Bayesian Spatially Varying Coefficient Models in the Presence of Collinearity

David C. Wheeler[1], Catherine A. Calder[1]
The Ohio State University[1]

Abstract

The belief that relationships between explanatory variables and a response variable in a regression model may vary within a study area has lead to the development of Bayesian regression models that allow for spatially varying coefficients (Gelfand et al., 2003). In the typical application of these spatially varying coefficient process (SVCP) models, marginal inference on the spatial pattern of regression coefficients is of central interest. In light of this, there is a need to assess the validity of these marginal posterior inferences, since these inferences can be misleading in the presence of explanatory variable correlation (i.e. collinearity). We present the results of a simulation study designed to assess the sensitivity of spatially varying coefficients to a range of levels of collinearity. The results show that the SVCP model is overall fairly robust to moderate levels of collinearity in terms of marginal coefficient inference, but degrades in coefficient accuracy with strong collinearity. We also illustrate that the posterior mean of the SVCP model coefficients can be viewed as ridge regression solutions with the amount of coefficient penalization controlled by numerous model parameters. Finally, we present an application of the spatially varying coefficient model to a cancer dataset, where the relationship between cancer rates and some explanatory variables is suspected to vary spatially.

Keywords: MCMC, regression, simulation study, spatial statistics

1. Introduction

In the statistics literature, Bayesian regression models with spatially varying coefficient processes have been introduced to model linear relationships between variables, where the relationships are not necessarily constant (Gelfand et al. 2003). For convenience, we refer to this Bayesian model as SVCP for spatially varying coefficient process model. The motivation for these models is that, in certain applications, regression coefficients could vary at the regional or local level. In the SVCP framework, the spatial varying coefficients are modeled in the form of a multivariate spatial process with the dependence between regression coefficients defined globally. Such an approach fits naturally into the Bayesian paradigm, as parameters are considered unknown random quantities.

One topic that is underrepresented in the spatial varying coefficient regression model literature is the validity of marginal inference on the regression coefficients, especially in the presence of correlation in the explanatory variables. The inherent assumption in works that apply spatially varying coefficient models to data is that the regression coefficients are free of strong dependence and, hence, appropriate for marginal interpretation. This is clearly an important assumption. To the best of our knowledge, there are no published papers that assess the validity of inferences derived from the SVCP model. However, it is well understood that in linear regression models, strong correlation in the explanatory variables can increase the variance of the estimated regression coefficients. This provides the motivation for this paper, in which we assess the accuracy of the estimated regression coefficients from the SVCP model through a simulation study, where the 'true' and fixed values of the regression coefficients are known. We evaluate the accuracy and coverage probabilities of the regression coefficients in both the absence of collinearity and in the presence of a range of levels of collinearity. We also illustrate that the posterior mean of the SVCP model coefficients can also be viewed as ridge regression solutions with the amount of coefficient penalization controlled by numerous model parameters. Finally, we present an application of the spatially varying coefficient model to a cancer dataset, where the relationship between cancer rates and some explanatory variables is suspected to vary spatially.

2. Bayesian Spatially Varying Coefficient Process (SVCP) Model

The Bayesian SVCP regression model is conveniently specified in a hierarchical manner, where the distribution of the data is specified conditionally on numerous unknown parameters, whose distribution is in turn specified conditionally on numerous other parameters. Following Gelfand et al. (2003), the SVCP model is

$$\left[\mathbf{Y} \mid \boldsymbol{\beta},\, \tau^2\right] = N(\mathbf{X}^T\boldsymbol{\beta},\, \tau^2\mathbf{I}), \qquad (1)$$

where $\mathbf{Y}$ is a vector of responses assumed to be Gaussian, conditional on the unknown parameters $\boldsymbol{\beta}$ and τ^2; $\boldsymbol{\beta}$ is a $np \times 1$ vector of regression coefficient parameters; and $\mathbf{X}^T$ is the $n \times np$ block diagonal matrix of covariates where each row contains a row from the $n \times p$ design matrix $\mathbf{X}^*$, along with zeros in the appropriate places (the covariates from $\mathbf{X}^*$ are shifted p places in each subsequent row in $\mathbf{X}^T$); $\mathbf{I}$ is the $n \times n$ identity matrix; and τ^2 is the error variance.

