2014

The Effect of Kernel and Bandwidth Specification in Geographically Weighted Regression Models on the Accuracy and Uniformity of Mass Real Estate Appraisal

Paul E. Bidanset

John R. Lombard
Old Dominion University, jlombard@odu.edu

Original Publication Citation
Bidanset, P. E., & Lombard, J. R. (2014). 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 & Administration, 10*(3), 5-14.

The Effect of Kernel and Bandwidth Specification in Geographically Weighted Regression Models on the Accuracy and Uniformity of Mass Real Estate Appraisal

BY PAUL E. BIDANSET AND JOHN R. LOMBARD, PH.D.

Local governments have the responsibility of fairly and uniformly taxing the properties within their jurisdictions, and they must be held accountable for the taxes levied upon property owners. Therefore, it is imperative that residential property assessments be accurate, fair, and defensible. In recent years, great strides have been made in the advancement of mass-appraisal techniques such as automated valuation models (AVMs). It has long been understood that heterogeneity across geographic stratums hinders conventional ordinary-least-squares-based multiple-regression analysis (MRA) models from accurately capturing variables' true effects (Ball 1973; Berry and Bednarz 1975; Anselin and Griffith 1988). While spatial consideration in the form of dummy variables and distance coefficients can help improve models, these techniques may fail to fully correct for spatial autocorrelation, and parameter averages may be skewed or cancel each other out as a result (Berry and Bednarz 1975; Fotheringham, Brunsdon, and Charlton 2002; McMillen and Redfearn 2010). Inaccuracy in parameter estimation in assessment models can lead to unfair valuation of properties creating a host of challenges for the taxing jurisdiction, not just the likelihood of additional costs in time and money spent defending valuations.

Sufficient research has shown locally weighted regression (LWR) methods improve traditional valuation model performance and predictability power (e.g., Brunsdon, Fotheringham, and Charlton 1996; McMillen 1996; Fotheringham, Charlton, and Brunsdon 1998). Geographically weighted regression (GWR) is one such LWR methodology that more accurately accounts for spatial

---

Paul E. Bidanset *is the Real Estate CAMA Modeler Analyst for the Office of the Real Estate Assessor in Norfolk, Virginia. He is concurrently pursuing a doctorate at the University of Ulster School of the Built Environment in Newtownabbey, Northern Ireland, United Kingdom. His doctoral research focuses on real estate valuation modeling. He holds a master's degree in economics from Old Dominion University, Norfolk, Virginia, and a bachelor's degree in economics from James Madison University, Harrisonburg, Virginia*

John R. Lombard, Ph.D., *is an associate professor at Old Dominion University where he teaches graduate courses in research methods, urban and regional development, and urban resource allocation. He also is Director of the ODU Center for Real Estate and Economic Development. He received his Ph.D. in economic geography from the State University of New York at Buffalo.*

heterogeneity (Fotheringham, Brunsdon, and Charlton 2002; LeSage 2004; Huang, Wu, and Barry 2010). The use of GWR in property tax modeling has become an area of study as well. GWR has been shown to provide assessment jurisdictions with more accurate valuations than MRA or other AVM techniques (Borst and McCluskey 2008; Moore 2009; Moore and Myers 2010; Lockwood and Rossini 2011; McCluskey et al. 2013). Lockwood and Rossini (2011) state GWR favorably reduces prediction errors that arise from edge effects of boundaries in global models, and that GWR-based models are more "in-tune with the market." Borst and McCluskey (2008) demonstrate the ability of GWR to detect submarkets within jurisdictions.

While GWR has been shown to improve upon several standard mass valuation methods, it is a relatively new technique in the appraisal community, and some researchers suggest the need for additional studies to further establish GWR's credibility (Lockwood and Rossini 2011). Therefore, further research aimed at evaluating and understanding GWR performance in valuation is necessary.

