UNDERSTANDING SPATIAL FILTERING FOR ANALYSIS OF LAND USE DATA SETS

Yiyi Wang
The University of Texas at Austin
6.510. Cockrell Jr. Hall
Austin, TX 78712-1076
yiyiwang@mail.utexas.edu

Kara M. Kockelman
(Corresponding author)
Professor and William J. Murray Jr. Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
Austin, TX 78712-1076
kkockelm@mail.utexas.edu
Phone: 512-471-0210

Xiaokun (Cara) Wang
Assistant Professor
Department of Civil and Environmental Engineering
4031 JEC Building, Rensselaer Polytechnic Institute
110 8th Street, Troy, NY 12180-3590 USA
Methods

The following paper is a pre-print and the final publication can be found in the *Journal of Transport Geography* No. 122-131, 2013.

ABSTRACT

This paper summarizes the literature on spatial filtering for analysis of spatial data, as proposed by Griffith (2000a). Given the scarcity of its application in transportation and its fledgling nature, preliminary case studies were conducted using continuous and discrete response data sets, for land values and land use, in comparison with results from spatial autoregressive models with distance decay parameters estimated using Bayesian techniques. For both the continuous land value and binary land use cases, the spatial filtering approach demonstrates great potential as a worthy competitor to more conventional Bayesian spatial-autoregressive models, offering high fit statistics, somewhat shorter computing times, and more straightforward computations.

Key words: spatial filtering, spatial econometrics, land use change, land valuation, spatial probit, spatial autoregressive regression

INTRODUCTION

Spatial relationships typically exist across persons and locations in transportation, land use, and demographic data sets. They can be summarized into two types of spatial effects: spatial heterogeneity and spatial autocorrelation (Anselin 1988).

Spatial heterogeneity refers to the variation or instability in observational units across a geographic region and implies that the functional forms and/or behavioral parameters vary by location (Anselin 1988). Geographically weighted regression (GWR) is commonly used to characterize such variations, by estimating parameters for each site or observational unit based on all observations within a neighborhood (optimally pre-determined using cross validation), as described in Fotheringham (2003). Ghosh et al. (2008) analyzed impervious (or heavily developed) land cover shares via a continuous-response GWR model, using data points across Minnesota's Twin Cities metro area. Páez (2006) calibrated a binomial probit GWR model with heteroscedastic error terms to characterize the development of 324 vacant 1-hectare grid cells near California's Bay Area Rapid Transit lines. Applications of multinomial GWR models for land use patterns can be found in McMillan and McDonald (1999) and Wang et al. (2011). The second type of spatial effect, spatial autocorrelation, arises primarily due to imperfect information on observational units and measurement errors. For example, information on variables like soil types, scenic beauty, school quality, and air traffic is regularly absent from models of land use value and change, resulting in correlations across nearby sites' error terms. Moreover, aggregated spatial data values (such as county-level average incomes) and arbitrary spatial boundaries (such as Census tract designations) can introduce forms of spatial autocorrelation. A common treatment for this effect is to directly specify a spatial structure, such as a spatial autoregressive (SAR) or spatial moving average (SMA)[1] models (Anselin 1988, Anselin and Hudak 1992, LeSage and Pace 2009). Work on discrete states of land use change with such specifications can be found in Chakir and Parent's (2009) spatial multinomial probit model (for cross-sectional data), Munroe et al.'s (2002) series of binary probit and random-effects probit models (using panel techniques), and Wang and Kockelman's (2009a, 2009b, 2009c) dynamic spatial ordered probit model with a temporal component.

Most applications to date rely on specific functional forms (such as SAR and SMA) and arbitrarily pre-determined weight structures to anticipate spatial structure in the data. Several issues can limit the use of specific functional forms in addressing spatial autocorrelation. McMillen (2004) noted how functional misspecifications may lead to spatial autocorrelation and advocated the use of non-parametric methods, to avoid *a priori* assumptions of model form. He used a Lagrange Multiplier (LM) test to examine residuals' spatial autocorrelation for a series of GWR applications with varying window sizes (or neighbor counts). Although reducing window sizes did not monotonically decrease Lagrange Multiplier test statistics, restricting the spatial relations to smaller neighborhoods (or window sizes) contributed to, in general, a smaller degree of spatial autocorrelation. Hence, McMillen argued that spatial autocorrelation may be induced by functional misspecification.

Computing effort is another important factor to consider, as demonstrated in Wang et al.'s (2010) pursuit of an estimable dynamic spatial multinomial probit specification. Essentially, more

[1] Spatial error models, where the weight matrix applies to the vector of error terms (as used in Zhou and Kockelman [2009], for example), are a type of SMA model.

complicated models require more complex estimation strategies, such as Bayesian sampling from large-size truncated normals (for latent response variables, in the case of multinomial probits, for example); issues of parameter identification, sample size limitations, and a model's functional flexibility can and do emerge. Another serious challenge relates to computing the log-determinant of a SAR specification: $|I_n - \rho W|$, where I_n is an *n* by *n* identity matrix, ρ is the degree of spatial autocorrelation and *W* is the connectivity or weight matrix. This is especially time-consuming when *n* is large. This log-determinant term is also known as the "normalizing factor" or the Jacobian term, and it ensures that the integral of a spatial model's likelihood function equals 1 (when integrating over the density or distribution of all unknown parameters) (Griffith, 2000a). These can be computationally costly to obtain, and Griffith (2000a) discussed three strategies for computing this term.

Eigenvector-based spatial filtering, as introduced by Griffith (2000a, 2000b, 2007), is a relatively new technique for analysis of spatial data sets, and it appears to offer much promise. Griffith (2008) formally established an indirect linkage between GWR and spatial filtering via interaction terms, and noted how GWR can be viewed as a special case of indirect spatial filtering. In other words, spatial filtering should be able to address apparent heterogeneity in behaviors by interacting eigenvectors (synthetic variables) and systematic covariates (*X*). SAR models rely on the assumption that responses of observational units are more likely to be influenced by nearby neighbors. They employ an *n* by *n* weight matrix (where *n* is the number of sample units) to describe this distance-decay spatial pattern (or any other pattern of decay, as in social networks). In contrast, spatial filtering addresses such spatial autocorrelation from a quasi *semi-parametric* point of view. Apart from the observed covariates, also known as the systematic component, spatial filtering techniques generate synthetic explanatory variables representing the data set's spatial structure. More flexibility is added to the model by bringing these synthetic variables (considered the model's non-parametric component [Tiefesldorf and Griffith 2007]) into the systematic part of a model. Different methods of generating these synthetic variables leads to three main types of spatial filtering in the literature: Dray et al.'s (2006) distance-based eigenvector procedure, Getis' (1990, 1995) G-statistics-based approach, and Griffith's (2000a, 2000b) eigenfunction-based procedure.