The prior distribution for the regression coefficient parameters is specified as

$$\left[\boldsymbol{\beta} \mid \boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}\right] = N(\mathbf{1}_{n\times 1} \otimes \boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}). \qquad (2)$$

The vector $\boldsymbol{\mu}_{\boldsymbol{\beta}} = (\mu_{\beta_0}, \ldots, \mu_{\beta_p})^T$ contains the means of the regression coefficients corresponding to each of the explanatory variables, of which there are p. The prior on the regression coefficients in the SVCP model takes into consideration the potential spatial dependence in the coefficients through the covariance, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$. For $\boldsymbol{\beta}_p = [\beta_{p1}, \ldots, \beta_{pn}]$, we assume a priori that each $\boldsymbol{\beta}_p$ follows an areal unit model (e.g., the CAR or SAR model; see Banerjee et al. 2004) or specify the prior on $\boldsymbol{\beta}_p$ using a geostatistical approach, where a distance-based covariance function is specified with a parameter. We consider a geostatistical prior specification of the regression coefficients and utilize an exponential spatial dependence function. The prior covariance matrix for the p different types of $\boldsymbol{\beta}$'s at each of n locations, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$, may have either a nonseparable or separable form. The separable form has two distinct components, one for the within site dependence between coefficients of the same type and one for the spatial dependence in the regression coefficients. Following Gelfand et al. (2003), we make the assumption of a separable covariance matrix for $\boldsymbol{\beta}$ with the form

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \mathbf{H}(\phi) \otimes \mathbf{T}, \qquad (3)$$

where $\mathbf{T}$ is a positive-definite $p \times p$ matrix for the covariance of the regression coefficients at any spatial location, $\mathbf{H}(\phi)$ is the $n \times n$ correlation matrix that captures the spatial association between the n locations, ϕ is the unknown spatial dependence parameter, and $\otimes$ denotes the Kronecker product operator. In the prior specification for $\boldsymbol{\beta}$, the Kronecker product structure results in a $np \times np$ positive definite covariance matrix because $\mathbf{H}(\phi)$ and $\mathbf{T}$ are both positive definite. The elements of the correlation matrix $\mathbf{H}(\phi)$, $H(\phi)_{ij} = \rho(s_i - s_j; \phi)$, are calculated from the (stationary and isotropic) exponential correlation function $\rho(h;\phi) = \exp(-h/\phi)$.

The specification of the Bayesian SVCP model in equations (1) and (2) is finalized with the specification of the prior distributions of the parameters. The prior for the coefficient means is normal with hyperparameters $\boldsymbol{\mu}$ and σ^2, $\left[\boldsymbol{\mu}_{\boldsymbol{\beta}}\right] \sim N(\boldsymbol{\mu}, \sigma^2\mathbf{I})$. The prior for the covariance matrix $\mathbf{T}$ is inverse Wishart with hyperparameters v and $\boldsymbol{\Omega}$, $\left[\mathbf{T}\right] \sim IW_v(\boldsymbol{\Omega}^{-1})$. The prior for the error variance is inverse gamma with hyperparameters a and b, $\left[\tau^2\right] \sim IG(a,b)$. These priors are conjugate and are used for computational convenience. The prior for the spatial dependence parameter ϕ is gamma with hyperparameters α and λ, $\left[\phi\right] \sim G(\alpha, \lambda)$. The posterior distributions of the SVCP model parameters are obtained using Markov chain Monte Carlo (MCMC). [DW1]

3. Simulation Study

In this section, we use a simulation study to evaluate the accuracy and coverage probabilities of the regression coefficients from the Bayesian SVCP model. The motivation for doing this study is to test the assumption that the inferences on the coefficients from the SVCP model are valid. The inferences are considered valid if the 95% credible interval for each estimated coefficient contains the true value approximately 95 percent of the time. We first evaluate the assumption of acceptable coverage when there is no collinearity in the model and then specify systematic increases in collinearity to inspect its effect on both the accuracy and the coverage probabilities of the regression coefficients and the strength of correlation in the estimated coefficients at each data point in the study area and across the entire study area.