The performance of kernel and bandwidth specification within GWR models has been explored in other disciplines—namely forestry and ecology (e.g., Guo, Ma, and Zhang 2008; Cho, Lambert, and Chen 2010). Thus far, optimal bandwidth/kernel combinations have not been examined side-by-side with respect to their potential impact on the statistical measures of equity and fairness as promulgated by the International Association of Assessing Officers (IAAO 2003). This research examines variations in GWR model performance across Gaussian and bi-square kernels with both fixed and adaptive bandwidths. Using residential data provided by the City of Norfolk, Virginia, this research evaluates the potential of GWR weighting specifications to further promote assessment fairness and equity.

Model Descriptions and Estimation Details

Geographically Weighted Regression Model

The traditional ordinary least squares (OLS) regression model is represented by

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i$$

where y_i is the i-th sale, β_0 is the model intercept, β_k is the k-th coefficient, x_{ik} is the k-th variable for the i-th sale, and ε_i is the error term of the i-th sale. The extension to OLS referred to as geographically weighted regression (Fotheringham, Brunsdon, and Charlton 2002) is depicted by the formula:

$$y_i = \beta_0(x_i, y_i) + \sum \beta_k(x_i, y_i)\, x_{ik} + \varepsilon_i$$

where (x_i, y_i) indicates the x,y coordinates of the i-th regression point. The x,y coordinates for this model are the latitude and longitude coordinates of each sale. GWR generates a regression equation for each observation weighted by location which takes into account spatially varying relationships.

Spatial Weighting Specifications—Kernels and Bandwidths

The kernels employed and examined in this research are the Gaussian kernel and bi-square kernel. Both kernels incorporate a distance decay function which allocates more weight to properties closer to a regression point than properties farther away (see figure 1). The bi-square kernel assigns a weight of zero to observations outside of the bandwidth, nullifying their impact on the local regression estimate.

The Gaussian weight and the bi-square weight are depicted as follows:

Gaussian Weight

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

Bi-square Weight

$$w_{ij} = [1-(d_{ij}/b)^2]^2 \text{ if } d_{ij} < b$$
$$= 0 \text{ otherwise.}$$

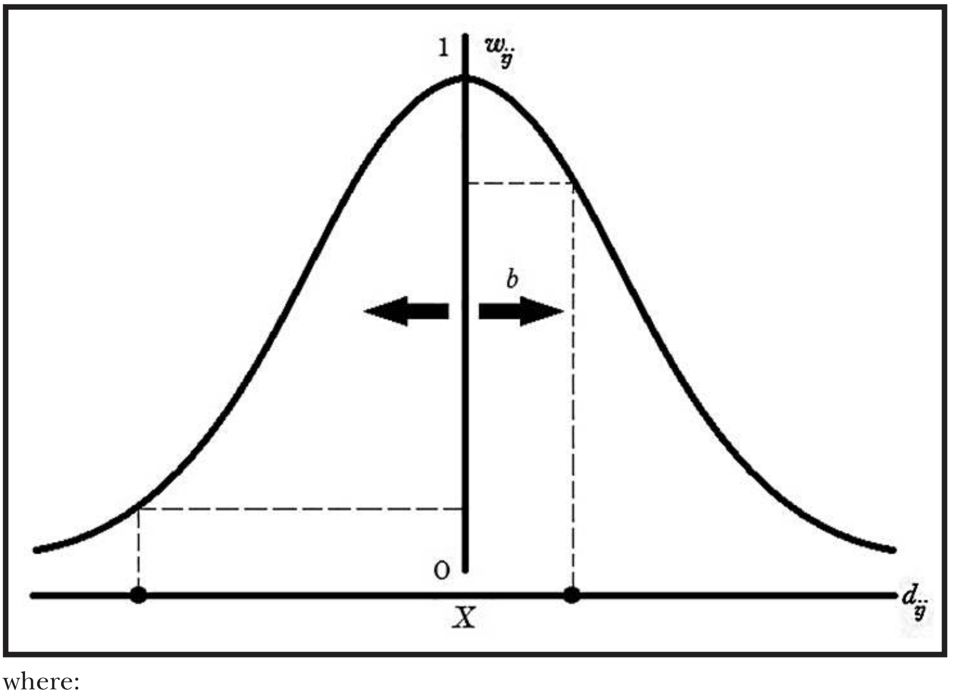
The Gaussian and bi-square kernels are tested using both fixed and adaptive bandwidths. In GWR, the size of the bandwidth is optimized by either distance (fixed kernel) or the number of neighboring observations (adaptive kernel). For example, the adaptive kernel selection process takes into account the density of observations and returns a value at each regression point for the optimal proportion of neighboring observations. During model calibration, bandwidths are tested and assigned cross validation (CV) scores; the bandwidth with the lowest CV score produces the lowest root mean square prediction error (Cleveland and Devlin 1988).