Dray et al.'s (2006) approach is also known as principal coordinate analysis of neighbor matrices (PCNM). Their work built upon Borcard and Legendre's (2002) original PCNM proposal by providing more formal mathematics and the linkage between the PCNM method and spatial autocorrelation structure. The original PCNM approach follows a 3-step procedure:

First, a pairwise Euclidean distance matrix (*D*) is computed between all sampled locations (where $D=[d_{ij}]$). Second, a threshold value *t* is chosen to construct a truncated distance-based matrix D^*, where $d^*_{ij} = d_{ij}$, if $d_{ij} \leq t$ and $d^*_{ij} = 4t$, if $d_{ij} > t$. Third, the doubly-centered square matrix $-\frac{1}{2}(I - \frac{11^T}{n})D_2^*(I - \frac{11^T}{n})$ is diagonalized[2].

[2] Diagonalization, also known as eigenvalue decomposition, is simply the application of linear algebra to solve the equation $Au = \lambda u$. In the case of spatial filtering, *A* is an *n* x *n* square matrix of a doubly-centered square matrix (e.g. $-\frac{1}{2}\left(I - \frac{11^T}{n}\right)D_2^*\left(I - \frac{11^T}{n}\right)$ for the original PCNM approach), u is an *n* by 1 vector, and λ is the corresponding eigenvalue for each specific eigenvector *v*.

Here I denotes an identity matrix, 1 is an n by 1 vector of ones, D_2^* is an element-by-element squared matrix of D^*). The principal coordinates are obtained by scaling each of eigenvectors (u_k) obtained in the last (third) step to have length $\sqrt{\lambda_k}$, where λ_k is the eigenvalue corresponding to the k^{th} eigenvector. All eigenvectors with positive eigenvalues were selected as synthetic control (supplementary explanatory) variables in the original PCNM approach (for use in the regression function's primary equation).

As discovered earlier (by de Jong et al. [1984]), there exists a linkage between eigenvalues and Moran's I (MI), an index that measures the strength of spatial autocorrelation in a single variable (Anselin 1988):

$$MI = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_{ij} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

where W=[w_{ij}] is a spatial weight matrix and y a variable observed across a number of n spatial units. Roughly speaking, the larger/smaller a positive/negative eigenvalue is, the stronger the positive/negative spatial autocorrelation the associated eigenvector represents. Recognizing this association, Dray et al.'s (2006) improvement to the original PCNM approach permits *negative* spatial autocorrelation (as represented by negative eigenvalues), while omitting eigenvectors with small (in absolute terms) eigenvalues. Although negative spatial autocorrelation is not common in the fields of transport and land use, it has been observed in competitive biological processes. Dray et al.'s version of the PCNM method also allows for weight matrices with non-zero diagonal elements.

The main difference between Getis' and Griffith's filtering approaches is the manner in which the original variables are decomposed. Getis (1990,1995) used the difference between observed and expected local spatial statistics, G_i, to separate spatial from non-spatial effects. In other words, observed variables (y) were filtered using $x_i^* = x_i \left[\frac{W_i}{n-1}\right] / G_i(d)$, where W_i is the sum of W's i^{th} row, $G_i(d)$ is the observed Getis-statistic value (computed as $\frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j}, \forall i \neq j$), and $\frac{W_i}{n-1}$ is its expected value. Therefore, the difference between x_i and its filtered counterpart, x_i^*, can be interpreted as a spatial variable, with a positive value indicating clustering of higher values of x and a negative value indicating clustering (or spatial autocorrelation) of lower x values. Also, these newly-generated spatial variables (x_i^*) are associated, but not correlated with x.

In contrast, Griffith (2000a, 2000b) used orthogonal and uncorrelated map patterns (represented as eigenvectors obtained from a doubly centered contiguity matrix) characterizing the data set's spatial (or other relationship) structure to filter out meaningful spatial forces. In an unusual paper, Getis and Griffith (2002) compared their two approaches using government expenditures per capita across U.S. states. The two filtering methods yielded similar goodness-of-fit statistics, although the z-score of Moran's I test statistic for residuals switched signs: it was weakly negative in Getis' model but weakly positive in Griffith's model. In addition, results from both approaches were similar to that of an SAR model. Importantly, Griffith's approach was deemed preferable, thanks to its flexibility for application in non-linear model specifications. By contrast, Getis' approach requires that analysts have variables with a natural origin and a linear model

specification, thereby limiting its use (Patuelli et al. 2011). Another advantage of Griffith's filtering approach is the orthogonality of the eigenvectors, facilitating stepwise variable addition to the model specification (also known as "forward regression", as applied in Griffith and Peres-Neto [2006]).

A potential drawback of Griffith's approach lies in computing the eigenfunctions (which are the eigenvectors and their associated eigenvalues) for large data sets, which can be a formidable task[3]. However, in comparison with the more widely used SAR approach to spatial autocorrelation, such computations may be reasonable. The paper makes such a comparison, and hereafter refers only to Griffith's eigenvector-based approach when discussing the mechanics of spatial filtering.

The following sections discuss in some detail the nature of Griffith's (2000a, 2000b) spatial filtering technique and its application. Two case studies, for analysis of continuous land value data and binary land development data, with comparison to a special, more flexible SAR model specification, are provided.

METHODOLOGY: MATHEMATICS AND APPLICATIONS OF SPATIAL FILTERING

The crux of spatial filtering lies in the linkage between eigenfunctions (i.e., the eigenvectors and corresponding eigenvalues) and spatial autocorrelation. Essentially, the (exogenously specified) *W* matrix's eigenvectors are used as supplemental covariates in the regression, to "filter" out spatial autocorrelation, thereby allowing for more efficient estimation of primary covariates' parameters. Eigenfunction decomposition has been widely used in fields like control theory and imaging, but its usage in spatial analysis is relatively new. A thorough interpretation of eigenfunctions from a regional/spatial perspective can be found in Griffith (1996), which uses a 9-by-9 regular square grid and 3 cases of Canada's urban census tracts to provide natural interpretations of the eigenvectors associated with the largest eigenvalues.

In the regular square grid case, Griffith's (1996) largest two eigenvectors represented a pair of orthogonal gradients, and the third largest depicted a centrally positioned, exponentially declining map pattern. The orthogonal pattern meant that those two eigenvectors could be expressed as a linear combination of the coordinates of their grid cells' centroids. The negative exponential pattern is equivalent to $E = a + b \cdot \exp\{c[(x - x_c)^2 + (y - y_c)^2]\}^{0.5}$, where x_c and y_c are the coordinates of the region's or data space's center, and *a*, *b* and *c* are parameters to be estimated. Similar map patterns were found for the three census tracts examples (for different Canadian cities), with different sample sizes (*n*) and city compactness. Moran's *I* statistic across model residuals fell each time an eigenvector was added to the set of explanatory factors. Of note is that in contrast to a square tessellation, the topographic center of an irregularly-shaped region does not necessarily coincide with its geometric center. Also, the salient gradient and concentric map patterns (for regular tessellation) tend to dissipate as *n* increases and more irregularity exists.

In addition, as *n* increased, greater agreement was observed between the eigenvalues of the weight matrix *W* and the eigenvalues of its transformation matrix $\Omega = (I - 11^T/n)W(I - 11^T/n)$, with their correlations ranging from 0.97 for the small *n* case to approximately 1.0 for

[3] Interestingly, when the data come from a regular square grid, computing the eigenvalues and vectors can be done directly and quickly, in closed form.

the large *n* case. The two sets of eigenvalues exhibited an interlaced pattern (with higher values aligning with lower values); and, as *n* increases, the eigenvalues of Ω showed upper bound convergence to the eigenvalue of *C* (Griffith, 1996).

