J Geograph Syst (2005) 7: 161–187
DOI: 10.1007/s10109-005-0155-6

ORIGINAL PAPER

David Wheeler · Michael Tiefelsdorf

Multicollinearity and correlation among local regression coefficients in geographically weighted regression

Received: 25 October 2004 / Accepted: 21 February 2005

Abstract Present methodological research on geographically weighted regression (GWR) focuses primarily on extensions of the basic GWR model, while ignoring well-established diagnostics tests commonly used in standard global regression analysis. This paper investigates multicollinearity issues surrounding the local GWR coefficients at a single location and the overall correlation between GWR coefficients associated with two different exogenous variables. Results indicate that the local regression coefficients are potentially collinear even if the underlying exogenous variables in the data generating process are uncorrelated. Based on these findings, applied GWR research should practice caution in substantively interpreting the spatial patterns of local GWR coefficients. An empirical disease-mapping example is used to motivate the GWR multicollinearity problem. Controlled experiments are performed to systematically explore coefficient dependency issues in GWR. These experiments specify global models that use eigenvectors from a spatial link matrix as exogenous variables.

This study was supported by grant number 1 R1 CA95982-01, Geographic-Based Research in Cancer Control and Epidermiology, from the National Cancer Institute. The author thank the anonymous reviewers and the editor for their helpful comments.

D. Wheeler (✉)
Department of Geography, The Ohio State University,
1036 Derby Hall, Columbus, OH 43210, USA
E-mail: wheeler.173@osu.edu

M. Tiefelsdorf
School of Social Sciences,
University of Texas at Dallas, Richardson, TX 75083, USA
E-mail: tiefelsdorf@utdallas.edu

Keywords Geographically weighted regression · Multicollinearity · Local regression diagnostics · Spatial eigenvectors · Experimental spatial design

1 Introduction

Geographically weighted regression (GWR) aims at identifying spatial heterogeneities in regression models of geo-referenced data. The spatial variability of the estimated local regression coefficients is usually examined to determine whether the underlying data generating process exhibits spatial heterogeneities or local deviations from a global regression model. A common procedure is to map the spatial GWR coefficient pattern associated with each exogenous variable. This approach, however, ignores potential dependencies among the local regression coefficients associated with different exogenous variables. Attention in this paper centers on these potential dependencies among the local coefficients. They can be expressed either as the correlation between pairs of local regression coefficients at *one location* or as the correlation between two *overall sets* of local coefficient estimates associated with two exogenous variables at all locations. Weak dependencies of either form hamper a substantive interpretation of the local GWR estimates, whereas strong dependencies induce artifacts that invalidate any meaningful interpretation and search for spatial heterogeneities because the regression coefficients are no longer uniquely defined.

The development of GWR started from local regression and smoothing techniques (Brunsdon et al. 1996; Fotheringham et al. 1998) and became increasingly more sophisticated by considering, for example, maximum likelihood estimation of the kernel bandwidths (Páez et al. 2002a), spatial autocorrelation among the residuals (Páez et al. 2002b), generalized linear model specifications (Fotheringham et al. 2002) and Bayesian GWR (LeSage 2001, 2004). Some diagnostic tests in GWR have also become more sophisticated, for instance, the development of formal test statistics for spatial nonstationarity and heterogeneity of the local model parameters (Leung et al. 2000). Aside from coefficient maps associated with single exogenous variables and local *t*-values, however, none of these developments pay much attention to fundamental regression diagnostics, such as residual analysis or the overall stability of the model. In general terms, model stability depends very much on the joint-distribution of the exogenous variables as demonstrated in the classical analysis by Longley (1967), where regression coefficients changed signs depending on whether specific exogenous variables or particular observations were in or excluded in the model. The uncertainties and numerical instabilities in this classical analysis are induced by the multicollinearity among the exogenous variables that lead to correlation among the estimated regression coefficients. In addition to these technical dimensions of multicollinearity, there are also more substantive ramifications affecting our ability to conduct an informed model interpretation: Fox (1997, p 351) states that "collinearity deals fundamentally with the inability to separate the effects of highly correlated variables" and Greene (2000, p 256)

highlights that "parameters are unidentified" and "different sets of parameters give the same $E(y_i)$."

Evaluating data in GWR for local multicollinearities and pairwise correlations between sets of local coefficients is even more important than in the traditional global regression model due to the increased complexities of the GWR estimation procedure that potentially induces interrelationships among the local estimates. In a worst-case scenario, these coefficient interrelationships lead to parameter redundancies, which clearly invalidate any attempt to interpret a single GWR coefficient independent of the remaining local estimates at the same location. Leung et al. (2000, p 14) point out that "the GWR technique in fact exerts some constraints on the parameters by the weighted least squares method." These constraints have the tendency to tie the GWR coefficient estimates locally and spatially together. A new class of GWR models would be required to formally express the effects of these constraints. This class would need to connect all location specific GWR regression equations simultaneously into a seemingly unrelated regression model, which allows jointly testing across several local models and a nested specification of a global model with local parameter variations.

Commonly used exploratory tools to uncover potential multicollinearity among the exogenous variables are bivariate correlation coefficients and bivariate scatter plots of pairs of exogenous variables. More statistical oriented tools that adopt a simultaneous view to diagnose multicollinearity in a fitted regression model are variance inflation factors (VIF), the correlation matrix of the estimated regression parameters (including the model intercept), and the condition index and variance-decomposition proportions by Belsley et al. (1980). Additional consequences of potential multicollinearity in a regression model are (a) a large change or even reversal in sign in one regression coefficient after another exogenous variable is added to the model or specific observations have been excluded from the analysis, (b) a counterintuitive sign in one regression coefficient, and (c) large parameter standard errors. It is essential to look for these effects of multicollinearity in global models and their local GWR counterparts when fitting and interpreting a GWR model. Furthermore, while a global regression model may have acceptable coefficient correlation levels and VIF levels, its GWR counterpart may have unacceptably high levels of correlation among the local GWR coefficients. These effects will be overlooked without a proper diagnostic analysis and may lead to an improper interpretation of the local regression model.

A critical reexamination of the basic properties of GWR becomes even more important since GWR is being established as a standard tool in exploratory spatial data analysis. Exploratory spatial data analysis should describe, on the one hand, the complex nature of spatial variation in order to capture facets of the underlying data generating process, which subsequently will provide input into a more refined model specification (Haining 1990, 2003). On the other hand, methods of exploratory spatial data analysis should be structurally neutral and not induce unsupported patterns onto the analysis results. There are numerous examples of the use of GWR as an exploratory tool. Nakaya (2001) uses the GWR approach for spatial interaction modeling with local distance decay and accessibility parameters.

Huang and Leung (2002) apply GWR to study regional industrialization in China. Longley and Tobón (2004) perform a comparative study of several global and local spatial estimation procedures including GWR to investigate the heterogeneity in patterns of intra-urban hardship. Interestingly, these applications interpret the local parameter patterns without reporting the level of correlation in the estimated regression coefficients, even though there appears to be coefficient correlation in some map patterns.

Our intent in this paper is (a) to raise the awareness of GWR users that they may encounter multicollinearity problems in their exploratory spatial data analyses, (b) to demonstrate potential effects of multicollinearity in specific GWR analyses, and (c) to provide a set of simple diagnostic tools to explore the severity of the multicollinearity problem. While we provide sufficient evidence of an inherent multicollinearity problem in GWR and hint at some underlying causes, an exhaustive discussion of the multicollinearity inducing mechanisms must be left to future research, which integrates the local GWR regressions into a simultaneous model. This paper briefly reviews the GWR modeling approach in Section 2. Also in this section, some relevant diagnostic tools are introduced to facilitate the analysis of the relationships among the estimated local regression coefficients. Section 3 presents a GWR model of bladder cancer mortality. This model motivated us to perform this GWR multicollinearity investigation through the visualization and description of the patterns and relationships among the estimated local regression coefficients. Section 4 presents numerous experiments using different sets of eigenvectors, which are extracted from a spatial connectivity matrix, to systematically investigate the effects of multicollinearity on GWR parameters. These sets of eigenvectors give results that are consistent with our empirical analysis of the bladder cancer model, and with results of our analyses of other data sets. The paper concludes with a summary of the findings and some recommendations to explore GWR multicollinearity effects in future research.

2 GWR fundamentals

This section gives a brief review of the basic GWR model specification and estimation procedure. Greater statistical details can be found in Fotheringham et al. (2002). In contrast to a global regression model $y_i = \beta_0 + \sum_{k=1}^{p} \beta_k \cdot x_{ik} + \varepsilon_i$, where the regression coefficients $\{\beta_0, ..., \beta_p\}$ are location invariant, the specification of a basic GWR model is

$$y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where y_i is the dependent variable at location i, x_{ik} is the value of the kth explanatory variable at location i, the β_{ik} is the local regression coefficient for the kth explanatory variable at location i, β_{i0} is the intercept parameter at location i, and ε_i is the random disturbance at location i, which may follow an independent normal distribution with zero mean and homogeneous var-

iance. Thus, in contrast to the global model with fixed β_ks, the regression coefficients β_{ik} are allowed to vary form location to location.