Measuring Vertical Equity and Uniformity in Valuation Models

As previously mentioned, the International Association of Assessing Officers (IAAO) maintains standards to provide a systematic means for determining assessment performance. These standards designate the coefficient of dispersion (COD) and the price-related differential (PRD) as the measures for evaluating assessment uniformity and equity, respectively (IAAO 2003). The COD quantifies uniformity of an appraisal stratum and is represented by the following formula:

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

Figure 1. The concept of a 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 j-th property at regression point i

b is the bandwidth

d_{ij} is the geographic distance between regression point *i* and property *j* (Fotheringham, Brunsdon, and Charlton 2002) where EP_i is the expected price of the i-th property, and SP_i is the sale price of the i-th property. The acceptability threshold for single-family homes set forth by IAAO is a COD value of 15.0, although values of 5.0 or less are suspect of sampling error or sales chasing (Gloudemans and Almy 2011).

The PRD is a coefficient of vertical equity and is calculated as follows:

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

The acceptability range set forth by IAAO for a stratum's PRD is between 0.98 and 1.03. Values above this range are evidence of regressivity; values below are evidence of progressivity (Gloudemans 1999).

The Akaike Information Criterion (AIC) is a commonly used measure of the relative performance of models. It is applied to the same sample and has the following calculation:

$$AIC_i = -2logL_i + 2k_i$$

where L_i is the maximum likelihood of the i-th model, and k_i is the number of free parameters of the i-th model.

The Data

The test data consisted of 2,450 valid single-family home sales in Norfolk, Virginia, from 2010 to 2012. (Using three years of sales is a recommended practice in the field of mass appraisal.) Valid sales must meet several criteria: the sale must be an arm's-length transaction where neither party is under duress to buy or sell; the property must be listed on the open market; and there can be no marital, blood, or previous relationship between the buyer and the seller. After a sale is completed and the new deed is registered with the Norfolk real estate assessor's office, a city appraiser pursues unbiased third-party verification of the conditions of the sale and the property characteristics. Valid sales and other types of property transfers (e.g., foreclosures, short sales, and the like) are marked accordingly. Invalid transfers, such as foreclosures, short sales, and government sales, were omitted from this analysis because they may not reflect the property's true market value—what assessors are required by law to determine.

Errors in data entry can potentially create outliers that will result in inaccurate models. For this reason, included observations met two criteria: a sale price greater than zero and properties with a positive net improvement sale price (sale price – assessed land value > 0). The sale price was converted to its natural logarithm and outliers were identified using an InterquartileRange×3 approach. In the IQR×3 method, the interquartile range (Q3-Q1) is calculated and multiplied by three. This value is then subtracted from the first quartile value (Q1) and added to the third quartile value (Q3) to create lower and upper

Table 1. Independent variables

Variable	Description
TLA	total living area (in square feet)
TLA2	total living area squared
TGA	total garage area (in square feet, detached + attached)
TGA2	total garage area squared
bldgcondFair	fair condition (average condition is default)
bldgcondGood	good condition
qualityclassFair	fair quality (average quality is default)
qualityclassGood	good quality
qualityclassVGd	very good quality
EffAge	effective age in years
EffAge2	effective age squared
Age	age in years
Age2	age squared
ForeclosureRatio	respective annual neighborhood ratio of foreclosures to valid transfers
RM12	reverse month of sale—spline 12
RM21	reverse month of sale—spline 21

bounds, respectively. Values outside of these bounds are outliers. Approximately two percent of observations were removed as a result of this procedure.