The matrix Ω is guaranteed to have an eigenvector of purely $1/\sqrt{n}$ values, corresponding to an eigenvalue of unity. One advantage of using Ω instead of *W* is the fact that the eigenvectors of Ω are orthogonal (and thus uncorrelated). In other words, Ω′s eigenvectors look a bit like the scaled principal components of the matrix *W*, though their mathematical derivations are quite different.

Spatial Filtering Based on Ω

Griffith (2000a) further investigated the use of Ω for filtered analysis of spatial data. The extreme (most negative and positive) values Moran's I for a specific spatial configuration (represented by connectivity matrix *C* or its row-standardized counterpart *W*) can be expressed as a function of Ω's eigenvalues (as per Tiefelsdorf and Boots [1995] and De Jong et al. [1984]):

$$\text{Moran's I} = \text{MI} = \frac{n}{1'W1} \cdot eigvalue(\Omega)$$

In other words, one could compute Moran's *I* for any set of numerical values (*y*) observed in a spatial size-*n* data set, and these are the normalized/scaled eigenvalues of Ω . Moreover, the first eigenvector (denoted as E_1) of matrix Ω is the vector of values yielding the strongest spatial autocorrelation (thus having the largest MI value) in the space *W*. It is thus the most important principal component of the spatial structure, as encoded by *W*. Ω's second eigenvector (E_2) offers the second largest eigenvalue or MI, and is orthogonal to (and thus uncorrelated with) the first eigenvector (Griffith, 2000a).

Using these principal components or top eigenvectors, Griffith (2000b) specified four linear spatial filter models (crime rates as a function of household incomes and house values in Columbus; family income as a function of population densities in Ottawa; population densities as a function of distance to CBD in Toronto; and wage rates as a function of unemployment rate and net migration rate in Ohio), with samples sizes ranging from 25 to 731. In general, noticeable spatial autocorrelation was present in the residuals of an unfiltered/standard ordinary least squares (OLS) model. The filtered OLS model performed moderately better than a SAR model with the same spatial weight matrix (*W*), as reflected by Moran's I for residuals in the filtered model being just half the size of that for the SAR model, on average. Griffith (2000b) then tested a 30-by-30 ($n = 900$) case of remotely sensed biomass values and controlled for all 279 eigenvectors that met his MI = 0.25 threshold, without any other, standard covariates. Neither this filtered OLS model nor a SAR model could address all of the spatial autocorrelation present in the data, however.

Selection of Eigenvectors

A key issue for filtered regression applications is the strategy to select meaningful eigenvectors to embody compelling spatial interactions. As noted above, Griffith (2000b) set a single MI threshold for all eigenvectors' inclusion. Griffith and Peres-Neto (2006) used a more efficient method by restricting attention to eigenvectors with positive spatial autocorrelation and then used stepwise regression to search among them for meaningful eigenvectors. Tiefelsdorf and Griffith (2007) compared the performance of two algorithms for selecting meaningful eigenvectors using

the natural log of bladder cancer mortality rates (as a function of the logs of lung cancer rates and population densities for 508 US State Economic Areas). As usual, stepwise regression was used to select the top-performing set of eigenvectors in order to optimally reduce the residual sum of squares, $RSS = \hat{\varepsilon}^{*T}\hat{\varepsilon}^{*}$, or increase the proportion of explained variation, $R^2 = 1 - \hat{\varepsilon}^{*T}\hat{\varepsilon}^{*}/[y^{\mathrm{T}}M_{(1)}y]$, where the $\hat{\varepsilon}^{*}$ are the estimates of the error terms and $M_{(1)}$ is an idempotent matrix: $I_n - \frac{1 \cdot 1^T}{n}$. To help ensure that the spatial pattern underlying each selected eigenvector is strong enough to merit inclusion in the regression equation, Tiefelsdorf and Griffith (2007) proposed minimizing the following Z-score objective function:

$$\min z[I(\hat{\varepsilon}^{*})] \equiv [\frac{y^T M_{(X|E)} W M_{(X|E)} y}{y^T M_{(X|E)} y} - E(I(\hat{\varepsilon}^{*}))]/[\mathrm{var}(I(\hat{\varepsilon}^{*}))]^{\frac{1}{2}}$$

where W is the spatial weight matrix, $E(I(\hat{\varepsilon}^{*}))$ is the expected value of residual's Moran's I (computed as $\frac{\mathrm{trace}(M_{(X|E)}WM_{(X|E)})}{n-k}$), and the variance of MI is $var\big(I(\hat{\varepsilon}^{*})\big) = \frac{2[(n-p)tr(M_{(X|E)}WM_{(X|E)})^2 - [tr(M_{(X|E)}WM_{(X|E)})]^2]}{(n-k)^2(n-k+2)}$, with $M_{(X|E)} \equiv I - X(X^TX)^{-1}X^T$. This is the t-statistic for a test of Moran's I (of the residuals) being statistically significant. In other words, minimizing this Z-score is akin to removing all evidence of spatial autocorrelation in the error terms.

Their first scenario involved a full set of eigenvectors, while their second focused on eigenvectors exhibiting positive spatial autocorrelation. They found that limiting their search to those exhibiting positive spatial autocorrelation greatly reduced the number of iterations to optimize the objective function used, both in minimizing the Z score on Moran's I and the RSS. Of interest is the fact that the spatial autocorrelation level (indexed by the Z score) did not monotonically fall under the RSS-minimization scheme. While Moran's I minimization scheme reached its conversion criterion faster than the RSS scheme, it required more computing and the difference in conversion rate was negligible when the search was confined to the set of eigenvectors exhibiting positive autocorrelation. Their RSS minimization scheme yielded four more eigenvectors, but the first eigenvectors selected under both strategies frequently identicalwith both tending to select eigenvectors with large spatial autocorrelation.

Applications of Spatial Filtering

Griffith (2004) compared his application of a logistic spatial filtering technique for binary response (y = 0,1) with results of Bayesian estimation and maximum pseudo-likelihood estimation and maintained that spatial filtering furnishes a successful way to account for spatial dependence in a logistic specification. Patuelli et al. (2011) added a spatial filtering component to a conventional random-effects model of German unemployment data over time (and across 439 districts). Chun and Griffith (2011) analyzed US interstate migration flows by utilizing a random-parameters model with spatial filter. In the context of network flows, the conventional n by n weight matrix is "enlarged" to an n^2 by n^2 matrix given there are n^2 combinations of origins and destinations for a region with n trip-generating and –attracting zones.

Spatial filtering's application to transportation is nascent. The majority of the spatial filtering literature comes from fields like ecology, optical imaging and economic analysis (e.g., Diniz-Filho and Bini 2005, Zhang et al. 2005, Tiefelsdorf and Griffith 2007, Jacob et al. 2008). Only

recently have Chun (2008) and Chun and Griffith (2011) used spatial filtering to model U.S. interstate migration flows.

