SpAM

SpAM (Spatial Analysis and Methods) presents short articles on the use of spatial statistical techniques for housing or urban development research. Through this department of Cityscape, *the Office of Policy Development and Research introduces readers to the use of emerging spatial data analysis methods or techniques for measuring geographic relationships in research data. Researchers increasingly use these new techniques to enhance their understanding of urban patterns but often do not have access to short demonstration articles for applied guidance. If you have an idea for an article of no more than 3,000 words presenting an applied spatial data analysis method or technique, please send a one-paragraph abstract to rwilson@umbc.edu for review.*

Evaluating Spatial Model Accuracy in Mass Real Estate Appraisal: A Comparison of Geographically Weighted Regression and the Spatial Lag Model

Paul E. Bidanset
University of Ulster and the City of Norfolk, Virginia

John R. Lombard
Old Dominion University

Abstract

Geographically weighted regression (GWR) has been shown to greatly increase the performance of ordinary least squares-based appraisal models, specifically regarding industry standard measurements of equity, namely the price-related differential and the coefficient of dispersion (COD; Borst and McCluskey, 2008; Lockwood and Rossini, 2011; McCluskey et al., 2013; Moore, 2009; Moore and Myers, 2010). Additional spatial regression models, such as spatial lag models (SLMs), have shown to improve multiple regression real estate models that suffer from spatial heterogeneity (Wilhelmsson, 2002). This research is performed using arms-length residential sales from 2010 to 2012 in

Abstract (continued)

Norfolk, Virginia, and compares the performance of GWR and SLM by extrapolating each model's performance to aggregate and subaggregate levels. Findings indicate that GWR achieves a lower COD than SLM.

Introduction

Ad valorem property taxes are a prominent source of government revenue in jurisdictions around the world. Taxing authorities are held accountable to ensure that these valuations are fair and equitable. In such roles, the optimization of the accuracy of mass real estate valuation approaches is critical.

Because of their precision and time- and cost-saving advantages, real estate mass appraisal methods that employ multiple regression-based models, known as automated valuation models (AVMs), are becoming increasingly prominent in industry practice and have received attention from the academic community. AVMs are used in a host of industries—both public and private—including loan origination, fraud detection, and portfolio valuation (Downie and Robson, 2007), and are promoted and advanced by such organizations as the International Association of Assessing Officers (IAAO). Statistical standards of equity established by such organizations give additional benchmarks by which modelers may test various approaches and methodologies.

Academic research has expanded regression models using geographically specific dummy variables and distance coefficients, and, although this approach has been shown to improve ordinary least squares (OLS)-based regression models, they often still suffer from biased coefficients and t-scores (Berry and Bednarz, 1975; Fotheringham, Brunsdon, and Charlton, 2002; McMillen and Redfearn, 2010). Some researchers (Fotheringham, Brunsdon, and Charlton, 2002) have used geographically weighted regression (GWR), a locally weighted regression technique, which has improved model performance by employing a spatial weighting function and allowing for coefficients to fluctuate across geographic space (Huang, Wu, and Barry, 2010; LeSage, 2004). Similarly, the spatial lag model (SLM)—a spatial autoregressive (SAR) model—addresses spatial heterogeneity by including an autocorrelation coefficient and spatial weights matrix (Anselin, 1988).

Because real estate markets behave differently across geographic space, AVMs free of spatial consideration often produce inaccurate, misleading results (Anselin and Griffith, 1988; Ball, 1973; Berry and Bednarz, 1975). GWR is prominently demonstrated throughout literature as a more accurate alternative to multiple regression analysis (MRA) AVMs (for example, Borst and McCluskey, 2008; Lockwood and Rossini, 2011; McCluskey et al., 2013; Moore, 2009; Moore and Myers, 2010). Similarly, SAR models have been sufficiently demonstrated to increase the predictive accuracy of such models (Borst and McCluskey, 2007; Conway et al., 2010; Quintos, 2013; Wilhelmsson, 2002). Descriptions of their methods and findings are summarized in exhibit 1.