The dependent variable, *Ln.ImpSalePrice*, is the natural log of the selling price of the house with the land market value removed. Moore and Myers (2010) utilize this subtraction of the assessed land value by treating it as an offset to help isolate the coefficients' impact on the improved property only. The transformation of the dependent variable into a natural log helps promote normal distribution among explanatory variables and allows for results to be measured in percentage terms as opposed to dollars. The predicted values are transformed from natural log form, and land value is added back in, prior to performing ratio tests (COD and PRD).

Table 1 consists of the independent variables and their descriptions. *TLA* is included in the model because house size is positively correlated with price. The inclusion of *TLA2* accounts for a nonlinear relationship of diminishing marginal returns to value. *TGA* includes square feet of both attached and detached garage space and is consistent with Moore and Myers (2010). *TGA2*, like *TLA2*, is a squared transformation to account for diminishing marginal returns. Dummy variables are used for *bldgcond* and *qualityclass*. The default for each is the "average" rating. The effective age (*EffAge*) is used more often than age because it takes into account the overall state of the cured depreciation relative to other improvements built around the same time (Gloudemans 1999). *EffAge2* is included in case there is a nonlinear relationship between *EffAge* and the dependent variable. With building condition (*bldgcond*), quality (*qualityclass*), and effective age (*EffAge*) accounted for, *Age* is then added to capture any potential premium on vintage or historic properties with *Age2*, the squared value, inserted to capture any diminishing marginal returns to value.

ForeclosureRatio is the observation neighborhood's annual rate of foreclosures to valid sales. This variable is expected to have a negative coefficient because a higher density of foreclosures to valid sales would likely provide a supply of cheaper substitutes similar to the observation as well as lower the desirability of the area through the potential for blight from poorly maintained bank-owned properties. *ForeclosureRatio* is stored for all properties based on the respective neighborhood and can be used when models are applied to future subject properties. (More accurate methods of accounting for nearby foreclosures, including spatial and temporal consideration, can be developed but are not the focus of this article.)

For the three years of sales data, 11 time-indicator three-year linear spline variables were constructed based on the reverse month of sale (*RM1* through *RM36*), with the most recent month of sale equal to *RM1*, the second most recent month of sale equal to *RM2*, and consecutive months continuing to the oldest month, which was assigned *RM36*. Only splines *RM12* and *RM21* of the data improved model performance (i.e., a difference in AIC of at least 2) and were the only splines added to the model. Such linear spline variables are commonly found to offer more explanatory power to real estate valuation models than traditional quarterly or monthly dummy variables (Borst 2013).

Analyses and maps were executed using the R software environment.

Results

Table 2 displays the performance statistics of the four weighting functions, while figure 2 provides a graphical representation of each model's performance. The Gaussian fixed bandwidth achieved the lowest AIC (–595.27), followed closely by Gaussian adaptive bandwidth (–574.34). Bi-square fixed bandwidth and bi-square adaptive bandwidth yielded significantly higher AIC scores (–221.05 and –225.68, respectively) which approach that of the global model (–202.04). The PRD remained nearly the same (with a difference range of 0.003) regardless of the kernel or bandwidth employed.

Table 2. Performance of spatial weighting function

Model	AIC	PRD	COD
Global	–202.04	1.01	10.31
Gaussian Fixed Bandwidth	–595.27	1.01	9.07
Gaussian Adaptive Bandwidth	–574.34	1.01	8.88
Bi-square Fixed Bandwidth	–221.05	1.01	10.31
Bi-square Adaptive Bandwidth	–225.68	1.01	10.29

The lowest COD was 8.88 (Gaussian adaptive bandwidth), with poorer performing models reaching up to 10.31 (table 2). The weighting scheme generated by the bi-square kernel places nearly equal weights on all observations, as evidenced by the respective models' approach of the global COD (10.29 and 10.31 versus 10.31). The nearly constant local R^2 values in the bi-square models (figures 3c and 3d) suggest the bandwidths calculated are so large that they capture nearly all neighbors. Interestingly, while COD and AIC appear highly correlated, the weighting scheme which achieves the lowest COD, and thereby the highest uniformity, only has the *second* best AIC.