This following discussion demonstrates the use of spatial filtering for analysis of land use and land value data, with a comparative look at continuous and binary SAR specifications (each with a distance-decay parameter, for added functional flexibility).

DATA SETS FOR ANALYSIS

Two related data sets were assembled to demonstrate spatial filtering techniques in a land use and land value context. One contains appraised values for private properties across Texas' Travis County in the year 2008, as obtained from County of Travis Central Appraisal District (TCAD). Five covariates were created for this regression: the shortest-path network distance from each parcel's centroid to Austin's central business district (*DistCBD*), the shortest Euclidean distance from each parcel centroid to the nearest freeway (*DistFwy*), the distance to the nearest major arterial (*DistMaj*), the distance to the nearest minor arterial (*DistMin*), and the parcel's *Slope*. Wang et al. (2011) provides more details on these covariates' calculations. Covariates of regional accessibility (using a logsum measure of destination choices in trip-making, as described in Wang et al. [2011] and Niemeier [1997]) and population density of each traffic analysis zone were also examined for inclusion in the models, but they exhibited too much collinearity with the constant term (due to a lack of variation in these variables across the sampled points). Table 1 summarizes key statistics of the land value data set.

The other data set characterizes when undeveloped, privately-held parcels in Travis County underwent development between 2003 and 2008. Undeveloped parcels are vacant parcels with the potential to develop, and thus exclude parks, preserved land, greenbelts and water. The original Travis County data set showed 48,445 undeveloped parcels in 2003, among which 42,589 parcels experienced no physical changes (subdivision or merging) over the 5 year period. Among these 42,589 parcels, 64.8 percent remained undeveloped through 2008. The vast majority (98.7%) of those that developed were developed into residential uses. In addition to the five covariates described above, the ratio between each parcel's perimeter and land area was controlled for in the binary-response model (where $y = 1$ if the parcel developed by 2008, and 0 otherwise). Table 2 provides summary statistics for this data set, with additional details provided in Wang et al. (2011).

Both of these data sets involve thousands of parcels, which turned out to be demanding for eigenvector calculation and selection (in reminding the model specifications, with synthetic control variables). So the sample sizes were reduced, as described below.

Table 1 Summary Statistics for Austin Land Value Data Set

	Unit	Min	Max	Mean	Median	StdDev
Land Value per Square Foot	$/SF	0	1,216	7.848	4.716	10.68
DistCBD	Mile	0	55.40	19.691	18.26	10.29
DistFwy	Mile	0	21.02	6.429	5.011	5.651
DistMajArt	Mile	0	6.139	0.803	0.483	0.855
DistMnrArt	Mile	0	10.62	1.770	1.474	1.368

Slope	Percent	0	65.49	7.614	5.654	6.703
# Obs.	307,331					

Table 2 Summary Statistics for Austin Land Development Data Set

	Unit	Min	Max	Mean	Median	StdDev
y	--	0	1	0.35	0	0.48
DistCBD	Mile	7.000E-01	43.84	19.95	18.55	10.44
DistMnrArt	Mile	3.307E-03	10.62	1.767	1.470	1.350
DistMajArt	Mile	2.354E-04	6.139	0.815	0.490	0.852
DistFwy	Mile	1.169E-02	21.02	6.611	5.414	5.661
Slope	Percent	0.01	74.60	7.829	5.829	6.874
PeriArea	1/ Ft.	0.001	0.462	0.042	0.044	0.023
# Obs.	42,589					

RESULTS OF DATA ANALYSIS

The *spdep* library (Bivand et al. 2011) of the open-source statistical software R includes methods to reduce residual spatial dependence in spatially filtered models. For example, the SpatialFiltering function, used here to select meaningful eigenvectors for the continuous land value model. It does this by maximizing the reduction in the MI value of residuals at each loop in the repeated-OLS estimation process. The process terminates when a maximum tolerance has been cleared for the Z score of the remaining MI value.

In calibrating this work's binary model of land development, R's Moran eigenvector (ME) fitering function for general linear models was used to find eigenvectors that represent significant spatial interactions. This function selects eigenvectors into the GLM model as synthetic variables until the MI of residuals of the continuous *y* (between 0 and 1, in this case) falls below a pre-specified tolerance level. In other words, both algorithms run through at most *n* of the *n* eigenvectors at each loop.

Such brute force selection of eigenvectors requires considerable computing effort, or smaller sample size. Given the very large data sets available for analysis here, two different methods of parcel selection were employed. For the continuous land value data set, one parcel was randomly selected from each of Austin's 590 traffic analysis zones (TAZs), ensuring reasonable geographic coverage of Travis County. For the binary land development data, a sample size of 1,000[4] was randomly selected across the county, resulting in greater coverage. For ease of visual representation, Theissen or Voronoi polygons[5] were created using the sampled parcels' centroids and ArcGIS polygon commands (to illustrate the results described below).

Land Value Model Results

[4] While the number of unselected eigenvectors falls as the selection process proceeds, *n* remains large in large data sets, over many iterations. R users report computing difficulty for sample size of 3,000+ on the R online forum, at https://stat.ethz.ch/pipermail/r-sig-geo/. A sample size of 3,000 resulted in excessively long computing times for the logit model

[5] Each Thiessen polygon has one input point, and any location within a polygon is closer to its input point than to that of any other polygon. Here, the input points are the centroids of each sampled parcel.

Table 3 summarizes the results of a SARα model, which is a conventional SAR model that allows for a distance decay parameter in its weight matrix: $y = \rho W_\alpha y + X\beta + \varepsilon$, where W_α is the row-standardized version of D with $D_{ij} = d_{ij}^{\alpha}$ if $d_{ij} < d_{\max}$, and $D_{ij} = 0$ if $d_{ij} \geq d_{\max}$ or $i = j$. $d_{\max}$ is the distance of the q^{th} nearest neighbor, with q selected to be 6 (somewhat arbitrarily, but, according to LeSage, it doesn't matter what q value one uses, since the indirect effects travel across neighbors in the spatial weight matrix). The model was estimated using Bayesian methods, as described in Wang et al. (under review). Table 4 shows the spatially filtered (SF) regression results, with all selected eigenvectors listed in Table 5. Both the SARα and SF models yield similar parameter estimates for the set of five covariates found in Table 1, but the SARα model tends to mask the statistical significance of covariates, as reflected in its consistently lower *t* statistics than those from SF model.

Using a tolerance level[6] of 0.1, 25 eigenvectors were selected into the SF-OLS model, yielding much higher R^2 and adjusted R^2 values. As one might expect in a setting with 25 added variables, all based on the weight matrix, Moran's *I* of the SF model's residual term is very low – and just 12 percent of MI value apparent in the SAR residuals, which still exhibit noticeable clustering.

The first four selected eigenvectors represent prominent spatial patterns across the Travis County TAZ map. The first depicts strong clustering in central and eastern regions. The second catches a clear cluster at the north-eastern corner and two smaller clusters at the northern and southern corners. The third captures two large clusters in the central and southwestern areas of the county, with the fourth mainly accounting for positive spatial interactions in the northwestern region.