Despite the popularity of both GWR and SLM models in housing research, to our knowledge, a study that simultaneously compares the performance of GWR and SLM using industry-accepted IAAO standards and that extrapolates each model's performance to aggregate and subaggregate levels has yet to be published. Farber and Yeates (2006) found GWR to have more accuracy and produce less spatially biased coefficients than SAR models, but no comparison has been made of how each performs against the other in the context of mass appraisal for tax assessments. A major finding of Bidanset and Lombard (2013)[1] is that traditional measures of hedonic model performance (for example, the Akaike Information Criterion [AIC], $\overline{R^2}$) do not necessarily indicate which model will perform the best given the assessment industry standards of uniformity (that is, coefficient of dispersion [COD]).[2] This article compares spatial regression techniques of the SLM and GWR and compares not only their prediction accuracy ability but also their attainment of IAAO equity and uniformity standards. Given the increasing availability of Geographic Information System, or GIS, data and advances incomputational ability to perform spatial AVMs, the understanding of the capability that each method lends to governments in reaching more accurate value estimations is critical.

Exhibit 1

Select Survey of Previous SAR Real Estate Research

Paper	Methodology	Results/Conclusions
Wilhelmsson, 2002	Compared OLS, SAR, and SEM.	SAR model improves model predictability of OLS model with spatial dummies but does not correct for spatial dependency.
Borst and McCluskey, 2007	Compared OLS-based and GWR alternatives with CSM.	CSM methodology is similar to the weights matrix used in an SLM and reduces baseline COD more than specified GWR model.
Conway et al., 2010	Developed spatial lag hedonic model to capture price effects of urban green space.	SLM improves OLS performance by helping to account for spatial autocorrelation.
Quintos, 2013	Used SLMs to create location-based base prices and location adjustment factors.	Spatial lags significantly improve OLS model performance.

COD = coefficient of dispersion. CSM = comparable sales method. GWR = geographically weighted regression. OLS = ordinary least squares. SAR = spatial autoregressive. SEM = spatial error model. SLM = spatial lag model.

[1] The final paper of this research, Bidanset and Lombard (forthcoming), is scheduled to be published in the *Journal of Property Tax Assessment and Administration*, volume 11, issue 3.

[2] AIC is a commonly used goodness-of-fit test of models applied to the same sample. It has the following calculation:

$$AIC_i = -2logL_i + 2K_i,$$

where L_i is the maximum likelihood of the ith model, and K_i is the number of free parameters of the ith model.

Model Descriptions and Estimation Details

The traditional OLS regression model is represented by

$$y_i = \beta_0 + \Sigma_k \beta_k x_{ik} + \varepsilon_i, \tag{1}$$

where y_i is the ith sale, β_0 is the model intercept, β_k is the kth coefficient, x_{ik} is the kth variable for the ith sale, and ε_i is the error term of the ith sale. The GWR extension is depicted by the following—

$$y_i = \beta_0(u_i, v_i) + \Sigma \beta_k(u_i, v_i) x_{ik} + \varepsilon_i, \tag{2}$$

where (u_i, v_i) indicates the latitude-longitude (xy) coordinates of the ith regression point. GWR creates a local regression allowing coefficients to vary at each observation. In this article, the xy coordinates of the respective sale represent each observation.

In matrix notation, the OLS model and GWR model are represented by equations 3 and 4, respectively.

$$Y = X\beta + \varepsilon, \text{ and} \tag{3}$$

$$Y = (\beta \otimes X)1 + \varepsilon, \tag{4}$$

where $\otimes$ denotes a logical multiplication operator; β is multiplied by the respective and corresponding value of X. This differentiates GWR from the constant vector of parameters (β) of the OLS model.