There are m potential prediction locations, which do not need to match the n calibration locations at which actual data records are observed. In particular, if $m \gg n$ then GWR is used for spatial interpolation. From Eq. 1, there are a total of $m{\cdot}(p+1)$ regression coefficients estimated in GWR, where p is the number of variables in the model. The intercept term is counted individually. It is apparent that the number of coefficient estimates may be substantially larger than the available degrees of freedom based on the given number of calibration observations. Thus GWR must compensate for the lack of degrees of freedom by imposing constraints on estimated regression coefficients (Leung et al. 2000, p 10). The regression coefficients are estimated for each calibration location independently by location-specific weighted least squares models. The matrix calculation for the estimated regression coefficients is

$$\hat{\beta}(i) = [\mathbf{X}^{\mathrm{T}} \cdot \mathbf{W}(i) \cdot \mathbf{X}]^{-1}\mathbf{X}^{\mathrm{T}} \cdot \mathbf{W}(i) \cdot \mathbf{y}, \tag{2}$$

where $\mathbf{W}(i) = \mathrm{diag}\,[w_1(i),..., w_n(i)]$ is the diagonal weights matrix that varies for any calibration or prediction location i, $\mathbf{X}$ is the matrix of exogenous variables with a first column of 1 s for the intercept, $\mathbf{y}$ is the vector of dependent variables, and $\hat{\beta}(i) = \left(\hat{\beta}_{i0}, \hat{\beta}_{i1}, \ldots, \hat{\beta}_{ip}\right)^{\mathrm{T}}$ is the vector of $p+1$ local regression coefficients at location i.

The weights matrix is specified as a local kernel function that models a distance decay effect from the n calibration locations to the prediction location i. There are many specifications of the kernel function (Fotheringham et al. 2002). One of the most commonly used kernel functions, and the one used in this analysis, is the bi-square nearest neighbor function

$$w(i) = \begin{cases} [1 - (d_{ij}/b)^2]^2 & \text{if } j \in \{N_i\} \\ 0 & \text{if } j \notin \{N_i\} \end{cases}, \tag{3}$$

where d_{ij} is the distance between the calibration location j and the prediction location i, b is the threshold distance to the Nth nearest neighbor, and the set $\{N_i\}$ contains the observations that are within the distance range of the threshold Nth nearest neighbor (see Fotheringham et al. 2002, p 58). The weights for observations beyond the distance of the Nth nearest neighbor are zero. Note, in contrast to the Gaussian distance decay weight function that was used by Leung et al. (2000), the bi-square weight function cannot capture a global model in which all observations are equally weighted. Thus, it serves well in investigating GWR multicollinearity effects, even if the underlying data generating process is assumed to be based on a global model that lacks spatial heterogeneity. In the parameter estimation, an observation j in close vicinity to observation i exerts a higher weight on the calculation of the regression coefficients than a more distant observation. Observations outside the Nth nearest neighborhood are excluded from the regression coefficient calculation in Eq. 2. The observations are geo-referenced as point locations. These points may be representative points of areal objects, such as

their geometric or population centroids. The N_0 is the neighborhood threshold parameter that must be estimated using a cross-validation approach. Cross-validation with the bi-square nearest neighbor kernel function can be summarized using the following equation:

$$N_0 = \min_N \sum_{i=1}^{n} [y_i - \hat{y}_{(i)}(N)]^2, \tag{4}$$

where $\hat{y}_{(i)}$ is the predicted value of observation i with the calibration observation i removed from the estimation, and N_0 is the value of N that minimizes the residual sum of squares. A golden section search method was used to find N_0 in this research. After the best kernel parameter N_0 is found, the GWR model parameters are estimated at each location using the calibrated kernel function.

To evaluate the local correlation among the estimated regression coefficients at location i, analogous to the regression parameter correlation in the global model, the local parameter correlation matrix can be calculated from the local covariance matrix. In line with Eq. 2.15 in Fotheringham et al. (2002, page 55) we calculate the covariance matrix among the local regression coefficients as

$$\mathrm{Cov}[\hat{\boldsymbol{\beta}}(i)] = \sigma^2 \cdot [\mathbf{X}^{\mathrm{T}} \cdot \mathbf{W}(i) \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^{\mathrm{T}} \cdot \mathbf{W}^2(i) \cdot \mathbf{X} \cdot [\mathbf{X}^{\mathrm{T}} \cdot \mathbf{W}(i) \cdot \mathbf{X}]^{-1}. \tag{5}$$

This expression assumes that the disturbances are distributed as $\varepsilon \sim N(\mathbf{0}, \sigma^2 \cdot \mathbf{I})$. Throughout their work Fotheringham et al. (2002) are using this i.i.d. assumption. This sets GWR aside from a locally weighted regression model that assumes $\varepsilon(i) \sim N\big(\mathbf{0}, \sigma_i^2 \cdot \mathbf{W}^{-1}(i)\big)$ and is used by Páez et al. (2002a, b). How justified these assumptions of independence are in the light of multiple GWR estimations at n calibration locations, is left for future research. From Eq. 5 we derive the local correlation matrix among the regression coefficients at the i-th calibration location by

$$\mathbf{R}(i) = \mathrm{diag}^{-\frac{1}{2}}\{\ \mathrm{Cov}[\hat{\boldsymbol{\beta}}(i)]\} \cdot \mathrm{Cov}[\hat{\boldsymbol{\beta}}(i)] \cdot \mathrm{diag}^{-\frac{1}{2}}\{\mathrm{Cov}[\hat{\boldsymbol{\beta}}(i)]\} \tag{6}$$

where $\mathrm{diag}\{\cdot\}$ extracts the main diagonal from the square covariance matrix. Subsequently, we refer to these correlations (6) as the *local coefficient correlations* at prediction location i.

The objective of the paper is to study the local coefficient correlations in dependence of the multicollinearity among the exogenous variables in the global model. We also employ alternative tools to address this research question, such as scatter plots between the kth and lth sets of n local regression coefficients for the kth and lth exogenous variables, and their correlation coefficients,

$$\mathrm{Corr}\Big(\Big\{\hat{\beta}_{1k}, \ldots, \hat{\beta}_{nk}\Big\}, \Big\{\hat{\beta}_{1l}, \ldots, \hat{\beta}_{nl}\Big\}\Big) = \frac{\sum_{j=1}^{n} (\hat{\beta}_{jk} - \bar{\hat{\beta}}_k) \cdot (\hat{\beta}_{jl} - \bar{\hat{\beta}}_l)}{\sqrt{\sum_{j=1}^{n} (\hat{\beta}_{jk} - \bar{\hat{\beta}}_k)^2 \cdot \sum_{j=1}^{n} (\hat{\beta}_{jl} - \bar{\hat{\beta}}_l)^2}}, \tag{7}$$

where $\bar{\hat{\beta}}_k = (1/n) \cdot \sum_{j=1}^{n} \hat{\beta}_{jk}$. This correlation is subsequently called the *overall correlation* of two sets of local regression coefficients. It is important to stress here that overall local coefficient scatter plots are frequently more informative than the plain linear correlation coefficient, since we observed on several occasions that overall sets of GWR coefficients are frequently related in a non-linear way. The following four matrices more clearly make the distinction between the correlations that are presented in this paper. The underlying $n \times (p+1)$ design matrix at the n calibration locations is

$$\mathbf{X}_{[n\times(p+1)]} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$$

and the resulting matrix of local GWR coefficients at calibration locations becomes

$$\mathbf{B}_{[n\times(p+1)]} = \begin{pmatrix} \beta_{10} & \beta_{11} & \cdots & \beta_{1p} \\ \beta_{20} & \beta_{21} & \cdots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n0} & \beta_{n1} & \cdots & \beta_{np} \end{pmatrix}.$$

It is apparent that GWR must implicitly impose constraints on this over-specified calibration problem in order to obtain $n \times (p+1)$ local parameter estimates from only n degrees of freedom. The exact nature of these constraints and how they tie the local estimates within one local GWR equation and across the set of n local models together is left to future research. The calculated local coefficient correlation matrix (6) at one location provides the correlation among the estimated parameters in a row of the local GWR coefficient matrix

$$\mathbf{B} = \begin{pmatrix} \beta_{10} & & \beta_{11} & & \cdots & & \beta_{1p} \\ \beta_{20} & \leftrightarrow & \beta_{21} & \leftrightarrow & \cdots & \leftrightarrow & \beta_{2p} \\ \vdots & & \vdots & & \ddots & & \vdots \\ \beta_{n0} & & \beta_{n1} & & \cdots & & \beta_{np} \end{pmatrix}$$

and these correlations are used in local coefficient correlation maps. The overall correlation (7) among sets of local GWR coefficients provides the correlation of pairs of parameter estimates over all locations of the local GWR coefficient matrix

$$\mathbf{B} = \begin{pmatrix} \beta_{10} & \leftrightarrow & \beta_{11} & \cdots & \beta_{1p} \\ \beta_{20} & \leftrightarrow & \beta_{21} & \cdots & \beta_{2p} \\ \vdots & & \vdots & \ddots & \vdots \\ \beta_{n0} & \leftrightarrow & \beta_{n1} & \cdots & \beta_{np} \end{pmatrix}.$$

The overall local coefficient scatter plot among sets of local GWR estimates can be generated immediately from standard GWR output. It thus serves as an efficient exploratory tool to detect local coefficient dependencies. In contrast, the local coefficient correlation matrices, or the local variance inflation factors, are not readily available in any standard GWR software package and require specialized computational procedures. Pairwise local coefficient correlations can also be mapped to investigate the spatial variability of the GWR coefficient dependencies.