The model we used to generate the data for the simulation study is

$$y(s) = \beta_1^*(s)x_1(s) + \beta_2^*(s)x_2(s) + \varepsilon(s), \quad (4)$$

where the x_1 and x_2 are the first two principal components from a random sample drawn from a multivariate normal distribution of dimension ten with a mean vector of zeros and an identity covariance matrix, and the errors $\boldsymbol{\varepsilon}$ are sampled independently from a normal distribution with mean 0 and variance of τ^{2*}. The star notation denotes the true values of the parameters used to generate the data. We note that there is no true intercept in the model used to generate the data and we do not fit an intercept in the simulation study. The data points are equally spaced on a 10×10 grid, for a total of 100 observations. The goal of the simulation study is to use the model in equation (4) to generate the data and then evaluate whether the regression coefficient estimates match $\boldsymbol{\beta}^*$ for the SVCP model.

For the simulation study, we start with no collinearity in the model and systematically add collinearity until the explanatory variables are nearly perfectly collinear. This is accomplished by replacing one of the original explanatory variables with one created from a weighted linear combination of the original explanatory variables, where the weight determines the amount of correlation of the variables. The formula for the generated weighted variable is

$$x_2^c = c \cdot x_1 + (1-c) \cdot x_2, \quad (5)$$

where x_2^c replaces x_2 in the model in equation (4) and c is a weight between 0 and 1.

The simulation study uses the following values to generate the data: the error variance $\tau^{2*} = 1$; the coefficient covariance matrix at all locations is set to $\mathbf{T}^* = \begin{pmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{pmatrix}$; the coefficient means are set to $\boldsymbol{\mu}_\beta^* = (1, 5)$; and the spatial dependence parameter $\phi^* = 10$. We fix these values and use them to simulate the true coefficients using the multivariate normal distribution. We simulate 200 realizations of the coefficient process in this simulation study, and hence assess the validity of $\boldsymbol{\beta}$ for 200 datasets. For each of the simulated datasets, we fit the SVCP model using an MCMC algorithm, which is run for 2000 iterations with a “burn-in” of 1000 iterations. Based on trace plots and Gelman’s $\hat{R}$ statistic (e.g. Gelman et al. 2003), the regression coefficients converged for individual realizations of the coefficient process after an initial 1000 “burn-in” iterations.

The prior specification for the SVCP model parameters is the following. We use a vague normal, $N(\mathbf{0}, 10^4\mathbf{I})$, for $\boldsymbol{\mu}_{\boldsymbol{\beta}}$, a three-dimensional inverse Wishart, $IW(3, .1\cdot\mathbf{I})$, for $\mathbf{T}$, and an inverse gamma, $IG(1, .01)$, for τ^2, where $\mathbf{I}$ is the identity matrix of dimension p. For the spatial dependence parameter ϕ, we use a gamma, $G(.021, .01)$, which has a mean of 2.12 and variance of 212. The hyperparameters for this gamma prior are chosen to have a large variance and a mean that solves the spatial correlation function set equal to .05, where the spatial range is set to half the maximum inter-point distance from the distance matrix $\mathbf{D}$. The spatial range is the distance beyond which the spatial association becomes negligible. Distributions that set prior mean spatial ranges to roughly half the maximum pairwise distance usually result in stable MCMC behaviour (Banerjee and Johnson 2005). The calculation for the mean is found from solving $\exp(-(1/2\cdot\max(\mathbf{D}))/\phi) = .05$ for ϕ.

To evaluate the accuracy and coverage probabilities of the regression coefficients, we calculate numerous summary statistics. For each realization of the process, we calculate the 95% credible intervals for each coefficient in the SVCP model. To calculate the 95% coverage probabilities, the true coefficients are first compared to the 95% intervals obtained for the respective coefficient in each data realization and then the total number of realizations that contain the true values are totalled and divided by the number of realizations. The means of the coverage probabilities are calculated for each explanatory variable from the corresponding coefficients to create a summary measure for all the realizations that is easy to present in a table. The accuracy of the regression coefficients is measured by calculating the root mean square error (RMSE) of the posterior mean of the coefficients at each realization. The average RMSE for all the realizations is calculated by averaging the RMSE’s from all of the individual realizations. We also calculate the overall correlation (C_{12}) between the two types of variable coefficients and the local coefficient correlation (C_{12}^s) for each realization, where the overall correlation is the Pearson correlation coefficient and the local coefficient correlation of the regression coefficients k and l across all locations is $T_{kl} \big/ \sqrt{T_{kk} T_{ll}}$.