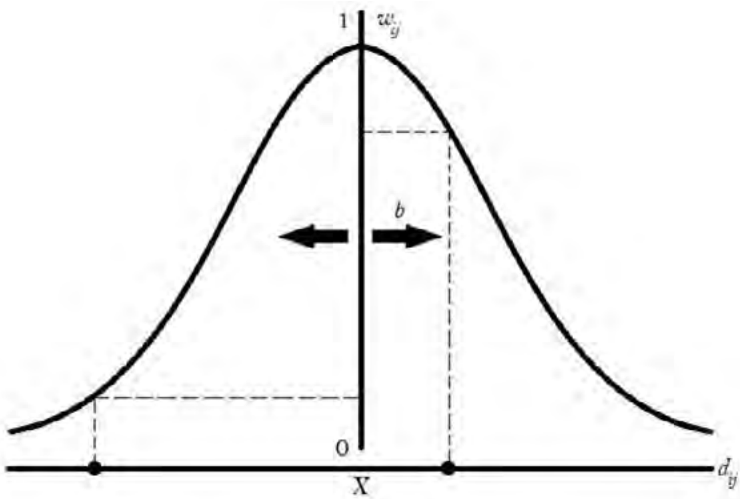
The GWR model will employ a Gaussian spatial kernel and a fixed bandwidth. Bidanset and Lombard (forthcoming) show that kernel and bandwidth combinations should be examined during the model calibration phase—specifically regarding effect on IAAO ratio study standards—to examine which produces the optimal results. With the current variables and data, the Gaussian kernel with a fixed bandwidth achieves the lowest COD and is used in comparison against other spatial weighting functions tested (that is, bisquare kernel with adaptive bandwidth, bisquare kernel with fixed bandwidth, and Gaussian kernel with adaptive bandwidth).

During model calibration, the fixed bandwidth used in the GWR model is selected by a procedure that identifies the bandwidth that will achieve the lowest AIC corrected value (Fotheringham, Brunsdon, and Charlton, 2002).

The Gaussian kernel incorporates a distance decay function that places a higher weight on properties more closely situated to the observation point (exhibit 2).

Exhibit 2

Spatial Kernel Used in Geographically Weighted Regression



where
X is the regression point,
● is a data point,
w_{ij} is the weight applied to the jth property at regression point i,
b is the bandwidth, and
d_{ij} is the geographic distance between regression point i and property j.

Source: Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, United Kingdom: John Wiley & Sons*

Gaussian weight—

$$w_{ij} = exp\ [-1/2(d_{ij}/b)^2]. \qquad (5)$$

The SLM is represented by the following equation (Borst and McCluskey, 2007; Can, 1992)—

$$Y = \rho WY + X\beta + \varepsilon, \qquad (6)$$

where W is a spatial weights matrix indicating distance relationship between observations i and j. The weights matrix establishes the effect nearby observations have on the subject property. The spatially lagged dependent variable is represented by the coefficient ρ. The weights matrix and the spatially lagged dependent variable help capture "spillover" effects from neighboring observations. In this article, a nearest neighbor matrix is derived to create a row standardized weights matrix.

Equity and Uniformity Measurement Standards

IAAO created and maintains standards that promote equity and fairness in real estate appraisals and assessments. The COD and the price-related differential (PRD) are two coefficients by which accuracy and fairness are measured.

For single-family homes, the IAAO set a maximum acceptability value of 15.0 for COD scores (IAAO, 2013). Values under 5.0 are indications of sales-chasing (cherry-picking sales that will produce optimal results) or sampling error (properties and areas more difficult to model are underrepresented; IAAO, 2013). The COD calculation is as follows—

$$COD = \frac{100}{n} \frac{\sum_{i=1}^{n} \left| \frac{EP_i}{SP_i} - Median\left(\frac{EP_i}{SP_i}\right) \right|}{Median\left(\frac{EP_i}{SP_i}\right)}, \tag{7}$$

where EP_i is the expected price of the ith property, and SP_i is the sales price of the ith property. The price-related differential is a score measuring vertical equity, represented by equation 8.

$$PRD = \frac{Mean\left(\frac{EP_i}{SP_i}\right)}{\sum_{i=1}^{n} EP_i \ / \ \sum_{i=1}^{n} SP_i}. \tag{8}$$

According to the IAAO Standard on Automated Valuation Models, PRD values of less than 0.98 suggest evidence of progressivity, while PRD values of more than 1.03 suggest evidence of regressivity (IAAO, 2003).