3 Multicollinearity in bladder cancer mortality data

To motivate the issue of multicollinearity in GWR, a simple model was built to explain white male bladder cancer mortality rates in the 508 State Economic Areas (SEA) of the United States for the years 1970–1994. The dataset comes from the Atlas of Cancer Mortality from the National Cancer Institute (Devesa et al. 1999) and contains age standardized mortality rates (per 100,000 person-years). The model consists of the explanatory variables population density and lung cancer mortality rate. Population density is used as proxy for environmental and behavioral differences with respect to an urban/rural dichotomy. It is expected, as several studies point out, that with an increase in the population density there is an increase in the rate of bladder cancer. Lung cancer mortality rates are used as proxy for the risk factor smoking, which is a known risk factor for bladder cancer. There is epidemiological evidence that an increase in smoking elevates the risk of developing bladder cancer, thus we expect a positive relationship between both variables. This approximation of smoking by lung cancer is reasonable, since the attributable risk of smoking for lung cancer is $>80\%$ and the attributable risk of smoking for bladder cancer is $>55\%$ (Mehnert et al. 1992). A global regression model was first built using bladder cancer mortality as the dependent variable with population density log transformed to linearize the relationship with the dependent variable. This model was then fit using a GWR approach implemented by the authors in Matlab. The summary results of the global regression are listed in Table 1. The risk factors are significantly positively related to the rate of bladder cancer, as expected. The variance inflation factors for the two global explanatory variable parameters are less than 2 and the correlation of the global

Table 1 Global regression and GWR results for the bladder cancer mortality model

Global						GWR		
Parameter	Estimate	Standard error × 2	p-value	VIF	Parameter correlation	Estimate mean	Inter-quartile range	*P*-value
Intercept	3.835	0.398	0.000			3.726	1.580	0.000
Smoking	0.030	0.013	0.000	1.534	−0.596	0.053	0.051	0.000
Pop. density	0.269	0.076	0.000	1.534	−0.596	0.124	0.291	0.000
R-square	0.25					0.73		

regression parameters is moderately negative at −0.59, whereas the correlation of the two variables is 0.59. These results suggest that multicollinearity is not a significant problem in the global model.

In order to test for heteroscedasticity with respect to the underlying population at risk pop_i in the 508 SEAs, a global weighted regression was fit using maximum likelihood estimation of the power parameter q for the weights $w_i = (pop_i)^q$, $i = 1, ..., 508$. While the parameter q is significant, it is distinctively closer to 0 than to 1, indicating no weights are more appropriate than the theoretically expected weights equal to the population at risk. In addition, regression coefficients of the unweighted and weighted models matched closely. Therefore, we will not account for model heteroscedasticity with respect to the areal population sizes in either the global model or the local GWR models.

The summary results of fitting this model with GWR are also listed in Table 1. The mean of the 508 local smoking regression coefficients is higher than in the global model and the mean of the local population density regression coefficients is lower than in the global model. The variance in both sets of local regression coefficients is higher in the GWR model based on a comparison of (1) the global standard errors and mean local standard errors and (2) two times the global standard errors and the inter-quartile range of the local parameters. The spatial heterogeneity in the GWR parameters is significant for the parameters of the explanatory variables as well as in the local intercepts according to the p-values from the Monte Carlo simulation test, where the test is described in Fotheringham et al. (2002, p. 93). The coefficient of determination increases notably from 0.25 of the global model to 0.73 of the local models.

It has been argued that one of the primary advantages of GWR is the ability to visualize the local regression coefficient estimates in order to identify local model heterogeneities. Figure 1 shows the map patterns for the GWR coefficients, which are associated with different explanatory variables. The two maps show a clear inverse map pattern correlation between the sets of local regression coefficients: in general, when the local smoking proxy parameter is high, the local population density parameter is low. This pattern is most noticeable in the West, Northeast, and portions of the Midwest immediately south of Lake Michigan.

The important question is whether this complementary relationship in the parameters is real, meaningful, and interpretable or an artifact of the statistical method. If the analyst does not ask this question and attempts to interpret the parameters, a likely interpretation is that in the West and Northeast smoking has a positive (statistically) relationship with bladder cancer mortality while population density has a counter-intuitive negative relationship with bladder cancer mortality. In addition, in parts of the Midwest and Oklahoma smoking has a counter-intuitive negative relationship with bladder cancer while population density has a positive relationship. This would lead to a serious mis-interpretation in these areas that is in gross contradiction to well-established etiological knowledge that smoking is a risk factor for bladder cancer.

Note that both choropleth map patterns of the local GWR coefficients must be cartographically symbolized by a bi-polar or diverging cartographic

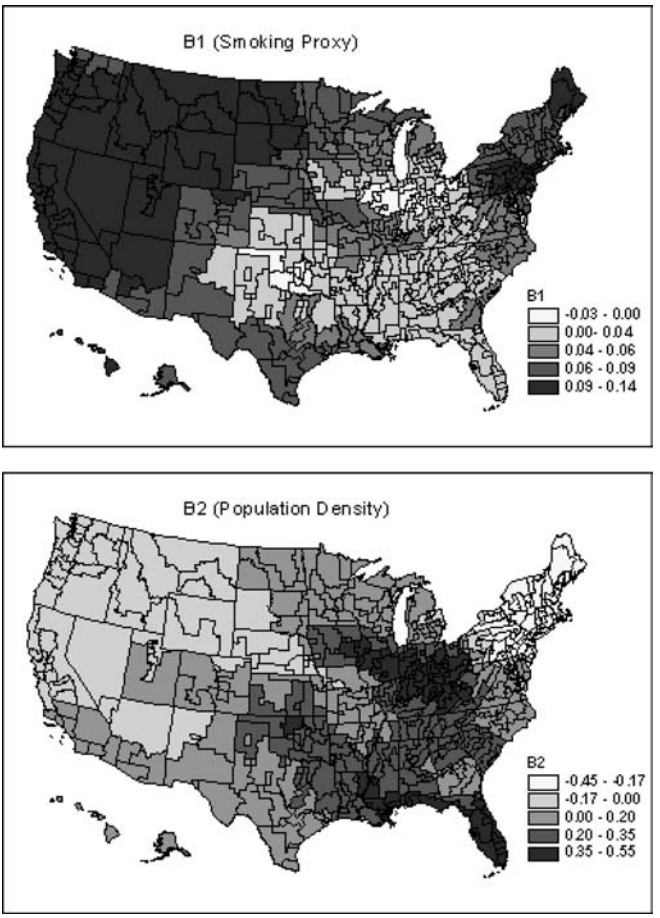

Fig. 1 Estimated GWR coefficients for a bladder cancer mortality model. The *top map* displays the spatial pattern of the local regression coefficients associated with the smoking proxy variable, while the *bottom map* displays the spatial pattern of the local regression coefficients associated with the log population density variable

map theme (Brewer et al. 1997). In a bi-polar map theme a particular value denotes a common reference around which the observed values are diverging. In our case positive and negative local GWR coefficients have a substantively different interpretation. Since bi-polar map themes are difficult to display in achromatic maps, we have opted for a connotation of observations below the reference values by a light gray scale whereas observations above the reference value are encoded by a heavy gray. A noticeable gap in the middle section of the gray scale enables us to distinguish immediately between both branches of the scale.

The next logical step in the analysis is to further explore the overall correlation between the sets of local regression coefficients. Figure 2 is a scatter plot of the local coefficient estimates for the two variables and shows a strong negative relationship. The dashed reference lines are the global

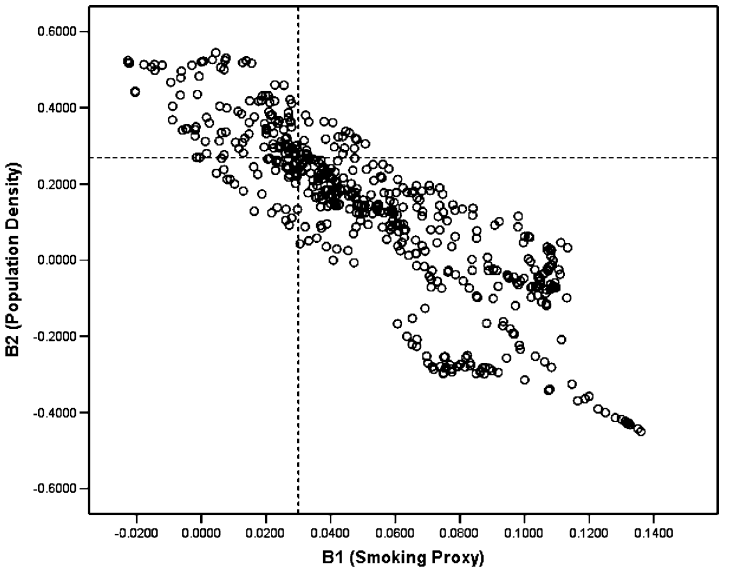

Fig. 2 Scatter plot of the local estimated regression coefficients associated with the smoking proxy and population density ($r = -0.85$). The dashed lines denote the levels of the related global parameter estimates

coefficient estimates and the plot shows that there is much variation around the global estimates. In addition, the local coefficient estimates do not center on their global counterparts. The overall correlation coefficient for the sets of local coefficients is -0.85 and helps quantify the visible negative relationship between the two map patterns illustrated in Fig. 1. While this scatter plot displays an overall linear relationship among the local GWR coefficients, some peculiar outlying clusters can be observed.

The local coefficient correlations based on Eq. 6 are mapped in Fig. 3 and it is clear that the strongest negative local parameter correlation is in the Midwest and parts of the Northeast. There are many locations in these areas with absolute magnitude correlation greater than 0.75. Figure 4 shows the distribution of the local coefficient correlations and highlights that many of the observations have a local correlation considerably stronger than the global coefficient correlation of -0.59 (the dashed line in the histogram). It is clear from these figures that the local coefficient correlation varies substantially over the study area and increases substantially when compared to the global coefficient correlation.