The results of the simulation study are listed in Table 1. The results show that the SVCP model does fairly well at covering the true coefficient values used to generate the data, and the coverage probabilities are not dramatically affected by collinearity. The coverage probabilities remain around .90 until there is very strong collinearity, when they increase to .93. However, the average RMSE of the coefficients from the many realizations increases systematically with increasing collinearity. The SVCP model also does a good job at controlling the average local coefficient correlation (C_{12}^s in the table), but does not do as well at controlling the overall correlation (C_{12} in the table) of the two types of variable coefficients. Overall, the SVCP model performs well in this simulation study with regard to inference on the regression coefficients, although the accuracy of the coefficients decreases with increasing collinearity.

4. SVCP Model as Ridge Regression

It is not completely unexpected that the SVCP model faired well in the simulation study in the presence of collinearity given the discussions in the literature of ridge regression solutions as Bayes estimates. Hoerl and Kennard (1970) first introduced ridge regression to overcome ill conditioned design matrices. Ridge regression coefficients minimize the residual sum of squares along with a penalty on the size of the squared coefficients as

$$\hat{\boldsymbol{\beta}}^R = \underset{\boldsymbol{\beta}}{\arg\min}\left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} x_{ik}\beta_k \right)^2 + \lambda \sum_{k=1}^{p} \beta_k^2 \right\}, \quad (6)$$

where λ is the ridge regression parameter that controls the amount of shrinkage in the regression coefficients (Hastie et al. 2001). Works by Lindley and Smith (1972) and Goldstein (1976) show that ridge regression coefficient estimates may be viewed as Bayesian regression coefficient posterior means under specific vague priors. Hastie et al. (2001) also describe the ridge regression solutions as Bayes estimates, where ridge regression uses independent normal distributions for each coefficient β_k prior. If the prior for each regression coefficient β_k is $N(0,\sigma^2)$, independent of the others, then the negative log posterior density of the regression coefficients $\boldsymbol{\beta}$ is equal to the expression in the braces in the ridge regression coefficient equation (6), with $\lambda = \tau^2 \big/ \sigma^2$, where τ^2 is the error variance. Specifically, in the Bayesian regression model with $y \sim N(\mathbf{X}\boldsymbol{\beta}, \tau^2\mathbf{I})$ and the independent prior $\beta \sim N(0,\sigma^2)$ for each coefficient, the posterior for the coefficients can be expressed as

$$[\boldsymbol{\beta} \mid \tau^2, \sigma_\beta^2; y] \propto \exp\left(-\frac{1}{2\tau^2} \sum_{i=1}^{n} (y_i - \sum_{k=1}^{p} x_{ik}\beta_k)^2 \right) \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^{p} (\beta_k - 0)^2 \right), \quad (7)$$

where for convenience of notation the variables have been centered. The negative log posterior density of $\boldsymbol{\beta}$ up to a constant is then found through algebra to be

$$\sum_{i=1}^{n} (y_i - \sum_{k=1}^{p} x_{ik}\beta_k)^2 + \frac{\tau^2}{\sigma^2} \sum_{k=1}^{p} \beta_k{}^2, \quad (8)$$

with the ridge shrinkage parameter $\lambda = \dfrac{\tau^2}{\sigma^2}$. This illustrates that the ridge regression estimate is the mean of the posterior distribution with a Gaussian prior and Gaussian data model, and that the ridge shrinkage parameter is a ratio of the error variance and common regression coefficient variance.