The Data and Variables

The data comprise 2,450 arms-length single-family home sales in Norfolk, Virginia, from 2010 to 2012 and their respective characteristics at the time of sale. City assessment staff review all transfers of real estate within the city of Norfolk and an unbiased third party confirms them. An arms-length transaction requires that neither party be under duress to buy or sell, the property is listed openly, and no previous relationship or affiliation exists between the buyer and the seller. Because assessment offices are required by law to value properties at fair market value—and non-arms-length transactions, such as foreclosures and short sales, do not necessarily reflect the true market—only arms-length transactions are included in the analysis. To promote the accuracy of results, outliers are identified and omitted using an IQRx3 approach (removing about 2 percent of observations). Furthermore, to reduce the likelihood of skewed results, observations are inspected to ensure no egregious errors, such as buildings with zero total living area, are present.

Exhibit 3 shows a list of the independent variables and their respective descriptions. *TLA* is the total area (in square feet) of livable space (excluding, for example, unfinished attics). *TGA* is total garage area (in square feet) of attached and detached garages. *Age* is the age of the building (in years). Regarding improvements built around the same time, the effective age (*EffAge*) represents the state of cured depreciation (Gloudemans, 1999). Each of these four variables is transformed to natural log form to allow for nonlinear relationships, such as diminishing marginal returns to price. A dummy variable *bldgcond* is included for the condition of the improvement, with a default of average. Using the reverse month of sale (*RM1* through *RM36*), 11 time-indicator 3-year linear spline variables are created, with *RM1* denoting the most recent month of sale and *RM36* denoting the oldest month of sale). Linear spline variables offer significantly more explanatory power than monthly, quarterly, or seasonally based variables (Borst, 2013). *RM12* and *RM21* improved model performance significantly and are included in the exhibit.

Ln.ImpSalePrice is the dependent variable, which is calculated by first subtracting the respective assessed land value from each sale price and then transforming this value to its natural logarithm. This method attempts to isolate the effects of the independent variables on the improvement alone (Moore and Myers, 2010).

Exhibit 3

Independent Variables

Variable	Description
ln.TLA	Total living area in square feet (natural log)
ln.EffAge	Effective age in years (natural log)
ln.Age	Age in years (natural log)
ln.TGA	Total garage area in square feet, detached + attached (natural log)
bldgcond	Condition of building (average is default)
RM12	12th reverse month spline variable
RM21	21st reverse month spline variable

Results

GWR achieves the most uniform results with the lowest COD of 9.12 (exhibit 4). The SLM follows with a COD of 10.86. Both models outperform the global model (12.51) with respect to uniformity. None of the models exceeds the IAAO maximum threshold of 15.00. PRD, although the highest with global (1.03) and the lowest with GWR (1.01), does not change very much across the three models. No model suggests evidence of regressivity or progressivity, although the global model is at the highest acceptable limit set by IAAO standards (1.03) before evidence of regressivity becomes present.

Across these models, rank of AIC is the same as rank of COD and PRD (exhibit 5).

Exhibit 4

Model Performance Results

Method	AIC	COD	PRD
Global	324.52	12.51	1.03
SLM	– 207.84	10.86	1.02
GWR	– 784.79	9.12	1.01

AIC = Akaike Information Criterion. COD = coefficient of dispersion. GWR = geographically weighted regression. PRD = price-related differential. SLM = spatial lag model.