The SF-OLS model took almost twice the time required for SAR model, since so many eigenvectors had to be repeatedly evaluated. When the tolerance level was raised to 4, the algorithm terminated after selecting 12 eigenvectors, and thus $\sum_{i=0}^{12}(590 - i)$ = 7,592 search loops, yielding a slightly lower but still satisfactory R^2 of 0.62 and a lower computing time (3,897 seconds) than that required by the SARα model. Thus, even under a moderate tolerance level, the SF-OLS technique may provide better fit with less computing burden than important, competing model specifications.

[6] The search algorithm terminates when the *Z* score for statistical significance of Moran's *I* falls below the analyst-provided tolerance value.

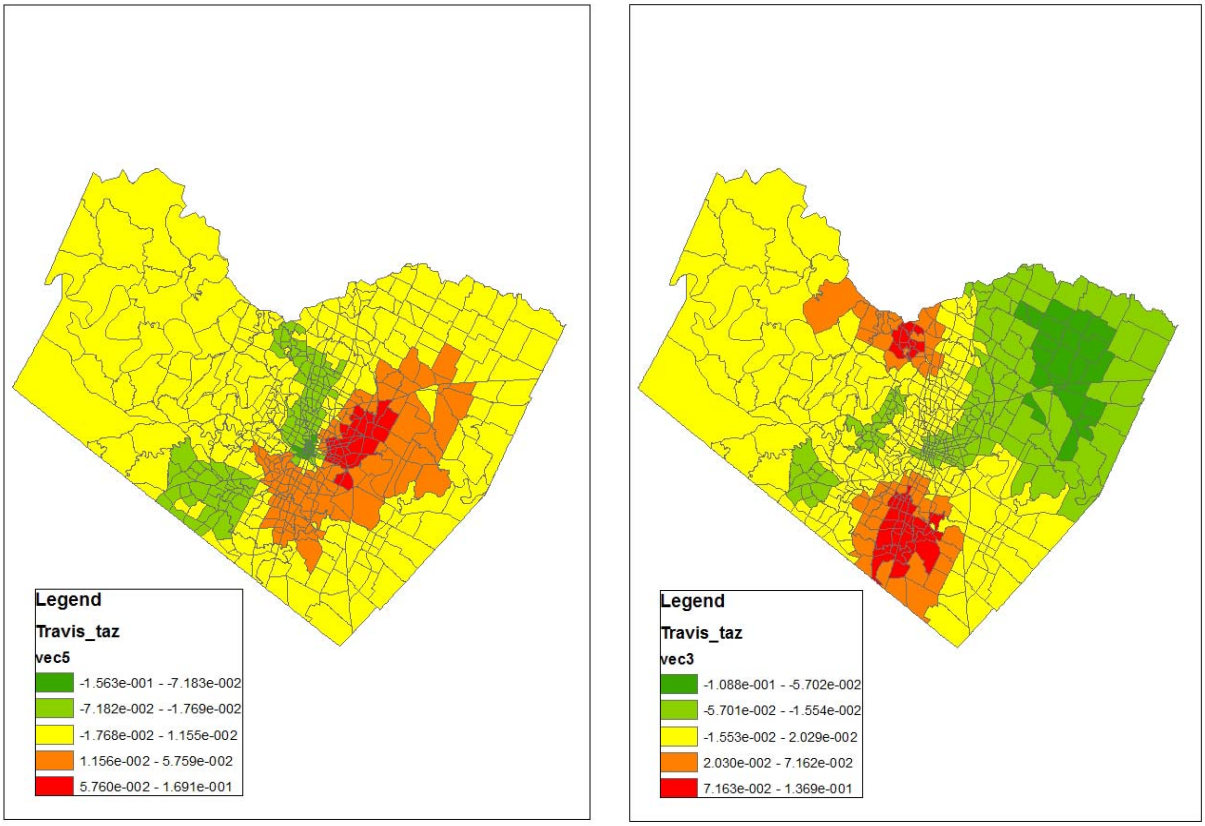

Figure 1. Top Two Selected Eigenvectors across Sampled Travis County Parcels (Left image shows loadings on the first selected eigenvector; right image shows loadings on the second.)

Land Development Model Results

As with the continuous-response modeling approach, top eigenvectors were simply added to the set of covariates in the main equation characterizing discrete response. In the case of a binary logistic model, the expression is as follows:

$$\text{Prob}(y_i = 1) = \frac{\exp\,(x_i\beta + E_i\gamma)}{1 + \exp\,(x_i\beta + E_i\gamma)}$$

where x_i is a 1 by k vector containing the systematic covariates for parcel i as well as a constant term. E_i is an 1 by p vector for the ith parcel's row of values in the top p eigenvectors selected. These E_i values characterize that parcel's "loadings" (a term from principal components analysis) on each of the included eigenvectors.

The ME search algorithm terminated after 12,922 search loops for a total of 4.81 minutes on an Intel(R) Xeon(R) 2.40GHz desktop computer, yielding 12 eigenvectors that represent significant spatial components as synthetic variables. As shown in Table 6, the likelihood of a parcel's development ($y = 1$) was estimated to be positively associated with distance to the CBD and distances to the nearest major arterial and freeway, suggesting that many parcels developing over the 2003-2008 period were on the periphery of Travis County, but relatively close to minor arterials and with somewhat lower slopes, on average, than parcels that remained undeveloped,

ceteris paribus. In addition, the larger the ratio of a parcel's perimeter to its area, the more likely it would undergo development.

Thiessen or Voronoi polygons were constructed around parcels' centroids to illustrate the eigenvector loadings across sampled parcels (so that parcels are visible, rather than being hidden among the tens of thousands available in the original data set). Figure 2 shows how the first two chosen/most relevant eigenvectors load. The first eigenvector captures clustering patterns at the northern and southern areas of the region under study, whereas the second depicts strong clustering pattern in the southwest corner, along with clusters sitting around the center.