Another question to consider is whether the GWR coefficient estimate correlation varies with the size and type of kernel, which acts as a locally smoothing window of the exogenous variables in the GWR model. Fig. 5 is a plot of the relationship between kernel size N and the overall correlation of the sets of GWR coefficient estimates. The dashed reference line denotes the best N value (75) for the empirical bladder cancer model in terms of lowest residual sum of squares found through a cross-validation search routine (4). The relationship between kernel size and overall correlation of coefficients

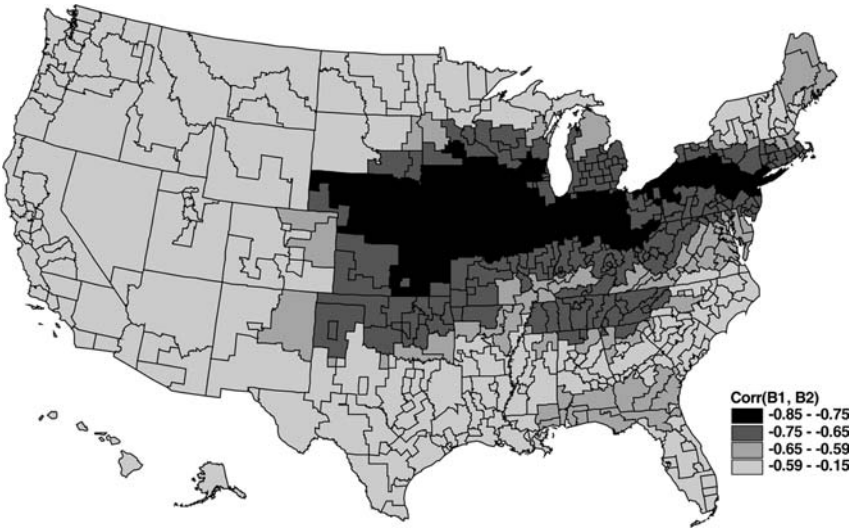

Fig. 3 Local coefficient correlations for the GWR coefficients associated with the smoking proxy and population density variables

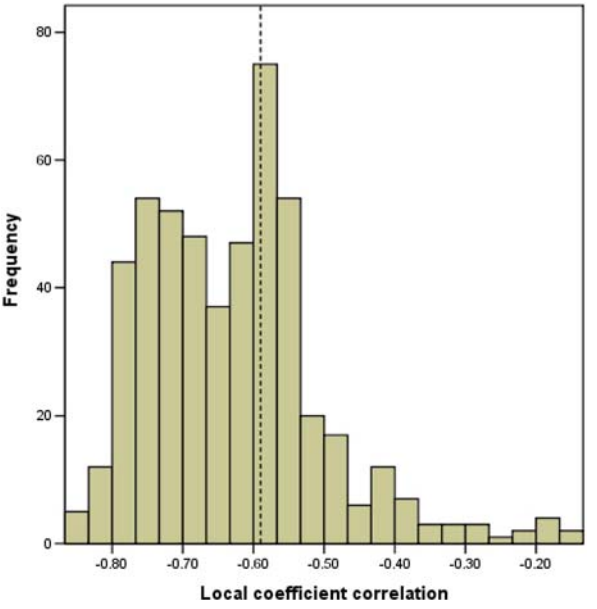

Fig. 4 Histogram of local coefficient correlations for smoking proxy and population density. The dashed reference line denotes the coefficient correlation in the global model

for this model is an almost monotonically decreasing function that reaches its minimum of just under -0.88 at a kernel size $N = 220$. A more thorough and exhaustive study is needed to better understand the influence of the

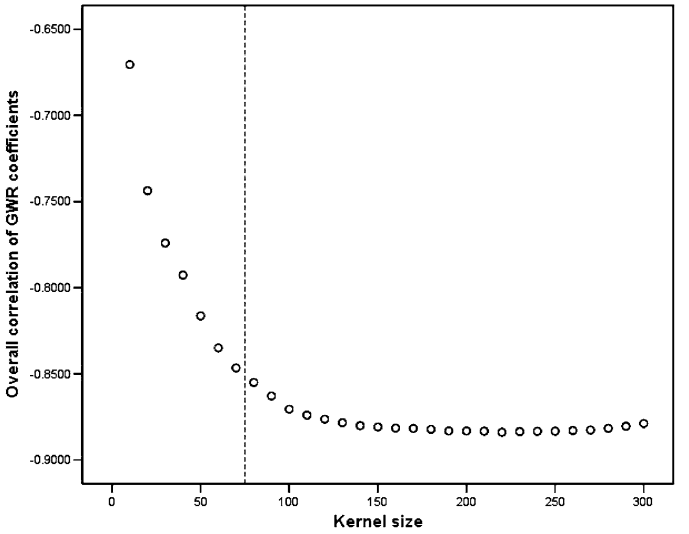

Fig. 5 Relationship between the kernel size and the overall correlation of the sets of GWR coefficients for smoking proxy and population density. The dashed line reflects the optimal bandwidth, which was determined by the cross-validation procedure

kernel function and its bandwidth on the estimated local regression coefficients and other model statistics.

4 Exploration of GWR multicollinearity in controlled experiments

The preliminary findings from the GWR model for bladder cancer mortality raise general questions about the presence of multicollinearity effects and the validity of the local regression coefficient estimates. Controlled experiments are necessary to further investigate the potential problem of multicollinearity in GWR and to gain a deeper understanding of the effects observed in the empirical bladder cancer model. In order to reduce the computational burden and use a more homogeneous study area with respect to the underlying spatial tessellation, these experiments were conducted for the 1990 Census layout of the $n = 159$ counties in Georgia. Fotheringham et al. (2002) used this tessellation in combination with a census dataset in their monograph as a benchmark tutorial on the GWR estimation procedure. Using the compact tessellation of Georgia's counties will overcome the spatial discontinuities of Alaska and Hawaii found in the bladder cancer example of the 508 SEAs and help avoid the potential problem of using the same distance threshold number of neighbors in Eq. 3 for the vastly expanding SEAs in the western United States and the densely packed SEAs in the East.

Several experiments are performed here with exogenous variables generated from the eigenvectors that are based on an $n \times n$ binary spatial link

matrix $\mathbf{C}$ of the 159 counties that captures the mutual adjacency relationships among the counties. The underlying eigenvector extraction was performed on the transformed and re-scaled link matrix $\mathbf{C}$ by

$$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} = \text{evec}\left(\left[\mathbf{I} - \frac{\mathbf{1}\cdot\mathbf{1}^{\mathrm{T}}}{n}\right]\cdot\frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}}\cdot\mathbf{C}\cdot\left[\mathbf{I} - \frac{\mathbf{1}\cdot\mathbf{1}^{\mathrm{T}}}{n}\right]\right)$$

in order for the eigenvectors to be orthogonal to the constant unity vector $\mathbf{1} = (1,1, \ldots, 1)^{\mathrm{T}}$, which is used in regression analysis to model the intercept. Eigenvectors of this transformed spatial adjacency matrix have specific properties that make them useful candidates for exogenous variables in our experiments: (1) they are uncorrelated among each other, so collinearity in a global data generating process is not an issue and all global regression coefficients are uncorrelated among each other (Fox 1997); (2) their means are zero and their variances are $(1/n)$, which we have rescaled for our analysis to one, thus, their basic distributional characteristics are directly comparable; (3) they exhibit particular spatial patterns and the spatial autocorrelation of these patterns with respect to Moran's I is identical to the associated eigenvalue of the eigenvector. In our experimentation, the first eigenvector $\mathbf{e}_1$ displays maximal positive autocorrelation with a Moran's I value of 1.1094, the last eigenvector $\mathbf{e}_{159}$ exhibits the greatest negative autocorrelation level with an I value of -0.5784, and the remaining eigenvectors are sorted in descending autocorrelation order within these two extremes. Details of this approach of generating uncorrelated spatial patterns with a given autocorrelation level and sample maps can be found in Boots and Tiefelsdorf (2000) and Griffith (2003). Note that the largest attainable autocorrelation level for the Georgia tessellation is clearly above one and that the smallest possible autocorrelation level does not reach minus one. Hence, as for most empirical tessellations, the standard bounds of the Pearson correlation coefficient do not apply (de Jong et al. 1984). Tiefelsdorf (forthcoming) uses spatial patterns derived from eigenvectors to motivate a test procedure for spatial pattern coherence.

The structure of the data generating process for the dependent variable $\mathbf{y}$ takes a simple form of a linear regression relationship with two exogenous variables $\mathbf{x}_1$ and $\mathbf{x}_2$

$$\mathbf{y} = \beta_0 + \beta_1\cdot\mathbf{x}_1 + \beta_2\cdot\mathbf{x}_2 + \boldsymbol{\varepsilon}, \tag{8}$$

where all global regression coefficients are set to $\beta_0 = \beta_1 = \beta_2 = 1$. The exogenous variables $\mathbf{x}_1$ and $\mathbf{x}_2$ as well as the disturbances ε are taken from the set of eigenvectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Note that the data generating process is specified strictly as a global model and any local heterogeneity in the GWR regression coefficients must be interpreted as local variation around the global data-generating process. We choose an eigenvector as a random disturbance vector instead of some randomly sampled data values because eigenvectors satisfy by default the theoretical requirement of a regression model that disturbances are uncorrelated with the exogenous variables in the model. Depending on the experimental design, the dis-

turbances had either no spatial autocorrelation (an associated eigenvalue very close to the expected value of Moran's I, i.e., $E(I) = -(1/(n-1))$, excluding all eigenvectors with an associated eigenvalue of 0, or some level of spatial autocorrelation. In the first run of the first experiment, the first exogenous variable $\mathbf{x}_1$ in the model is an eigenvector ($\mathbf{x}_1 \equiv \mathbf{e}_3$) that exhibits a positively autocorrelated spatial pattern. In the second run of the first experiment, the first exogenous variable is another positively autocorrelated eigenvector ($\mathbf{x}_1 \equiv \mathbf{e}_4$). The second variable $\mathbf{x}_2$ in the model is a linear combination of either the third and the first or the fourth and first eigenvector using the formula (see Boots and Tiefelsdorf 2000, p. 327)

$$\mathbf{x_2} = \begin{cases} \sin(\theta) \cdot \mathbf{e_3} + \cos(\theta) \cdot \mathbf{e_1} & \text{for the 1st run,} \\ \sin(\theta) \cdot \mathbf{e_4} + \cos(\theta) \cdot \mathbf{e_1} & \text{for the 2nd run,} \end{cases} \tag{9}$$

where θ is specified to control the level of correlation between $\mathbf{x}_1$ and $\mathbf{x}_2$, which is given by $\text{corr}(\mathbf{x}_1, \mathbf{x}_2) = \sin(\theta)$.