Exhibit 5

Local R^2 Maps by Spatial Weighting Function

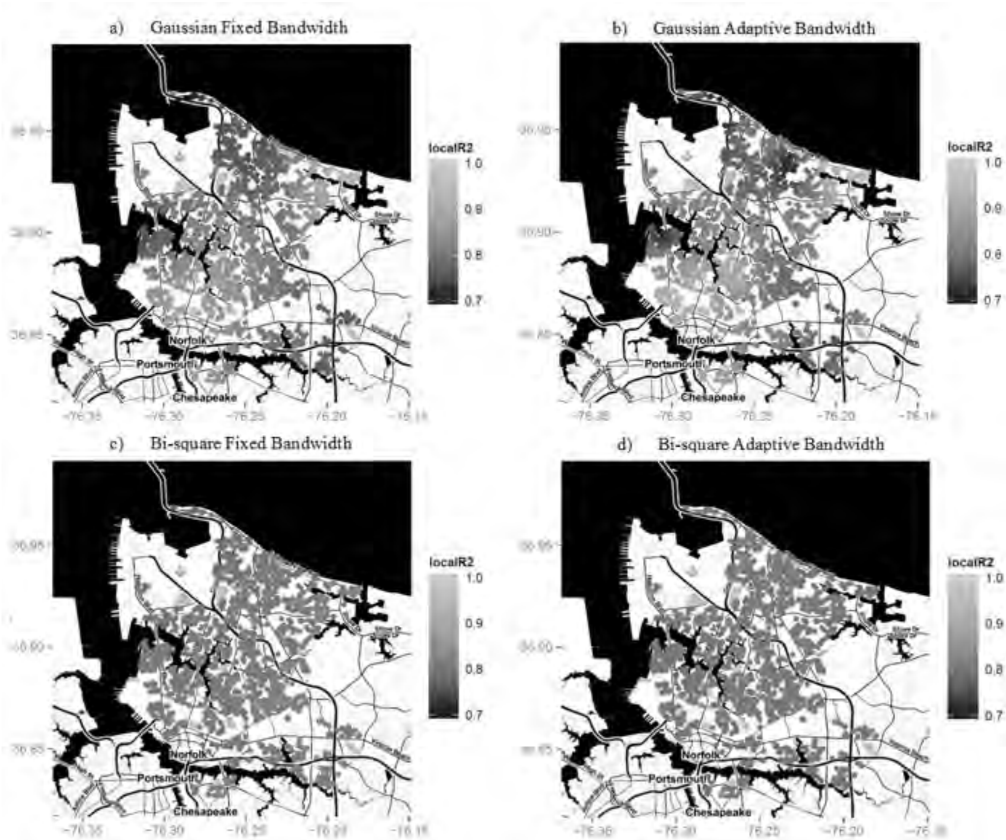
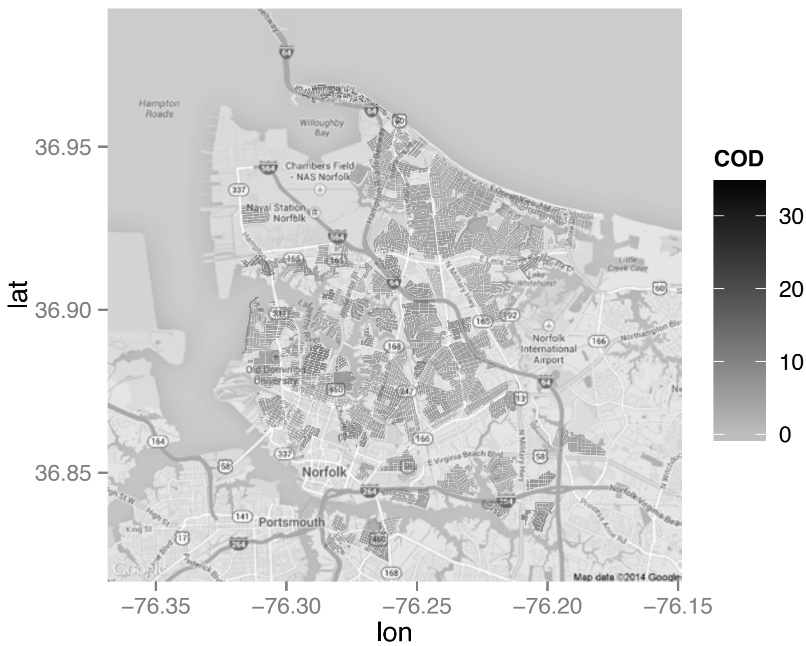


Exhibit 6 (three maps—6a, 6b, and 6c) shows the COD for each Norfolk neighborhood. These neighborhoods are identified by city authorities and are delineated by neighborhood shapefiles provided by the city. Because neighborhoods are on average composed of more similar homes (age, architecture, size, condition, proximity to various parts of the city, and so on), they serve as submarkets for further analysis and evaluation of model performance. Understanding how various models perform across neighborhoods of varying compositions enables modelers to calibrate modeling techniques that optimize individual submarkets. Because the geographic location of a

Exhibit 6

COD Disaggregated by Neighborhood (1 of 2)