The plotted local R^2 maps in figure 3 demonstrate that weighting schemes that achieve superior overall AIC and COD scores do not necessarily achieve the highest predictability power within sub-geographic areas. (The locations of the following noteworthy neighborhood areas are provided in figure 4.) The bi-square kernels clearly outperform the Gaussian kernels in the Larchmont and Willoughby neighborhoods while the Gaussian kernels perform better in the East Ocean View (New) and Winona neighborhoods. Additionally, compared to Gaussian fixed bandwidth, Gaussian adaptive bandwidth alleviates predicted standard errors in more geographically isolated areas such as the Willoughby and Glenwood Park neighborhoods. Nevertheless, despite a lower overall COD, Gaussian adaptive bandwidth still produces higher predicted standard er-

Figure 2. Graphical representation of spatial weighting function performance

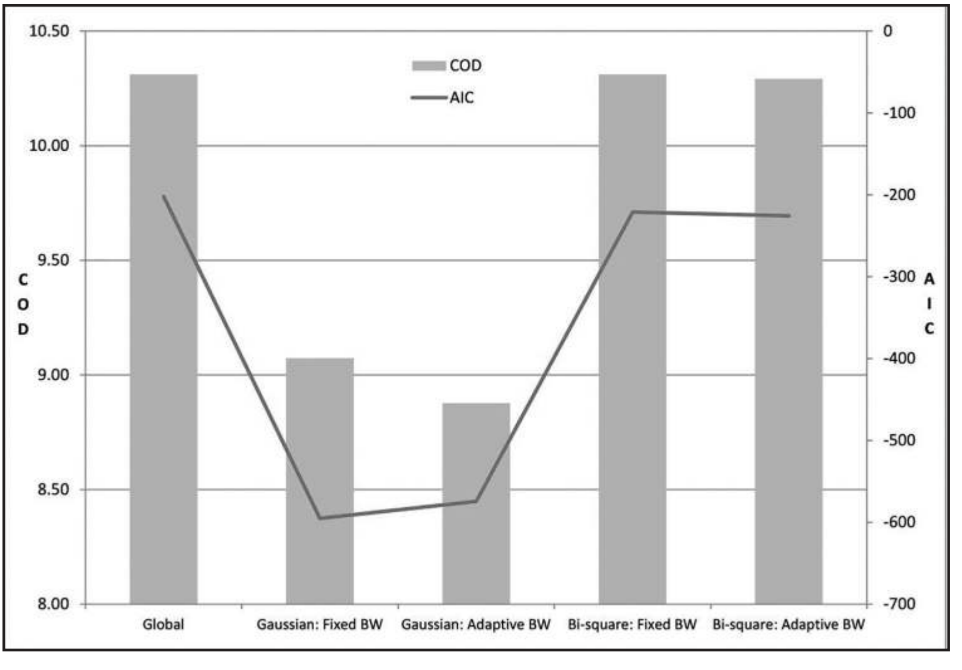


Figure 3. Local R^2 maps by spatial weighting function

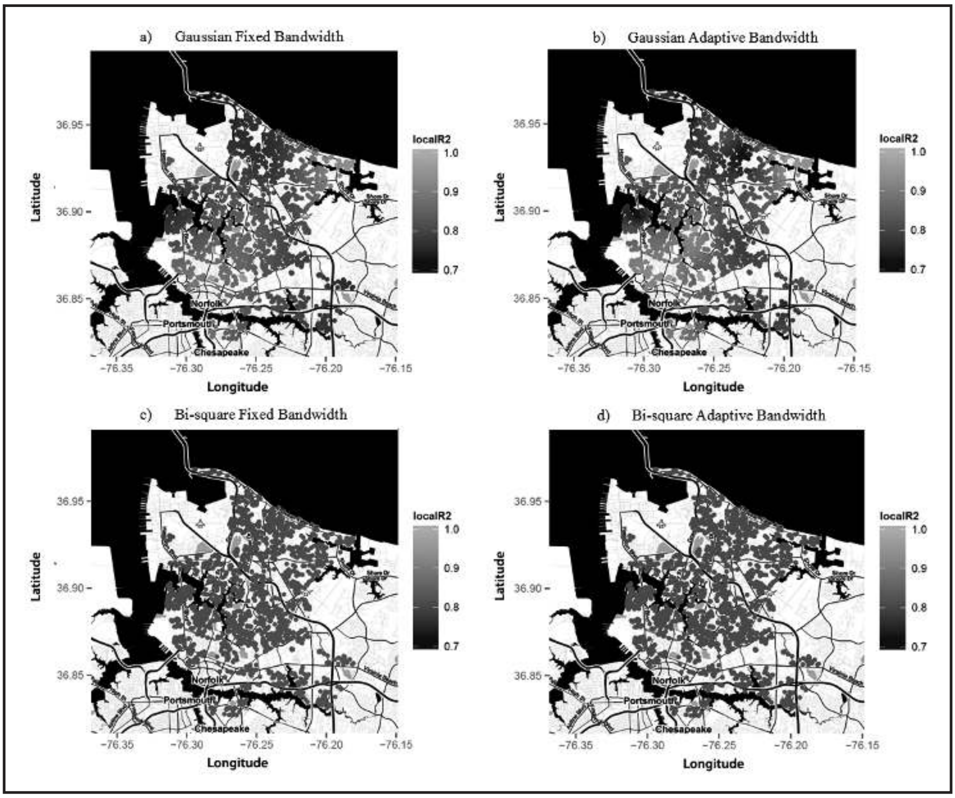
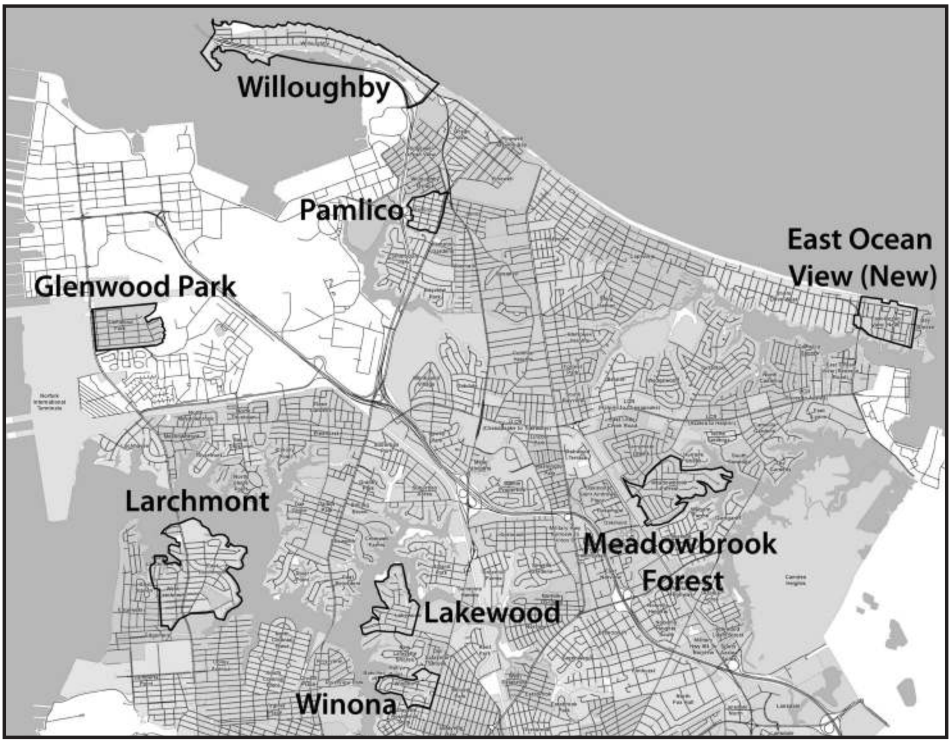


Figure 4. Location of neighborhoods featured in results comparisons



rors in other parts of the city (e.g., the Lakewood, Pamlico, and Meadowbrook Forest neighborhoods).