Various experiments are presented here. Each experiment has a different motivation and/or objective to investigate effects and causes of the local regression coefficient correlation. *Experiment 1* involves a systematic increase in the level of global correlation among the exogenous variables using the combined eigenvector approach of Eq. 9. The objective for Experiment 1 is to investigate if there is a systematic relationship between the overall coefficient correlation levels with increasingly positive or negative global correlation levels between the exogenous variables. Specifically of interest is whether a moderate level of positive global variable correlation induces strong negative correlation in local regression coefficients. *Experiment 2* focuses on the correlation patterns among the locally weighted exogenous variables. It involves fitting a GWR model with an orthogonal eigenvector pair as exogenous variables and a spatially uncorrelated error term to investigate the relationship between locally smoothed exogenous variables and local coefficient correlations. The objective of this experiment is to see to what extent local kernel weighted exogenous variable correlation influences the correlation of the local regression coefficients. *Experiment 3* uses a series of increasingly positive spatially autocorrelated disturbances $\boldsymbol{\varepsilon}$ in a GWR model with two uncorrelated exogenous variables to explore another possible cause for correlation of local regression coefficients.

4.1 Experiment 1: correlation among exogenous variables and bandwidth variation

The first experiment entails increasing the correlation in the exogenous variable $\mathbf{x}_2$ with the variable $\mathbf{x}_1$ in a controlled way by increasing θ in Eq. 9 and observing the level of correlation in the GWR coefficients. Note that if $\theta = 0$ then $\mathbf{x}_2 = \mathbf{e}_1$ with $\mathbf{x}_1$ uncorrelated with $\mathbf{x}_2$ and if $\theta = \pi/2$ then $\mathbf{x}_2 = \mathbf{e}_3$. Two eigenvectors, 1 and 3, with high levels of spatial autocorrelation were selected for explanatory variables because they have clearly distinguishable

patterns. In addition, eigenvector 4 was selected as an exogenous variable paired with eigenvector 1 in another run of the experiment because the two eigenvectors exhibit a different co-patterning than eigenvectors 1 and 3. The three eigenvector patterns are illustrated in Fig. 6. The spatial autocorrelation for eigenvector 1 has a Moran's I value of 1.109, the spatial autocorrelation for eigenvector 3 has a Moran's I value of 1.021, and the spatial autocorrelation for eigenvector 4 has a Moran's I value of 0.9861. The error term was equal to standardized eigenvector 61, which had no spatial autocorrelation as indicated by a Moran's I close to $E(I) = -1/(159-1)$.

The bi-square nearest neighbor kernel function was used and a fixed $N = 143$ was selected. In the experiment θ was systematically increased from -0.95 to 0.95 by 0.1 increments except for the step from $|0.90|$ to $|0.95|$ where the increment of 0.05 was chosen. Overall this leads to a correlation between $\mathbf{x}_1$ and $\mathbf{x}_2$ which ranges from -0.81 to 0.81. The overall correlation of the sets of GWR coefficients was recorded at each value of θ for each eigenvector pair and the results are plotted in Fig. 7. There are several interesting things

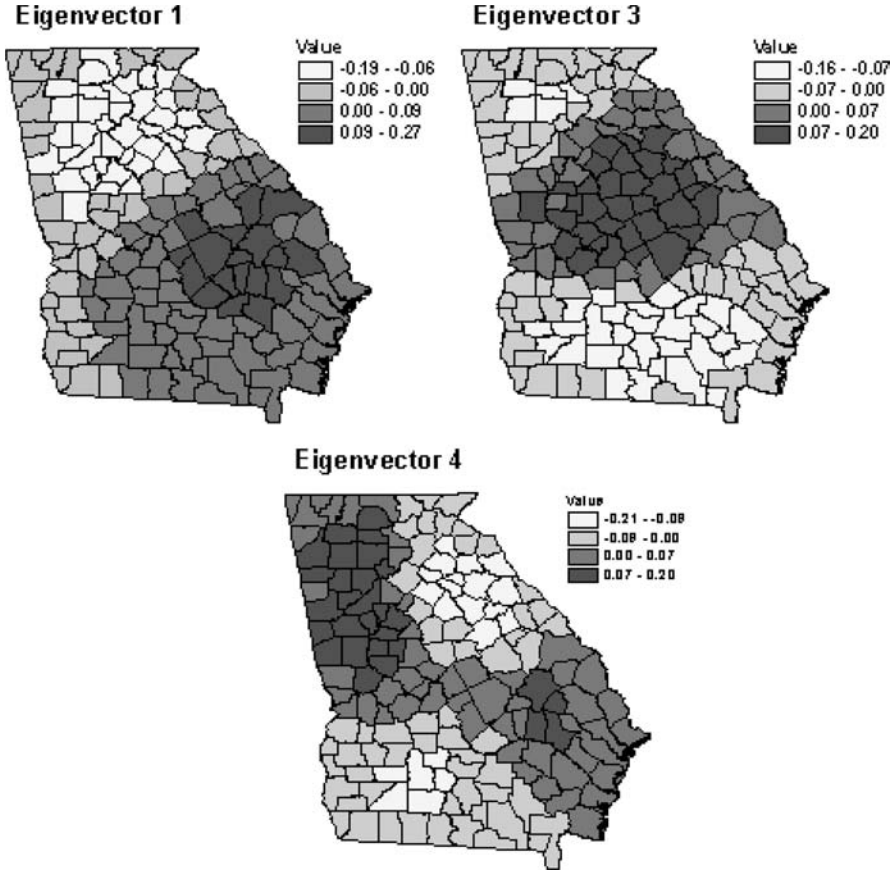

Fig. 6 Spatial patterns of the orthogonal eigenvectors associated with eigenvalue 1 ($I = 1.109$), eigenvalue 3 ($I = 1.021$), and eigenvalue 4 ($I = 0.9861$)

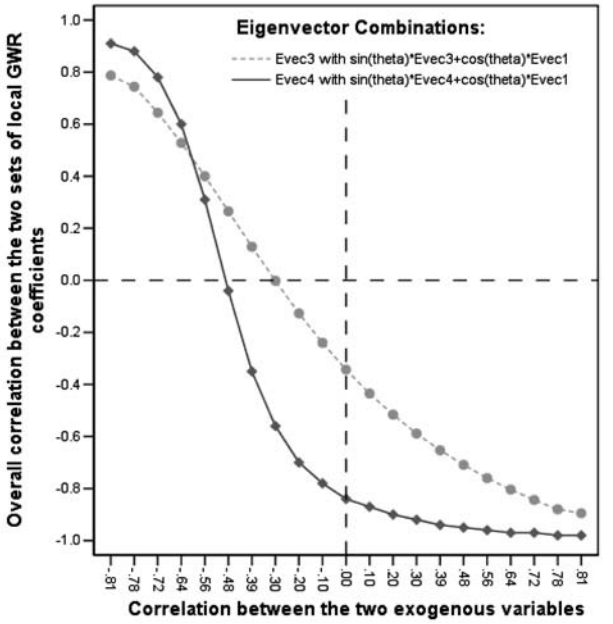

Fig. 7 Relationship between the correlation in the exogenous variables and the overall correlation between the sets of associated GWR coefficients for two different eigenvector pairs

to note in the figure. The figure shows a clear relationship between the amount of collinearity in the exogenous variables and the overall correlation between the sets of local GWR coefficients associated with both exogenous variables. The overall correlation between the coefficients becomes consistently more negative as the correlation in the exogenous variables becomes more positive. More interestingly, the overall correlation of the sets of GWR coefficients is -0.34 for eigenvectors 1 and 3 and is -0.84 for eigenvectors 1 and 4 when the exogenous variables are globally uncorrelated. Therefore, there can be strong negative overall correlation in the local regression coefficients even if the exogenous variable pair is uncorrelated. The figure also shows that there can be a fairly rapid increase in the strength of the overall correlation among the local regression coefficients. The overall correlation in the GWR coefficients is at a level of -0.8 when the global correlation of $\mathbf{x}_1$ and $\mathbf{x}_2$ reaches 0.64 for eigenvectors 1 and 3. The overall correlation is dangerously high when $\mathbf{x}_1$ and $\mathbf{x}_2$ are correlated above the 0.2 level for eigenvectors 1 and 4. In this portion of the graph the coefficients for eigenvector pair 1 and 4 are almost perfectly negatively correlated.

Figure 8 shows scatter plots of the local regression coefficient estimates at four levels (0.0, 0.4, 0.6, 0.8) of θ in the experiment for eigenvectors 1 and 3. The reference lines on the axes display the true global parameters. The plots show the increasingly more negative relationship in the sets of coefficient estimates as θ increases. Also, the variance of the local GWR coefficients increases as the global correlation in the exogenous variables increases.