(a) Global Results



(b) SLM Results

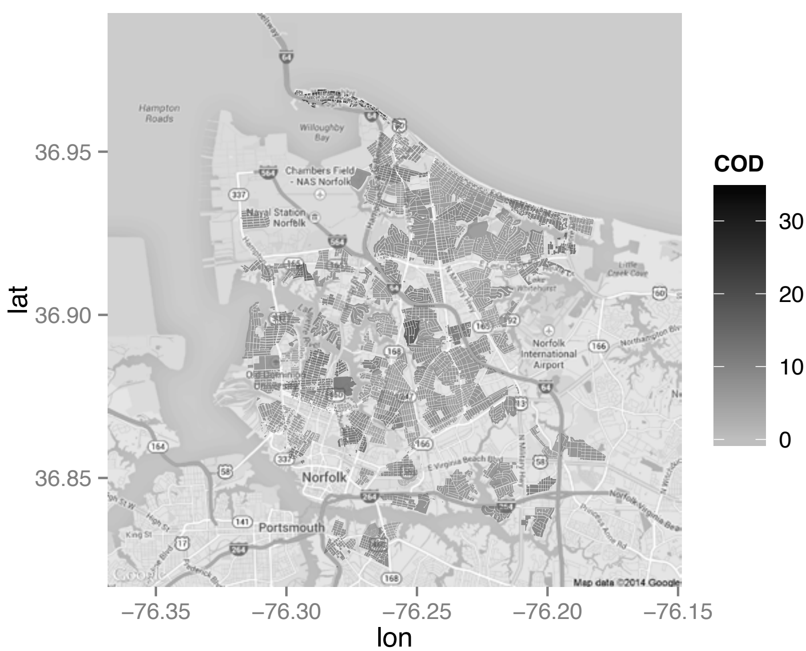
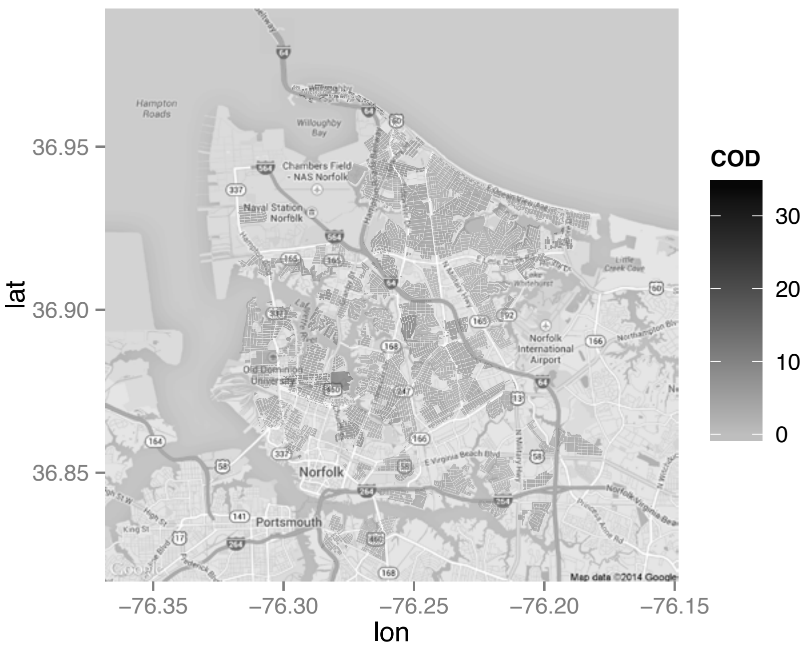


Exhibit 6

COD Disaggregated by Neighborhood (2 of 2)

(c) GWR Results



COD = coefficient of dispersion. GWR = geographically weighted regression. lat = latitude. lon = longitude. SLM = spatial lag model.

neighborhood can be correlated with socioeconomic and demographic conditions, such disaggregation enables assessors to further ensure all markets are treated without discrimination—yet another step toward promoting equitable valuations.

Darker shaded areas indicate higher COD values (decreased uniformity in value predictions) and lighter shaded areas represent lower COD values (increased uniformity in value predictions). The global model produces, overall, many dark gray- to black-shaded neighborhoods of low uniformity (exhibit 6a). The SLM model (exhibit 6b), although it alleviates only a few neighborhoods of high COD values, actually makes many neighborhoods worse.

