieson, R. E., and J. R. White, 1989. The existence and capitalization of neighbourhood externalities: A reassessment. *Journal of Urban Economics*, 25(1): 68-76.
- Guntermann, K. L., and P. F. Colwell, 1983. Property values and accessibility to primary schools. *Real Estate Appraiser and Analyst*, 49(1): 62-68.
- Hägerstrand, T., 1970. What about people in regional science? *Papers of the Regional Science Association*, 14: 7-16.
- Hotelling, H., 1933. Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, 24: 417-441; 498-520.
- Jackson, J. R., 1979. Intraurban variation in the price of housing. *Journal of Urban Economics*, 6: 464-479.
- King, Leslie J., 1969. *Statistical Analysis in Geography*, Prentice-Hall, Englewood Cliffs, N.J.
- Korem, 1998. MapLogix Version 2.0. Topological Map Editing Utility for MapInfo Professional. [HTTP://www.korem.com/maplogix/index.html](http://www.korem.com/maplogix/index.html)
- Landau, U., J. N. Prashker and M. Hirsh, 1981. The effect of temporal constraints in household travel behaviour. *Environment and Planning A*, 13: 435-444.

- Miller, H. J., 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, 5(3): 287-301.
- Niedercorn, J. H., and N. S. Ammari, 1987. New evidence on the specification and performance of neoclassical gravity models in the study of urban transportation. *The Annals of Regional Sciences*, 21(1): 56-64.
- Rodriguez, M., C. F. Sirmans and A. P. Marks, 1995. Using geographic information systems to improve real estate analysis. *The Journal of Real Estate*, 10(2): 163-173.
- Sirpal, R., 1994. Empirical modelling of the relative impacts of various sizes of shopping centers on the values of surrounding residential properties. *Journal of Real Estate Research*, 9(4): 487-505.
- Tabachnick, Barbara G. and L.S. Fidell, 1996. *Using Multivariate Statistics*. HarperCollins College Publishers, New York, p. 636.
- Tiefelsdorf, M., and B. Boots, 1997. A note on the extremities of local Moran's I_i and their impact on global Moran's I . *Geographical Analysis*, 29(3): 248-257.
- Thériault, M., D. Leroux and M. H. Vandersmissen, 1998. Modelling travel route and time within GIS: Its use for planning. In *Simulation Technology: Science and Art. Proceedings of the 10th European Simulation Symposium*. A. Bargiela and E. Kerckhoffs Eds., The Nottingham-Trent University and Society for Computer Simulation International, October 26-28, 402-407.
- Thrall, G. I., 1993. Using GIS to rate the quality of property tax appraisal. *Geo Info Systems*, 3(3): 56-62.
- Thrall, G. I., and A. P. Marks, 1993. Functional requirements of a geographic information system for performing real estate research and analysis. *Journal of Real Estate Literature*, 1(1): 49-61.
- Timmermans, H., and R. G. Golledge, 1990. Applications of behavioural research on spatial problems: preferences and choice. *Progress in Human Geography*, 14: 311-354.