Apparently, some spatial structure effects are present in the estimation of the GWR regression coefficients as illustrated by the distinctive ribbons of points in the scatter plots. The more extreme case of overall correlation in regression coefficients using eigenvectors 1 and 4 is shown in Fig. 9 for $\theta = 0.0$ and $\theta = 0.7$. The figure reinforces the impression that the sets of local regression coefficients are almost perfectly linearly dependent when the exogenous variables are moderately globally correlated.

To ensure that the results were not specific to the selected kernel size, the experiment was repeated using two sample kernel sizes, 40 and 159. The results for both these kernel sizes were similar to those from the initial run of the experiment ($N = 143$). More specifically, the trends in the overall correlation in the regression coefficients shown in Figs. 7 and 8 were still prevalent with the other kernel sizes.

The GWR regression coefficients for the exogenous variables eigenvector 3 and the combined eigenvector using eigenvector 1 with $\theta = 0.7$ ($r = 0.64$ for explanatory variables) are mapped in Fig. 10. The maps of the correlated exogenous variables $\mathbf{x}_1$ and $\mathbf{x}_2$ in the top of the figure show similar but not exact patterns. The negative relationship ($r = -0.8$) in the GWR coefficients is clearly evident in the regression coefficient maps in the bottom of the figure. The GWR coefficients for the first exogenous variable $\mathbf{x}_1$ are highest in the north and south and lowest in the east. The GWR coefficients for the second exogenous variable are lowest in the north and south and highest in the east. The patterns in the two exogenous variables are most similar in the north, south, and central areas. Comparing the exogenous variable maps of $\mathbf{x}_1$ and $\mathbf{x}_2$ and local coefficient maps reveals that when both variables have similar patterns of values in an area, GWR associates the effect in that area to only one parameter and the other parameter is pushed away from this and actually gets the opposite effect. One can speculate that a similar mechanism is at work in the bladder cancer mortality example. The local regression coefficients in this experiment are clearly misleading and cannot be interpreted individually.

4.2 Experiment 2: local correlation of the exogenous variables

It is reasonable to suspect that the global correlation level in the explanatory variables may not be representative of the correlation observed among the exogenous variables at a local scale. This local scale correlation among the independent variables is specified in this investigation by a weighted spatial window based on the same local kernel function that is used in the GWR model. This may offer an explanation for the relatively large local coefficient correlations in some areas and small local coefficient correlations in other areas. The local kernel weighted exogenous variables at each calibration location are a direct transformation of the global exogenous variables and therefore the correlation of these locally weighted variables can be directly linked to the local coefficient correlations. In general, one can expect that the smoothing effects of the weights accentuate positive local correlation in the exogenous variables. To investigate the relationship between the local cor-

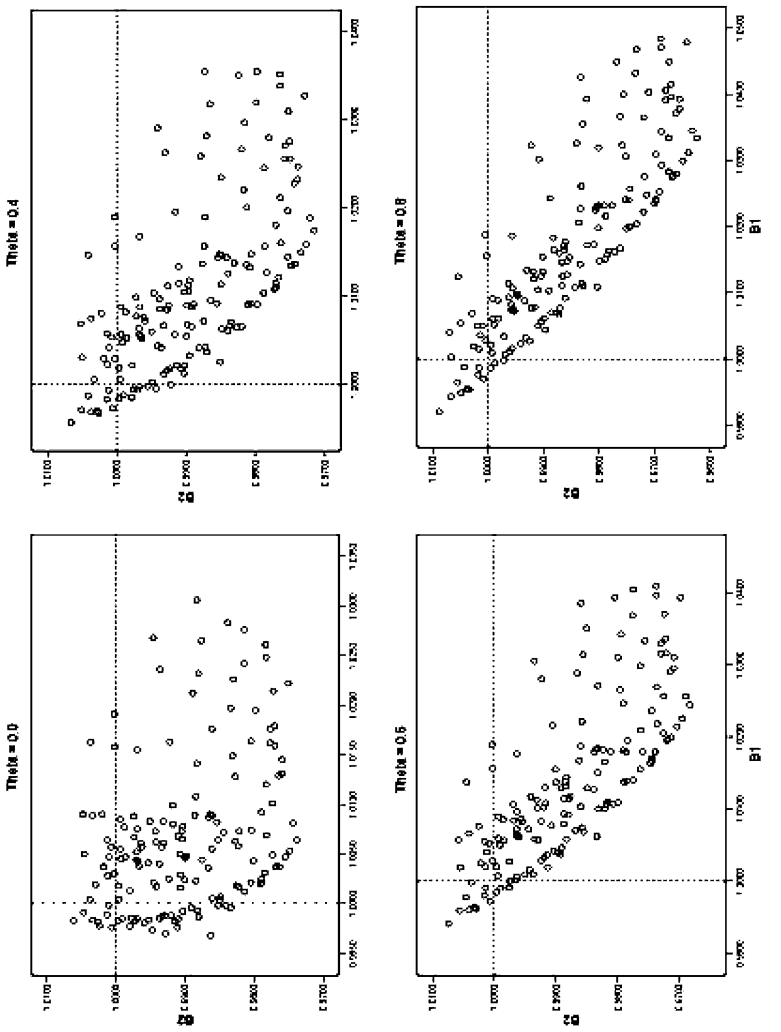

Fig. 8 Scatter plots of the local regression coefficients at four levels of $\theta \in \{0.0,0.4,0.6,0.8\}$ for eigenvector pair 1 and 3

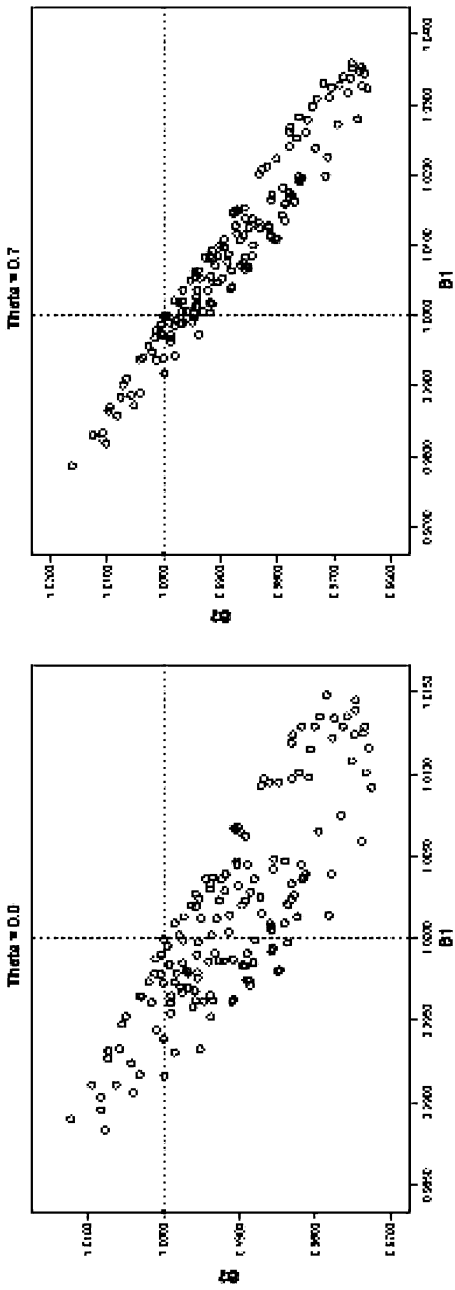

Fig. 9 Scatter plots of the local regression coefficients at two levels of $\theta \in \{0.0, 0.7\}$ for eigenvector pair 1 and 4

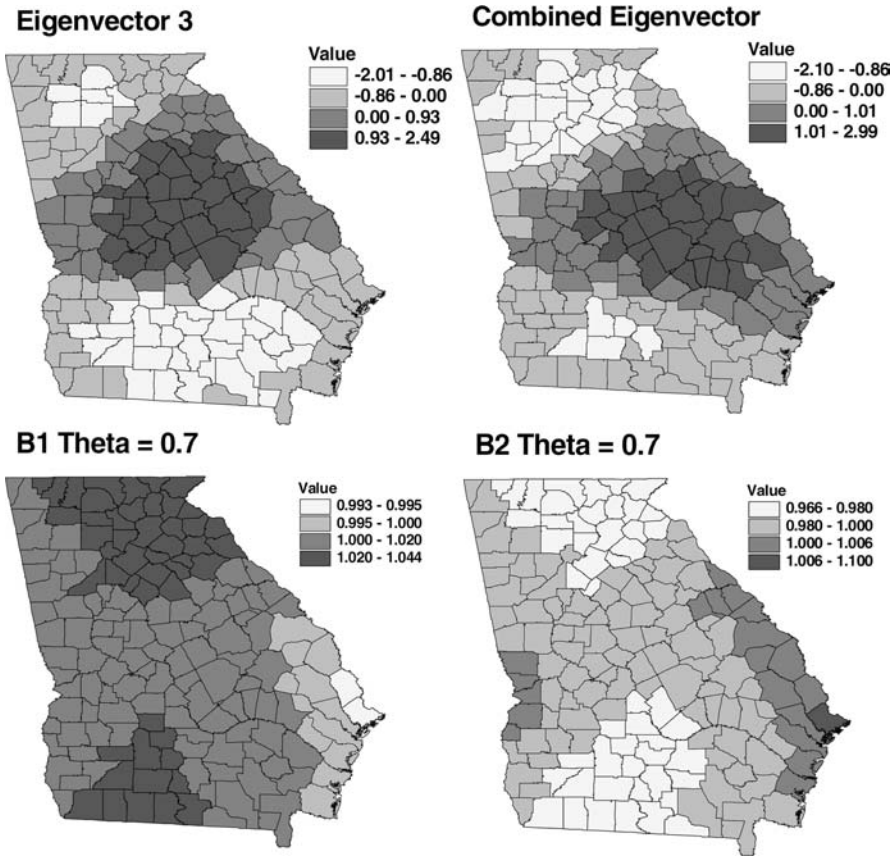

Fig. 10 Geographically weighted regression coefficient estimates for a model with eigenvector 3 and a combination eigenvector using eigenvector 1 ($\theta = 0.7$). The reference level for the eigenvector patterns is zero, whereas the reference level for the estimated GWR coefficients is $\beta_1 = 1$ and $\beta_2 = 1$

relations of weighted exogenous variables and local coefficient correlations, two GWR models were fit for both eigenvector pairs (1 and 3, 1 and 4) from experiment 1. The spatially independent error term is specified by eigenvector 61. The eigenvectors in either set remain pairwise uncorrelated ($\theta = 0.0$) and the level of positive spatial autocorrelation in the exogenous variable is an indicator of its local spatial smoothness. The correlation among the local regression coefficients is calculated as described earlier in Sect. 2 by Eqs. 5 and 6.