The global model is more uniform than SLM (for example, at about [36.89, -76.25]), but the SLM outperforms the global model and GWR (exhibit 6c) directly to the east of Old Dominion University.

Exhibit 6c reveals the GWR model overall achieves a much smoother distribution of lower COD values, as evidenced by the lighter gray colors and less severe contrast of shades.

Although GWR achieves the lowest citywide COD, the global model outperforms GWR at about (36.95, -76.16). The global model and SLM outperform GWR at about (36.85, -76.255). Similar to findings of Bidanset and Lombard (forthcoming), this variation in COD suggests that, although a model achieves optimal aggregate results, it may still be outperformed within subaggregate geographic regions. Several areas, such as the northeastern peninsula labeled "Willoughby Spit," are drastically improved with GWR, and the COD is reduced to an IAAO-acceptable level (less than 15.00). Waterfront homes in neighborhoods are grouped into a separate neighborhood shapefile. In each map of exhibit 6, the waterfront homes in Willoughby Spit are significantly less uniform than the nonwaterfront homes.

Conclusions

Using arms-length residential sales from 2010 to 2012 in Norfolk, Virginia, this article compares the performance of GWR and SLM, specifically regarding IAAO levels of uniformity and equity at aggregate and subaggregate geographic levels. Findings suggest that GWR achieves more uniform results (lower COD) overall than SLM, and both achieve more uniform results than the spatially unaware global model. Although a model may produce optimal overall results, disaggregation into submarkets (for example, neighborhoods) reveals that it can still be outperformed within subgeographic areas by other models that produce inferior overall results. Compared with the global model, the SLM model actually increases the COD for a number of neighborhoods, despite having a lower overall citywide COD. This variation of models across geographic space supports findings of Bidanset and Lombard (2013) and suggests that modelers should explore various models' performance in various locations to optimize equity and uniformity in assessment jurisdictions overall.

Furthermore, waterfront estimations of value are included in land values, which, as previous literature suggests, are subtracted from total value in an attempt to isolate the explanatory variables' effects on the price of the building only. The differences between waterfront and nonwaterfront properties' uniformity suggest that this method does not fully account for such effects and, therefore, should be included in the model, perhaps in the form of a dummy variable.

Further GWR- and SLM-performance research is needed. Variations in SLM weights matrix style, such as binary, global standardized, and variance stabilization, and their effect on COD and PRD could be examined. In addition, more research that uses different variable selections and different markets of varying size and characteristics could be explored. Temporal variations and weighting schemes should also be evaluated to measure potential effects on the behavior of spatial models.

Acknowledgments

The authors thank Bill Marchand (Chief Deputy Assessor) and Deborah Bunn (Assessor) of the City of Norfolk, Virginia, Office of the Real Estate Assessor, for the opportunity to conduct this research.

Authors

Paul E. Bidanset is a Ph.D. student at the University of Ulster, School of the Built Environment, Newtownabbey, United Kingdom, and a real estate CAMA modeler analyst for the City of Norfolk, Virginia, Office of the Real Estate Assessor.

John R. Lombard teaches graduate courses in research methods, urban and regional development, and urban resource allocation in the Department of Urban Studies and Public Administration at Old Dominion University (ODU) and serves as the Director of the ODU Center for Real Estate and Economic Development.

References

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Springer.

Anselin, Luc, and Daniel A. Griffith. 1988. "Do Spatial Effects Really Matter in Regression Analysis?" *Papers in Regional Science* 65 (1): 11–34.

Ball, M.J. 1973. "Recent Empirical Work on the Determinants of Relative House Prices," *Urban Studies* 10 (2): 213–233.

Berry, Brian J., and Robert S. Bednarz. 1975. "A Hedonic Model of Prices and Assessments for Single-Family Homes: Does the Assessor Follow the Market or the Market Follow the Assessor?" *Land Economics* 51 (1): 21–40.

Bidanset, Paul E., and John R. Lombard. Forthcoming. "The Effect of Kernel and Bandwidth Specification in Geographically Weighted Regression Models on the Accuracy and Uniformity of Mass Real Estate Appraisal," *Journal of Property Tax Assessment and Administration*.