The view of ridge regression solutions as Bayes estimates suggests that the Bayesian SVCP model coefficients can be viewed as ridge regression estimates because of the normal prior for the regression coefficients in the SVCP model. The posterior distribution for the coefficients of the SVCP model can be expressed with convenient notation as

$$[\beta \mid \tau^2, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}; y] \propto \exp(-\frac{1}{2}[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + (\boldsymbol{\beta} - (\mathbf{1} \otimes \boldsymbol{\mu}_{\boldsymbol{\beta}}))^T \cdot \tau^2 (\mathbf{H}(\phi) \otimes \mathbf{T})^{-1} (\boldsymbol{\beta} - (\mathbf{1} \otimes \boldsymbol{\mu}_{\boldsymbol{\beta}}))]). \quad (9)$$

The negative log posterior density of $\boldsymbol{\beta}$ up to a constant is then

$$(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+(\boldsymbol{\beta}-(\mathbf{1}\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}}))^T\tau^2\cdot (\mathbf{H}(\phi)\otimes\mathbf{T})^{-1}(\boldsymbol{\beta}-(\mathbf{1}\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}})), \quad (10)$$

where the shrinkage term is unconventionally a matrix $\boldsymbol{\lambda}$ that is calculated as $\tau^2\cdot\mathbf{H}^{-1}(\phi)\otimes\mathbf{T}^{-1}$. Therefore, the amount of shrinkage on $\boldsymbol{\beta}$ towards the mean $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ depends on τ^2, ϕ, and $\mathbf{T}$ in the SVCP model.

5. Bladder Cancer Mortality Example

To further motivate the issue of collinearity in spatially varying coefficient regression models and show an example with actual data, we use a simple SVCP model to explain white male bladder cancer mortality rates in the 508 State Economic Areas (SEAs) of the United States for the years 1970 to 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). In modeling these data we are interested in the spatially varying effects of log population density and lung cancer mortality rate, where population density is log transformed to linearize the relationship with the dependent variable. The model is

$$y(s)=\beta_0(s)+\beta_1(s)\cdot \mathrm{SMOKE}(s)+\beta_2(s)\cdot \mathrm{LNPOP}(s)+\varepsilon(s). \quad (11)$$

Lung cancer mortality rate is used as a proxy for 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, hence, 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). Population density is used as a proxy for environmental differences with respect to an urban/rural dichotomy. It is expected, as several studies have pointed out, that with an increase in the population density there is an increase in the rate of bladder cancer.

A traditional regression model was first built for bladder cancer mortality. The risk factors are both significantly positively related to the rate of bladder cancer. The estimated regression coefficients are 0.03 for smoking and 0.27 for population density. The variance inflation factors, a traditional method for diagnosing collinearity (e.g. Neter et al. 1996), for the two global explanatory variable parameters are less than 2 and the correlation of the global regression parameters is moderately negative at -0.59, whereas the correlation of the two variables is 0.59. These results suggest that collinearity is not an issue in this regression model.

Collinearity does appear to be a problem with these data, however, when used in a frequentist spatially varying coefficient regression model called geographically weighted regression (GWR). GWR is a collection of weighted least squares models at the study area locations, where the weights are calculated using a kernel function that is inversely related to distance (see Fotheringham et al. 2002 for details). Wheeler and Tiefelsdorf (2005) illustrate that the GWR estimated coefficients for these bladder cancer data are negative for both explanatory variables in some parts of the study area and the coefficients for the two variables exhibit moderate to strong overall correlation. For illustrative purposes, the GWR estimated coefficients are shown in Figure 1. The coefficients for population density are negative for most of the Northeast. These negative coefficients are counter to previous studies, intuition, and the traditional regression estimates. As Lindley and Smith (1972) point out, when data are correlated, least-squares regression can "produce regression estimates which are too large in absolute value, of incorrect sign and unstable with respect to small changes in the data." The weighted least-squares procedure of GWR likely suffers from the same condition.

In contrast to the GWR coefficients, the SVCP model estimated coefficients do not indicate a considerable problem with collinearity. To estimate the model parameters, we use 2000 iterations in the MCMC, with a "burn-in" of 1000 iterations. Based on trace plots and the $\hat{R}$ statistics, the regression coefficients converged within 1000 iterations of the Gibbs sampler. The prior specification for this model is as follows. We use a vague normal, $N(\mathbf{0}, 10^4\mathbf{I})$, for $\boldsymbol{\mu}_{\boldsymbol{\beta}}$, a four-dimensional inverse Wishart, $IW(4, .1\cdot\mathbf{I})$, for $\mathbf{T}$, and an inverse gamma, $IG(1, .01)$, for τ^2, where $\mathbf{I}$ is the identity matrix of dimension p. For the spatial dependence parameter ϕ, we use a gamma, $G(.103, .01)$, which has a mean of 10.37 and variance of 1037.