Figure 11 shows scatter plots of the local coefficient correlations for the two exogenous variables and the correlation of the locally weighted exogenous variables at the n calibration locations for the two eigenvector pairs. In both plots there is a clear negative relationship between the local correlations of the weighted exogenous variables and the local regression coefficient correlations. Areas with stronger positive weighted exogenous variable correlation have stronger negative local coefficient correlation. The association

is stronger for eigenvectors 1 and 4, although the relationship is less linear. One can suspect that latent spatial structure effects are responsible for the ribbons in both scatter plots. The result of this experiment supports the idea that the correlations among locally weighted exogenous variables are one of the driving forces behind local coefficient correlations in the GWR coefficients. This experiment was repeated for other eigenvector pairs and their results are consistent with this finding.

4.3 Experiment 3: spatial autocorrelation in the error term

All the experiments performed to this point had a specified error term with no spatial autocorrelation. To investigate the effect of using spatially autocorrelated disturbances in the model, an experiment was designed with standardized eigenvectors 1 and 3 as the exogenous variables in a model and one eigenvector from the range of positively spatially autocorrelated eigenvectors as the error term in the model. This is the model in experiment 1 with $\theta = 0$, but with a spatially autocorrelated error term with a varying autocorrelation level. For spatially uncorrelated disturbances, the overall correlation of the GWR coefficients was -0.34 (see Experiment 1). There are 58 positively spatially autocorrelated eigenvectors remaining to be used for the error term after eigenvectors 1 and 3 are included in the model. Each one of these remaining eigenvectors is used as the error term in 58 other GWR models in the experiment.

The correlation of the sets of GWR coefficients is plotted in Fig. 12 along with the corresponding autocorrelation level of the eigenvector used as the error term in each of the 59 models. Figure 12 shows that there is an overall slightly positive relationship between the level of spatial autocorrelation in the error term and the level of overall correlation in the GWR coefficients. However, the overall correlation of the GWR coefficients jumps erratically between both sides of the neutral reference line (no correlation) throughout the range of error autocorrelation levels. Therefore, it is not possible to use the autocorrelation level of the error term to determine whether the overall correlation in the regression coefficients will be positive or negative. The strongest overall correlation in the local coefficients (-0.91) is with an error term of eigenvector 15 (spatial autocorrelation level of Moran's $I = 0.686$). Figure 13 is a scatter plot of the local coefficient estimates for both exogenous variables when eigenvector 15 is used as the error term. The plot has reference lines for the true global regression coefficients. There is a negative nonlinear pattern in the regression coefficients and the true global coefficient values are not jointly included in any pair of the local coefficient estimates. This indicates the potential for biased GWR coefficients when the error term is spatially autocorrelated. The results of this experiment suggest that in some cases, models with spatially autocorrelated error terms may lead to GWR coefficients that exhibit an artificially strong dependence.

One final note is that the overall GWR R^2 goodness of fit (see Fotheringham et al. 2003) tends to increase as the spatial autocorrelation in the error term increases. This is shown in Fig. 14, where there is a nonlinear

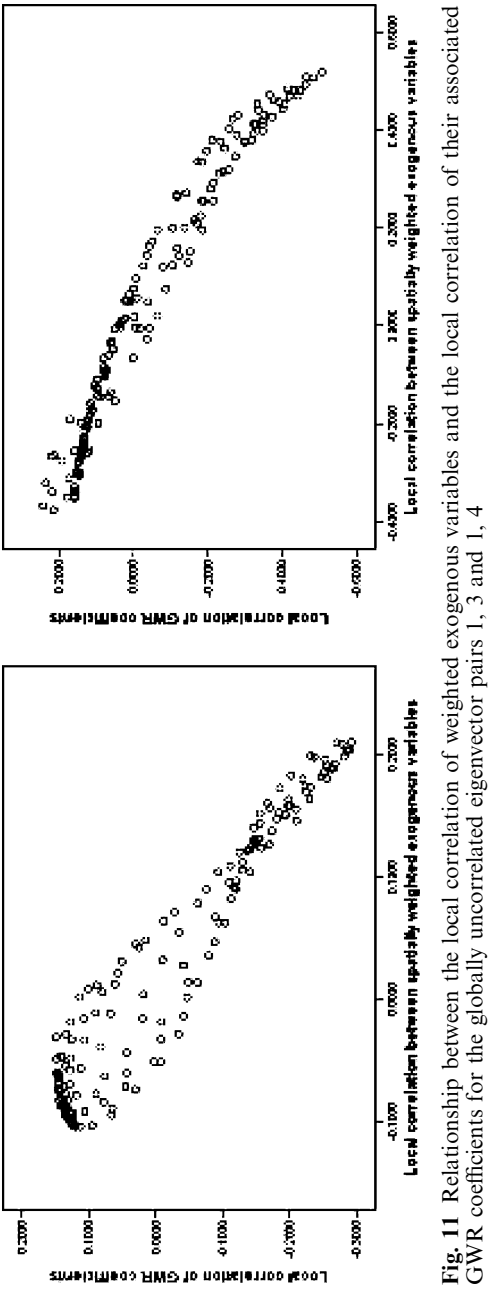

Fig. 11 Relationship between the local correlation of weighted exogenous variables and the local correlation of their associated GWR coefficients for the globally uncorrelated eigenvector pairs 1, 3 and 1, 4

positive relationship between R^2 and the error term autocorrelation. There is also a relatively large jump in R^2 when the autocorrelation level reaches 0.87. This suggests that GWR becomes a better predictor when the global regression errors are spatially autocorrelated. Thus, GWR utilizes implicitly the redundancy (see Haining 1991) among the observations that is induced by spatial autocorrelation, whereas global predictors such as the simultaneous autoregressive model explicitly incorporate spatial autocorrelation among the observations.

5 Conclusions

Little attention has been paid to standard diagnostic techniques in GWR. This paper demonstrates the need for diagnostic tools, especially regarding the issues of multicollinearity that may arise in GWR. This paper also found that the effects of multicollinearity are substantially stronger in the GWR model than in global regression models. In GWR, moderate to strong correlation of two explanatory variables makes their associated local parameters almost completely interdependent. This correlation of local regression coefficients potentially invalidates any interpretation of individual GWR parameter estimates and can facilitate misleading conclusions if the situation is not properly diagnosed. It is not a new finding that local spatial estimates have the tendency to be correlated among each other due to implicit constraints that tie local estimates together, such as the sharing of common data in local estimation procedures. Ord and Getis (1995) discuss the correlations among local G_i-statistics. Tiefelsdorf (2000, Figs. 11.4 and 11.5) shows that local Moran's I_i-statistics are correlated among each other, dependent on the

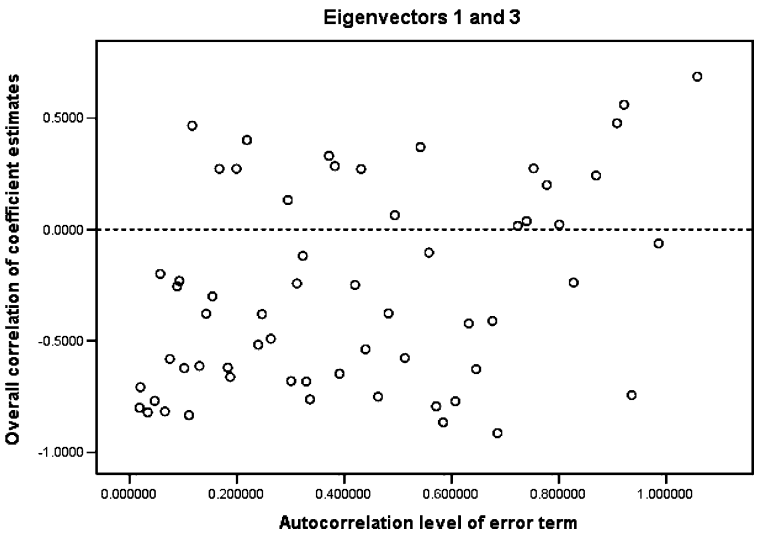

Fig. 12 Scatter plot of the overall correlation between the sets of local coefficient estimates versus the spatial autocorrelation level in the error terms for the model with eigenvector 1 and eigenvector 3 as exogenous variables