———. 2013. "Optimal Spatial Weighting Functions of Geographically Weighted Regression Models Used in Mass Appraisal of Residential Real Estate." Paper presented at the International Geographic Union Conference 2013: Applied GIS and Spatial Modelling. Leeds, United Kingdom, May 30.

Borst, Richard A. 2013 (April). *Optimal Market Segmentation and Temporal Methods: Spatio-Temporal Methods for Mass Appraisal*. Fairfax, VA: International Property Tax Institute.

Borst, Richard A., and William J. McCluskey. 2008. "Using Geographically Weighted Regression To Detect Housing Submarkets: Modeling Large-Scale Spatial Variations in Value," *Journal of Property Tax Assessment and Administration* 5 (1): 21–51.

———. 2007. "Comparative Evaluation of the Comparable Sales Method With Geostatistical Valuation Models," *Pacific Rim Property Research Journal* 13 (1): 106–129.

Can, Ayse. 1992. "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics* 22 (3): 453–474.

Conway, Delores, Christina Q. Li, Jennifer Wolch, Christopher Kahle, and Michael Jerrett. 2010. "A Spatial Autocorrelation Approach for Examining the Effects of Urban Greenspace on Residential Property Values," *The Journal of Real Estate Finance and Economics* 41 (2): 150–169.

Downie, Mary Lou, and Gill Robson. 2007. *Automated Valuation Models: An International Perspective*. London, United Kingdom: The Council of Mortgage Lenders.

Farber, Steven, and Maurice Yeates. 2006. "A Comparison of Localized Regression Models in a Hedonic House Price Context," *Canadian Journal of Regional Science* 29 (3): 405–420.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, United Kingdom: John Wiley & Sons.

Gloudemans, Robert J. 1999. *Mass Appraisal of Real Property*. Chicago: International Association of Assessing Officers.

Huang, Bo, Bo Wu, and Michael Barry. 2010. "Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices," *International Journal of Geographical Information Science* 24 (3): 383–401.

International Association of Assessing Officers (IAAO). 2013. *Standard on Ratio Studies*. Kansas City, MO: International Association of Assessing Officers.

———. 2003. *Standard on Automated Valuation Models (AVMs)*. Chicago: International Association of Assessing Officers.

LeSage, James P. 2004. "A Family of Geographically Weighted Regression Models." In *Advances in Spatial Econometrics*, edited by Luc Anselin, Raymond Florax, and Sergio Rey. Berlin, Germany: Springer-Verlag: 241–264.

Lockwood, Tony, and Peter Rossini. 2011. "Efficacy in Modelling Location Within the Mass Appraisal Process," *Pacific Rim Property Research Journal* 17 (3): 418–442.

McCluskey, W.J., M. McCord, P.T. Davis, M. Haran, and D. McIlhatton. 2013. "Prediction Accuracy in Mass Appraisal: A Comparison of Modern Approaches," *Journal of Property Research* 30 (4): 239–265.

McMillen, Daniel P., and Christian L. Redfearn. 2010. "Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions," *Journal of Regional Science* 50 (3): 712–733.

Moore, J. Wayne. 2009. "A History of Appraisal Theory and Practice Looking Back From IAAO's 75th Year," *Journal of Property Tax Assessment & Administration* 6 (3): 23–50.

Moore, J. Wayne, and Josh Myers. 2010. "Using Geographic-Attribute Weighted Regression for CAMA Modeling," *Journal of Property Tax Assessment & Administration* 7 (3): 5–28.

Quintos, Carmela. 2013. "Spatial Weight Matrices and Their Use As Baseline Values and Location-Adjustment Factors in Property Assessment Models," *Cityscape* 15 (3): 295–306.

Wilhelmsson, Mats. 2002. "Spatial Models in Real Estate Economics," *Housing, Theory and Society* 19 (2): 92–101.

Additional Reading

Dubin, Robin, R. Kelley Pace, and Thomas G. Thibodeau. 1999. "Spatial Autoregression Techniques for Real Estate Data," *Journal of Real Estate Literature* 7 (1): 79–96.

Fotheringham, A. Stewart, Martin E. Charlton, and Chris Brunsdon. 1998. "Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis," *Environment and Planning A* 30: 1905–1927.

———. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity," *Geographical Analysis* 28 (4): 281–298.