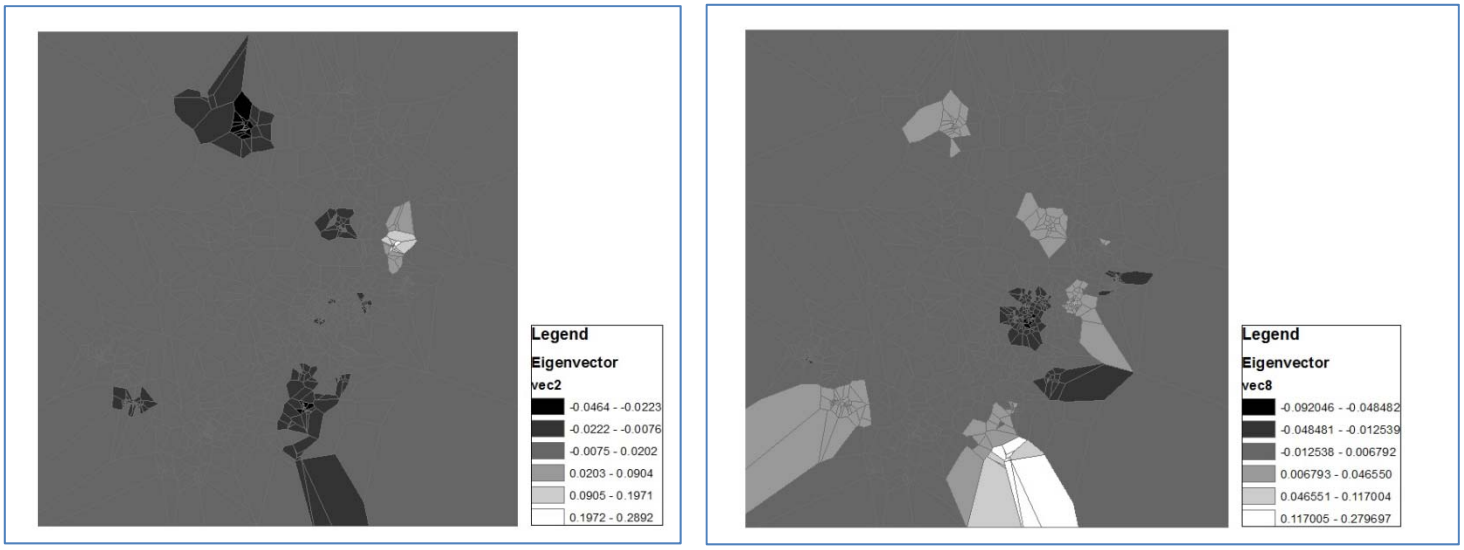

Figure 2. Top Two Selected Eigenvectors across Sampled Travis County Parcels (Left image shows loadings on the first selected eigenvector; right image shows loadings on the second.)

The same set of sampled points was analyzed using a binary probit SAR model with a distance decay parameter (to be estimated) in its weight matrix, with estimation results shown in Table 7. A spatial autoregressive probit model with a distance decay parameter (α) in the weight matrix was expressed as: $y^* = \rho W_\alpha y^* + X\beta + \varepsilon$, where y^* is the unobserved response variable with a positive value leading to a $y = 1$ outcome, and zero otherwise. ρ is the spatial autocorrelation coefficient characterizing the strength of spatial association in response values (after controlling for X factors), X is an n by K covariate matrix, with n denoting the number of observational units (e.g., parcels over space) under study and K the dimension of the parameter vector β. The error term, ε, is assumed to have an iid normal distribution: $\varepsilon \sim N(0, \sigma^2 I_n)$. To ensure identification (as present in any latent-response model), the homoscedastic error term is set to unity. Here, W_α is a function of the distance decay parameter α, such that $w_{ij,\alpha} = d_{ij}^{-\alpha}$, if $d_{ij} < d_{max}$, and $w_{ij,\alpha} = 0$ otherwise. W_α is the row-standardized version of D.

Both parameter estimates and quasi-t stats were smaller in magnitude than in the SF model, presumably due to larger relative scaling on the latent error term (which requires scaling for model identification). Essentially, when there is more noise in a probit or logit model, the slope parameters fall, relative to the noise term (whose scale is pre-specified). Interestingly, computing time was much higher for Bayesian estimation of the SARPα model than for the probit with spatial filtering.

Table 3 Parameter Estimates and Impact Estimates for SARα Model of Land Value

	Mean	Stdev	T-Stats	Direct Effect	Stdev_Dir	T-Stats	Indirect Effect	Stdev_Indir	T-Stats
Constant	8.865	1.168	7.589	9.451	1.251	7.557	14.244	2.693	5.289
DistCBD	-0.651	0.126	-5.163	-0.694	0.135	-5.152	-1.046	0.249	-4.197
DistMin	-0.271	0.378	-0.718	-0.288	0.402	-0.716	-0.434	0.620	-0.700
DistMaj	1.043	0.849	1.228	1.117	0.906	1.233	1.683	1.402	1.201
DistFwy	0.811	0.255	3.173	0.860	0.273	3.156	1.296	0.455	2.850
Slope	-0.181	0.147	-1.234	-0.192	0.156	-1.227	-0.289	0.244	-1.186
ρ	0.549	0.032	17.32						
α	-0.864	0.338	-2.557						
σ	101.3	6.203	16.33						
R^2	0.290								
Adjusted R^2	0.284								
Moran's I of ε	0.512								
Run time	4,369 seconds								

Table 4 Regression Results for Spatially Filtered Least-Squares Model of Land Value

	Estimate	Stdev	T-Stats	Pr(>|t|)
Constant	24.09	0.799	30.158	< 2e-16
DistCBD	-1.747	0.104	-16.84	< 2e-16
DistFwy	2.221	0.229	9.696	< 2e-16
DistMaj	2.051	0.776	2.642	0.00846
DistMin	-0.8192	0.347	-2.359	0.01866
Slope	-0.2761	0.136	-2.037	0.04209
vec5	-176.8	9.254	-19.1	< 2e-16
vec3	-58.99	9.254	-6.374	3.86E-10
vec4	-53.27	9.254	-5.756	1.42E-08
vec2	49.86	9.254	5.387	1.06E-07
vec14	37.54	9.254	4.057	5.68E-05
vec31	37.50	9.254	4.052	5.79E-05
vec48	38.77	9.254	4.189	3.25E-05
vec57	34.17	9.254	3.692	0.00024
vec19	26.24	9.254	2.836	0.00474
vec33	26.92	9.254	2.909	0.00377
vec61	-32.43	9.254	-3.504	0.00049
vec6	21.04	9.254	2.274	0.02335
vec7	20.49	9.254	2.214	0.02724
vec16	20.90	9.254	2.258	0.02431
vec15	20.48	9.254	2.213	0.02728
vec46	23.87	9.254	2.579	0.01016
vec23	20.47	9.254	2.212	0.02734
vec40	-22.29	9.254	-2.409	0.01634
vec75	28.06	9.254	3.032	0.00254
vec18	18.81	9.254	2.032	0.04259
vec39	20.71	9.254	2.238	0.02561
vec84	28.32	9.254	3.06	0.00232
vec66	-23.46	9.254	-2.535	0.01153
vec86	26.56	9.254	2.87	0.00426
vec92	25.73	9.254	2.78	0.00562
vec12	15.40	9.254	1.664	0.09666
R^2	0.668			
Adjusted R^2	0.65			
Run time	8,421 seconds			
Moran's I of ε	0.065			

Table 5 Eigenvectors Selected for Spatially Filtered Least-Squares Model of Land Value (with Tolerance = 0.1).