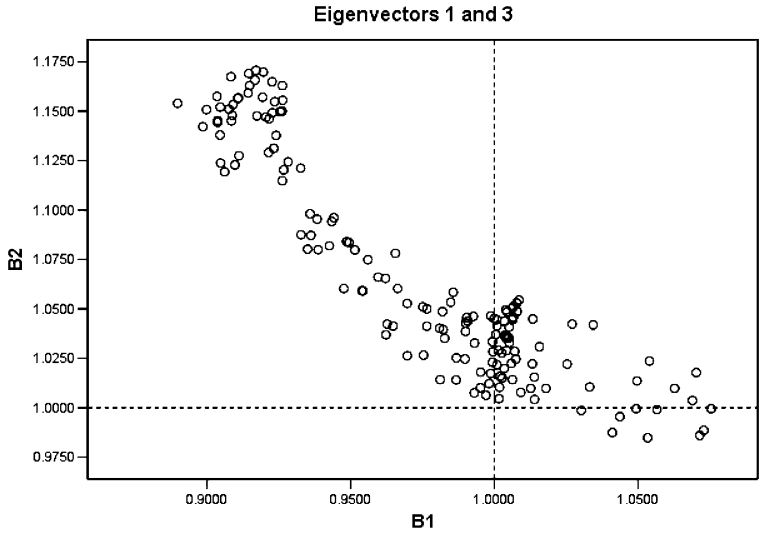

Fig. 13 Scatter plot of local coefficient estimates for eigenvector 1 and eigenvector 3 with a spatially autocorrelated error term (autocorrelation level 0.686)

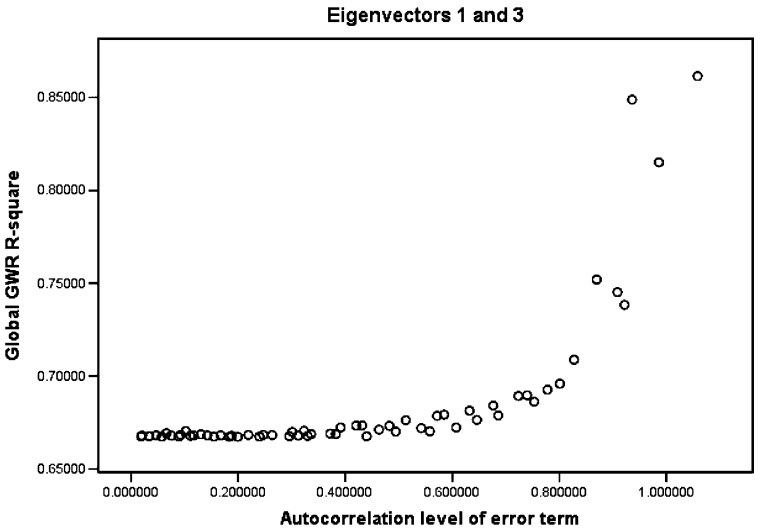

Fig. 14 Relationship between the spatial autocorrelation level of the error term and global GWR R^2 for the model with exogenous eigenvectors 1 and 3

spatial lag. Furthermore, in the context of spatial interaction models, Tiefelsdorf (2003) highlights the case where the correlation between a set of local distance decay parameters and other estimated parameters voids any substantive interpretation of the spatial heterogeneity in local distance decay effects.

The potential repercussions of multicollinearity in GWR require a careful application of the technique and the use of diagnostic tools. This paper

presented some useful multicollinearity diagnostics, including (a) scatter plots of the local parameter estimates (see Fig. 2), (b) maps of the local parameter correlation (see Fig. 3), (c) histograms of local parameter correlations (see Fig. 4), and (d) scatter plots of correlation between two local kernel weighted explanatory variables and the local parameter correlations (see Fig. 11). One may also produce local VIF maps for each variable to show areas of large parameter variance inflation. One reviewer of this paper demonstrated the usefulness of a local variant of Belsley et al. (1980) variance-decomposition proportions to investigate local multicollinearity in GWR. The exploratory spatial repertoire of the GeoDa software (Anselin 2003) can be used to investigate the observed spatial structure effects in our scatter plots (see, for instance, Figs. 2, 8 and 11).

Additional research is needed to summarize the effect of spatially autocorrelated errors on the GWR parameters. As our initial results presented here show, spatially autocorrelated errors can produce severely correlated local regression coefficients. In addition, the effects of different bandwidths, alternative specifications of the spatial weights function, and their interaction with exogenous variables and the error term require further investigations. Potential spatial structure effects, which are apparently inherent in GWR, also require further investigation. The analysis presented in this paper only considered simple models with two explanatory variables. Future research should study, in controlled experiments, the joint impact of multicollinearity in models with several exogenous variables and spatially varying regression coefficients.

Currently, GWR ignores that the local models must relate to a global reference model in order to express the local parameters as variation around their global counterparts. This link between a global estimate and local estimates is in line with the interpretation of local Moran's I_i as variation around a global spatial autocorrelation level (Anselin 1995; Tiefelsdorf forthcoming), and with the link between local and global distance decay parameters in spatial interaction models (Tiefelsdorf 2003). In order to tie the local models to their global reference model, a system of seemingly unrelated regression equations may be built that allows for a global reference model and simultaneous testing of local parameter estimates across several local models. Gelfand et al. (2003) demonstrate a simultaneous approach using a Bayesian framework. However, in how well the Bayesian approach deals with multicollinearity issues is largely unstudied. Clearly, more work is needed in these research areas.

References

Anselin L (1995) Local indicators of spatial association – LISA. Geogr Anal 27(2):93–115

Anselin L (2003) An introduction to spatial autocorrelation analysis with GeoDa. GeoDa Documentation. URL: http://www.csiss.org/clearinghouse/GeoDa/

Belsley DA, Kuh E, Welsch RE (1980) Regression diagnostics: Identifying, influential data and sources of collinearity. Wiley, New York

Boots BN, Tiefelsdorf M (2000) Global and local spatial autocorrelation in bounded regular tessellations. J Geogr Syst 2:319–348

Brewer CA, MacEachren AM, Pickle LW, Hermann D (1997) Mapping mortality: evaluating color schemes for choropleth maps. Ann Assoc Am Geogr 87:411–438
Brunsdon C, Fotheringham AS, Charlton M (1996) Geographically weighted regression: A method for exploring spatial nonstationarity. Geogr Anal 28(4):281–298
de Jong P, Sprenger C, van Veen F (1984) On extreme values of Moran's *I* and Geary's c. Geogr Anal 16:17–24
Devesa SS, Grauman DJ, Blot WJ, Pennello G, Hoover RN, Fraumeni JF Jr (1999) Atlas of cancer mortality in the United States, 1950–94. National Cancer Institute, Bethesda. URL: http://www3.cancer.gov/atlasplus/
Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically weighted regression: The analysis of spatially varying relationships. Wiley, West Sussex
Fotheringham AS, Charlton M, Brunsdon C (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ Plan A 30(11):1905–1927
Fox J (1997) Applied regression analysis, linear models, and related methods. Sage Publications, Thousand Oaks
Gelfand A, Kim HJ, Sirmans CJ, Banerjee S (2003) Spatial modeling with spatially varying coefficient processes. J Am Stat Assoc 98(462):387–396
Greene WH (2000) Economic analysis. Prentice-Hall, Upper Saddle River, New Jersey
Griffith D (2003) Spatial autocorrelation and spatial filtering. Springer, Berlin Heidelberg New York
Haining RP (1990) Spatial data analysis in the social and environmental sciences. Cambridge University Press, Cambridge
Haining, RP (1991) Bivariate correlation with spatial data. Geogr Anal 23:210–27
Haining RP (2003) Spatial data analysis: theory and practice. Cambridge University Press, Cambridge
Huang Y, Leung Y (2002) Analysing regional industrialisation in Jiangsu province using geographically weighted regression. J Geogr Syst 4:233–249
LeSage JP (2001) Econometrics toolbox for MATLAB. URL: http://www.spatial-econometrics.com/
LeSage JP (2004) A family of geographically weighted regression models. In: Anselin L, Florax RJGM, Rey SJ (eds) Advances in spatial econometrics. Methodology, tools and applications. Springer, Berlin Heidelberg New York, pp 241–264
Leung Y, Mei CL, Zhang WX (2000) Statistical tests for spatial nonstationarity based on the geographically weighted regression model. Environ Plan A 32:9–32
Longley JW (1967) An appraisal of least squares programs from the point of the user. J Am Stat Assoc 62:819–841
Longley PA, Tobón C (2004) Spatial dependence and heterogeneity in patterns of hardship: an intra-urban analysis. Ann Assoc Am Geogr 94:503–519
Mehnert WH, Smans M, Muir CS, Möhner M, Schön D (1992) Atlas of cancer incidence in the former German Democratic Republic 1978–1982. Oxford University Press, New York
Nakaya T (2001) Local spatial interaction modelling based on the geographically weighted regression approach. GeoJournal 53:347–358
Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geogr Anal 27:286–306
Páez A, Uchida T, Miyamoto K (2002a) A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. Environ Plan A 34:733–754
Páez A, Uchida T, Miyamoto K (2002b) A general framework for estimation and inference of geographically weighted regression models: 2. Spatial association and model specification tests. Environ Plan A 34:883–904
Tiefelsdorf M (2000) Modelling spatial processes: the identification and analysis of spatial relationships in regression residuals by means of Moran's *I*. Springer, Berlin Heidelberg New York
Tiefelsdorf M (2003) Misspecifications in interaction model distance decay relations: A spatial structure effect. J Geogr Syst 5:25–50
Tiefelsdorf M (forthcoming) Specification and distributional properties of the spatial cross-correlation coefficient $C_{\varepsilon_1,\varepsilon_2}$. Environ Plan A (conditionally accepted)