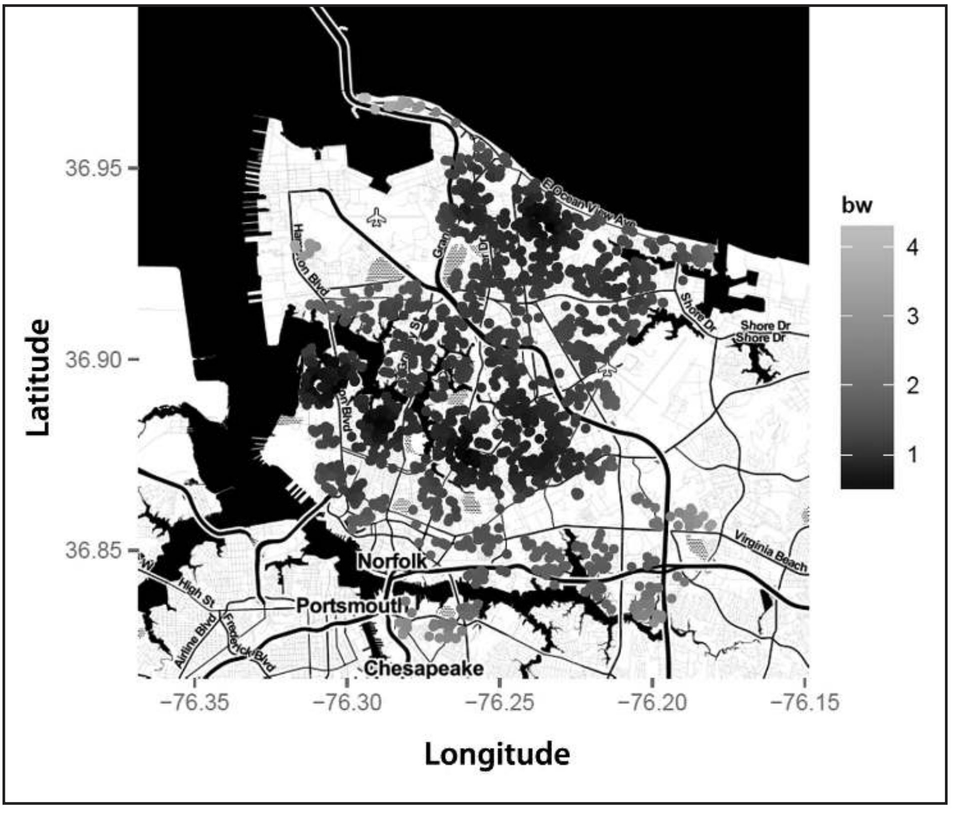
Figure 5 shows the amount of bandwidth that was used for Gaussian adaptive bandwidth at each observation. When applying GWR models to additional data, for example, in a holdout sample or later valuations, the geographically weighted data are applied to new fit points where kernel and bandwidth methods can be specified.

Conclusion

This research, using valid sales of single-family homes in Norfolk, Virginia, from 2010 to 2012, evaluated the varying predictability power that different kernel specifications in geographically weighted regression models lend to mass appraisal of real estate, and the potential improvement each lends to taxing entities in attaining equity, uniformity, and ultimately defensibility in their property assessments. Specifically, the weighting specifications that were studied were the Gaussian kernel with fixed bandwidths, the Gaussian kernel with adaptive bandwidths, the bi-square kernel with fixed bandwidths, and the bi-square kernel with adaptive bandwidths. The model applying a Gaussian kernel and adaptive bandwidths produced results that were most uniform by IAAO standards.

Appraisal uniformity is affected by the spatial weighting scheme chosen by the modeler. COD was shown to fluctuate with kernel specification, but PRD remained unaffected. The PRD results indicated none of the models suffered from vertical inequity, and even though COD scores differed, each model still achieved acceptable tax uniformity. Building upon previous research, these findings suggest that careful kernel and bandwidth specification in GWR models may greatly enhance taxing jurisdictions' ability to more efficiently reach unifor-

Figure 5. Bandwidth size (in kilometers)—Gaussian adaptive bandwidth



mity in their assessments, and thereby reduce the administrative and legal costs associated with inaccurate real estate valuations.