Step	Selected Eigenvector	Eigenvalue	Moran's *I* of Residuals	*Z* Score	Pr(*ZI*)	R^2	*r*
0	0	0.000	0.526	26.742	0.000	0.281	0.000
1	5	1.020	0.313	13.705	0.000	0.498	-176.8
2	3	1.038	0.276	11.610	0.000	0.522	-59.0
3	4	1.022	0.244	9.768	0.000	0.542	-53.2
4	2	1.051	0.213	7.936	0.000	0.559	49.9
5	14	0.923	0.197	7.040	0.000	0.569	37.5
6	31	0.758	0.184	6.298	0.000	0.579	37.5
7	48	0.607	0.173	5.667	0.000	0.589	38.8
8	57	0.535	0.166	5.247	0.000	0.597	34.2
9	19	0.880	0.157	4.809	0.000	0.602	26.2
10	33	0.739	0.149	4.411	0.000	0.607	26.9
11	61	0.497	0.143	4.027	0.000	0.615	-32.4
12	6	1.010	0.136	3.711	0.000	0.618	21.0
13	7	1.003	0.129	3.409	0.001	0.621	20.5
14	16	0.907	0.123	3.110	0.002	0.624	20.9
15	15	0.912	0.117	2.816	0.005	0.627	20.5
16	46	0.611	0.112	2.525	0.012	0.630	23.9
17	23	0.829	0.106	2.242	0.025	0.633	20.5
18	40	0.662	0.101	1.956	0.050	0.637	-22.3
19	75	0.408	0.096	1.672	0.094	0.642	28.1
20	18	0.890	0.090	1.405	0.160	0.645	18.8
21	39	0.675	0.085	1.140	0.254	0.648	20.7
22	84	0.354	0.081	0.873	0.383	0.653	28.3
23	66	0.465	0.077	0.624	0.533	0.657	-23.5
24	86	0.349	0.073	0.383	0.702	0.662	26.6
25	92	0.322	0.069	0.174	0.862	0.667	25.7
26	12	0.957	0.065	-0.022	0.982	0.668	15.4

Notes: *r* indicates the regression coefficient associated with each eigenvector. Pr(ZI) is the probability of the permutation-based standardized deviate for the given value of the alternative argument.

Table 6 Regression Results for Spatial Filtered Binary Logit Model of Land Development

	Estimate	Stdev	T stat.	z value	Pr(>|z|)	Marg. Effect
Constant	-1.945	0.292	-6.655	-6.655	2.84E-11	-
DistCBD	0.289	0.038	7.507	7.507	6.06E-14	0.156
DistMin	-0.560	0.122	-4.574	-4.574	4.79E-06	-0.302
DistMaj	0.498	0.308	1.616	1.616	0.106	0.268
DistFwy	0.456	0.156	2.928	2.928	0.003	0.246
Slope	-0.087	0.026	-3.377	-3.377	0.001	-0.047
Periarea	7.479	3.220	2.322	2.322	0.020	4.027
vec2	-11.398	3.656	-3.117	-3.117	0.002	-6.137
vec8	-18.906	4.500	-4.201	-4.201	2.65E-05	-10.180
vec20	-12.685	2.741	-4.628	-4.628	3.68E-06	-6.830
vec32	-11.481	2.961	-3.877	-3.877	0.000	-6.182
vec18	12.871	3.677	3.500	3.500	0.000	6.930
vec66	-10.554	2.920	-3.614	-3.614	0.000	-5.683
vec38	-7.053	2.419	-2.915	-2.915	0.004	-3.798
vec15	6.293	2.232	2.819	2.819	0.005	3.388
vec45	5.105	2.212	2.308	2.308	0.021	2.749
vec7	8.343	3.089	2.701	2.701	0.007	4.492
vec60	-6.914	2.912	-2.374	-2.374	0.018	-3.723
vec84	-6.606	2.529	-2.612	-2.612	0.009	-3.557
vec22	4.832	3.235	1.494	1.494	0.135	2.602
AIC	1170.2					
Likelihood Ratio	$254.3 > \chi^2_{0.05}(19) = 30.14$					
Moran's I of ε	0.103					
Run Time	288.6 seconds					

Significance codes: *** is for 0.000, ** is for 0.001, * is for 0.01, * is for 0.05, and 0.1
‘ ’

Table 7 Parameter Estimates and Impact Estimates for SARPα Model of Land Use.

Covariates	Mean	St.Dev.	T-Stat.	Direct Effect	Stdev_Dir	T-Stat.	Indirect Effect	Stdev_Indir	T-Stats
Constant	-0.571	0.139	-4.102	-	-	-	-	-	-
Dist2CBD	0.060	0.016	3.652	0.023	0.00103	34.744	0.025	3.24E-04	76.1
DistMin	-0.092	0.053	-1.753	-0.039	0.00115	-27.226	-0.038	3.18E-05	-1192
DistMaj	0.311	0.167	1.868	0.133	0.00515	27.188	0.128	9.66E-05	1327
DistFwy	0.038	0.063	0.608	0.016	0.00134	26.907	0.016	1.67E-06	9418
Slope	-0.025	0.014	-1.708	-0.011	3.89E-04	-27.077	-0.010	5.05E-06	-2000
PeriArea	4.689	1.737	2.699	1.974	0.070	28.196	1.928	0.006	342
ρ	0.526	0.044	12.086						
α	-0.670	0.191	-3.505						
AIC	2961.0								
Moran's I of e	0.046								
Run time	12,354 seconds								

SUMMARY AND CONCLUSIONS

The value of using spatial filtering for regression model estimation seems clear: the relatively straightforward decomposition of a spatial weight matrix offers a set of vectors that characterize the data set's spatial relationships. Inclusion of these values in the model's primary behavioral equation accounts for much information that would otherwise remain largely latent, in the model's error terms, even in SAR-type models, as illustrated here. Other benefits of a spatially filtered approach include avoidance of normalizing factor computation that SAR techniques require. However, spatial filtering is computationally burdensome in other ways, such as exhaustively identifying the top eigenvectors for model inclusion, particularly in large data sets (since the number of vectors equals sample size).

This paper compared results of spatial filtering models and relatively flexible SAR models (each allowing for a distance decay parameter in their weight matrices). Parameter estimates, inference, predictive power, and computing times improved. Despite improved model fit, spatially filtered models may rely on a large number of proxy variables, at odds with model parsimony. While such proxies may perform fairly well in substituting for missing covariates (characterizing spatial structure of the data), the interpretation of these synthetic variables remains largely ambiguous. Underlying causes of such dependence is largely unknow: are they generated by pure spatial interactions (e.g., attributes of nearby parcels and their land owners impacting others' outcomes) or it is caused by omitted explanatory variables (such as soil and air quality).

In conclusion, spatial filtering methods for spatial data sets are an appealing method for removing spatial relationships in analysis of spatial data, allowing the analyst to focus on the marginal effects of covariates of interest while appreciating the various components that characterize the data set's spatial structure. Such techniques are shown here in a positive comparison to results of reasonably flexible spatial autoregressive model specfications. Opportunities for continued investigation include analysis of count and categorical response variables, use of simulated data sets and other spatial techniques for more rigorous understanding of the benefits of a filtered approach.

ACKNOWLEDGEMENT

The authors are very grateful to Dr. Michael Tiefelsdorf and Dr. Daniel Griffith for nuances on spatial filtering, and to Dr. Roger Bivand for valuable information on model implementation. In addition, we are thankful to the U.S. National Science Foundation for providing financial support under Award SES-0818066, and to Mrs. Annette Perrone for her editorial assistance. Of course, we alone are responsible for any omission and errors.

REFERENCES

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers: Norwell, MA.

Anselin, L. and Hudak, S. (1992) "Spatial econometrics in practice: A review of software options." *Regional Science and Urban Economics* 22(3): 509-536.