Traditional measures of model performance, such as the Akaike Information Criterion (AIC) and adjusted R^2, are most often the focus of GWR implications for real estate modeling. As this research suggests, modelers should extrapolate beyond indicators of model fit such as the AIC to include measures of uniformity and equity. Therefore, analysts should explore varying kernel and bandwidth combinations during the calibration phase of modeling. Furthermore, geographic disaggregation into local R^2 values reveals that a model which consistently produces superior overall results (lower COD, lower AIC) can still be outperformed within sub-geographic areas by a suboptimal city-wide aggregate model. It would behoove taxing entities to evaluate which weighting specifications perform best for each submarket, and subsequently stratify assessment models for geographical variations within a single taxing jurisdiction.

This article sets the stage for a wealth of additional research. The implications for optimal weighting specifications can be applied to other locally weighted regression techniques used in real estate modeling, such as temporal or attribute weighting. Other kernel functions (e.g., Epanechnikov, triangular, or uniform) can be included for analysis as well.

Acknowledgment

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

References

Anselin, L., and D.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, B.J., and R.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.

Borst, R. 2013. Optimal market segmentation and temporal methods. Spatio-temporal Methods in Mass Appraisal, seminar presented by the International Property Tax Institute, Fairfax, VA, April 23–25.

Borst, R.A., and W.J. McCluskey. 2008. Using geographically weighted regression to detect housing submarkets: Modeling large-scale spatial variations in value. *Journal of Property Tax Assessment & Administration* 5 (1): 21–51.

Brunsdon, C., A.S. Fotheringham, and M.E. Charlton. 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 28 (4): 281–298.

Cho, S.H., D.M. Lambert, and Z. Chen. 2010. Geographically weighted regression bandwidth selection and spatial autocorrelation: An empirical example using Chinese agriculture data. *Applied Economics Letters* 17 (8): 767–772.

Cleveland, W.S., and S.J. Devlin. 1988. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* 83 (403): 596–610.

Fotheringham, A.S., C. Brunsdon, and M. Charlton. 2002. Geographically weighted regression: The analysis of spatially varying relationships. Chichester, West Sussex, England: John Wiley & Sons.

Fotheringham, A.S., M.E. Charlton, and C. Brunsdon. 1998. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30 (11): 1905—1927.

Gloudemans, R.J. 1999. Mass appraisal of real property. Chicago: International Association of Assessing Officers.

Gloudemans, R., and R. Almy. 2011. Fundamentals of mass appraisal. Kansas City, MO: International Association of Assessing Officers.

Guo, L., Z. Ma, and L. Zhang. 2008. Comparison of bandwidth selection in application of geographically weighted regression: A case study. *Canadian Journal of Forest Research* 38 (9): 2526–2534.

Huang, B., B. Wu, and M. 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.

IAAO. 2003. Standard on automated valuation models (AVMs). *Assessment Journal* 10, no. 4 (Fall): 109–154.

LeSage, J.P. 2004. A family of geographically weighted regression models. In *Advances in spatial econometrics*, ed. by L. Anselin, R.J.G.M. Florax, and S.J. Rey, 241–264. Berlin: Springer-Verlag.

Lockwood, T., and P. Rossini. 2011. Efficacy in modelling location within the mass appraisal process. *Pacific Rim Property Research Journal* 17, No. 3 (September): 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, D.P. 1996. One hundred fifty years of land values in Chicago: A nonparametric approach. *Journal of Urban Economics* 40 (1): 100–124.

McMillen, D.P., and C.L. Redfearn. 2010. Estimation and hypothesis testing for nonparametric hedonic house price functions. *Journal of Regional Science* 50 (3): 712–733.

Moore, J.W. 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.W., and J. Myers. 2010. Using geographic-attribute weighted regression for CAMA modeling. *Journal of Property Tax Assessment & Administration* 7 (3): 5–28.