Bivand, R. (2011) spdep: Spatial dependence: weighting schemes, statistics and models. Accessed time: July 6 2011. URL: http://cran.r-project.org/web/packages/spdep/index.html.

Boots, B. and Dufournaud, C. (1994) A programming approach to minimizing and maximizing spatial autocorrelation statistics. *Geographical Analysis* Vol. 26:54-66.

Chakir, R. and Parent, O. (2009) Determinants of Land Use Changes: a Spatial Multinomial Probit Approach. *Papaers in Regional Science* Vol. 88 (2): 327-344.

Chun, Y. (2008) Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographical Systems*, Vol. 10 (4): 317-344.

Chun, Y., Griffith, D. A (2011) Modeling Network Autocorrelation in Space-Time Migration Flow Data: An Eigenvector Spatial Filtering Approach. *Annals of the Association of American Geographers*, Vol. 101 (3): 523-536.

Cressie, N. (1991) *Statistics for Spatial Data*, Wiley, New York.

De Jong, P., Sprenger, C., van Veen, F. (1984) On extreme values of Moran's I and Geary's c. *Geographic Analysis*. Vol. 16: 17-24.

Diniz-Filho, J. A. and Bini, L. M. (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Global Ecology and Biogeography*, Vol. 14 (2): 177-185.

Dray, S., Legendre, P., and Peres-Neto, P. R. (2006) Spatial modeling: a comprehensive framework for principal coordinate analysis of neighbor matrices (PCNM). *Ecological Modeling*, Vol. 196: 483-493.

Fotheringham, S. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons Ltd., West Sussex, England, 2003.

Getis, A. (1990) Screening for spatial dependence in regression analysis. *Papers of the Regional Science Association* Vol.69: 69–81.

Getis, A. (1995) Spatial filtering in a regression framework: experiments on regional inequality, government expenditures, and urban crime. In *New Directions in Spatial Econometrics*. Berlin: Springer.

Getis, A. and Griffith, D. (2002) Comparative spatial filtering in regression analysis. *Geographical Analysis* Vol. 34 (2): 130-140.

Ghosh, D. and Manson, S. (2008) Robust Principal Component Analysis and Geographically Weighted Regression: Urbanization in the Twin Cities Metropolitan Area (TCMA). *URISA Journal* 20 (1): 15-25.

Griffith, D. (1984) Re-examining the question "Are locations unique?", *Progress in Human Geography*, 8: 82-94.

Griffith, D. A. (1996) Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying geo-referenced data. *The Canadian Geographer*, Vol. 4: 351-367.

Griffith, D. A. (2000a) A linear regression solution to the spatial autocorrelation problem. Journal of Geographical Systems, Vol. 2:141-156.

Griffith, D. A. (2000b) Eigenfunction properties and approximations of selected incidence matrices employed in spatial analyses. *Linear Algebra & Its Applications*, Vol. 321: 95-112.

Griffith, D. A. (2004) A spatial filtering specification for the auto-logistic model. *Environment and Planning A*, Vol. 36: 1791-1811.

Griffith, D. A., Peres-Neto, P. R. (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analysis. *Ecology*, Vol. 87 (10): 2603-2613.

Griffith, D.A. (2008) Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A*. Vol.40: 2751-2769.

Hollander, M., Wolfe, D. A. (1999) *Nonparametric Statistical Methods*, 2nd Edition. Wiley.

Jacob, B. G., Griffith, D. A., Novak, R. J. (2008) Decomposing malaria mosquito aquatic habitat data into spatial autocorrelation eigenvectors in a SAS/GIS® module. *Transactions in GIS*, Vol. 12 (3): 341-364.

Kaiser, M., Cressie, N. (1997) Modeling Poisson variables with positive spatial dependence. *Statistics and Probability Letters* 35:423-432.

Kosfeld, R. and Dreger, C. (2006) Thresholds for employment and unemployment: A spatial analysis of German regional labor markets, 1992–2000. *Papers in Regional Science*, Vol. 85 (4): 523-542.

Niemeier, D. A. (1997) Accessibility: An Evaluation Using Consumer Welfare. *Transportation*, Vol. 24, pp. 377-396.

Patuelli, R., Griffith, D. A., Tiefelsdorf, M., Nijkamp, P. (2011) Spatial Filtering and Eigenvector Stability: Space-Time Models for German Unemployment Data. *International Regional Science Review*, Vol. 34 (2): 253-280.

Tiefelsdorf, M. and Griffith, D. A. (2007) Semiparametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A*, Vol. 39: 1193-1221.

Yuan, J. and Mills, K. (2005) A cross-correlation-based method for spatial–temporal traffic analysis. *Performance Evaluation*, Vol. 61 (2-3): 163-180.

Zhang, Y., Brooks, D. H., Franceschini, M. A., Boas, D. A. (2005) Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *Journal of Biomedical Optics*, Vol. 10 (1): 011014-1-11.

McMillen, D.P. and McDonald, J.F. (1999) Land Use before Zoning: The Case of 1920's Chicago. *Regional Science and Urban Economics* 29 (4): 473-489.

McMillen, D. P. (2004) Employment densities, spatial autocorrelation, and subcenters in large metropolitan areas. *Journal of Regional Science*, Vol. 44 (2): 225-243.

Munroe, D.K., Southworth, J. and Tucker, C. M. (2002) The Dynamics of Land-Cover Change in Western Honduras: Exploring Spatial and Temporal Complexity. *Agricultural Economics* 27 (3): 355-369.

Páez, A. (2006) Exploring Contextual Variations in Land Use and Transport Analysis Using a Probit Model with Geographical Weights. *Journal of Transport Geography* 14: 167–176.

Tiefelsdorf, M, Griffith, D. A. (2007) Semiparametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A*, Vol. 39: 1193-1221.

Tiefelsdorf M., Boots, B. (1995) The exact distribution of Moran's I. *Environment and Planning A* 27: 985-999.

Wang, X., Kockelman, K. M., Lemp, J. D. (2011) The dynamic spatial multinomial probit model: analysis of land use change using parcel-level data. Proceedings of the 90th Annual Meeting of the Transportation Research Board. Available at http://www.ce.utexas.edu/prof/kockelman/public_html/TRB11DSMNP.pdf.

Wang, X., and Kockelman, K. (2009a) Bayesian Inference for Ordered Response Data with a Dynamic Spatial-Ordered Probit Model. *Journal of Regional Science* 49 (5): 877-913.

Wang, X. and Kockelman, K. (2009b) Application of the Dynamic Spatial Ordered Probit Model: Patterns of Land Development Change in Austin, Texas. *Regional Science* 88 (2): 345-366.

Wang, X. and Kockelman, K. (2009c) Application of the Dynamic Spatial Ordered Probit Model: Patterns of Ozone Concentration in Austin, Texas. *Transportation Research Record* No. 2132: 13-24.

Zhou, B., and Kockelman, K. (2009) Predicting the Distribution of Households and Employment: A Seemingly Unrelated Regression Model with Two Spatial Processes. *Journal of Transport Geography* 17: 369-376.