The SVCP model coefficients are mapped in Figure 2 and are all non-negative for smoking and non-negative for population density for all but two of the 508 SEAs, where the two negative population density SEAs are very close to zero. While there is some overall correlation in the SVCP model coefficients for the two variables, the complimentary pattern is not as strong as with the GWR coefficients. In addition, the variance of the coefficients is not as large with the SVCP model. The SVCP model here achieves a similar penalization effect to that of ridge regression, but has the advantage that the shrinkage parameters are estimated from the data and not through a separate estimation procedure such as cross-validation.

6. Conclusions

In the statistics literature, there has been an increasing interest in recent years in spatially varying relationships between variables. Attempts at modeling these relationships have produced Bayesian regression models with spatially varying coefficient processes (SVCP). While this type of model has been applied to real datasets in the literature, there has been a noticeable lack of emphasis on the validation of marginal inferences derived from its coefficients. We use a simulation study to assess the validity of the regression coefficients from the Bayesian SVCP model, while considering the presence of collinearity, which is often a feature of real data. Our results show that the SVCP model coefficients are fairly robust to collinearity in terms of coverage probabilities, but do degrade in accuracy with increasing collinearity. It is not completely unexpected that the Bayesian SVCP model accommodates collinearity given previous work that shows that ridge regression coefficient estimates may be viewed as Bayesian regression coefficient estimates under specific vague priors. Moreover, we illustrate that the posterior mean of the Bayesian SVCP model regression coefficients are of a ridge regression solution form.

In an example with bladder cancer mortality rates in the United States, we demonstrate that the SVCP model produces fewer regression coefficients with signs that are counter to both the least-squares regression coefficients and intuition than does a frequentist regression model called geographically weighted regression (GWR) that also has spatially varying coefficients. In addition, the SVCP model regression coefficients for different types of variables appear to be less spatially dependent and have less variance than those from GWR. Apparently, the shrinkage features of the Bayesian SVCP model lead to more robust and fewer misleading marginal inferences.

References

Banerjee S, Johnson GA (2005) Coregionalized single- and multi-resolution spatially-varying growth curve modelling with application to weed growth. UMN Biostat Tech Report.

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. John Wiley & Sons: West Sussex.

Gelfand AE, Kim H, Sirmans CF, Banerjee S (2003) Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association* 98: 387 – 396.

Gelman A, Carlin JB, Stern HS, Rubin DB (2003) *Bayesian Data Analysis* (2nd ed.). Chapman & Hall: London.

Goldstein M (1976) Bayesian analysis of regression problems. *Biometrika* 63 (1): 51 – 58.

Hastie T, Tibshirani R, Friedman J (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag: New York.

Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for non-orthogonal problems. Technometrics 12: 55 – 67.

Lindley DV, Smith AFM (1972) Bayes estimates for the linear model. Conference Proceedings, Royal Statistical Society.

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.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W (1996) *Applied Linear Regression Models*. Irwin: Chicago.

Wheeler D, Tiefelsdorf M (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems* 7: 161 – 187.

Weight	X Corr	Mean CP (β1)	Mean CP (β2)	C_{12}	C^{s}_{12}	RMSE(β)	RMSE(y)
0.0	0.000	0.90	0.91	0.138	0.064	0.508	0.894
0.1	0.127	0.90	0.89	0.199	0.093	0.514	0.907
0.3	0.442	0.89	0.89	0.377	0.147	0.536	0.907
0.5	0.755	0.90	0.89	0.597	0.212	0.568	0.899
0.7	0.937	0.91	0.91	0.777	0.228	0.651	0.904
0.9	0.995	0.93	0.93	0.835	0.224	1.235	0.892

Table 1. Results of the simulation study. The columns listed in order are the correlation weight, explanatory variable correlation, average coverage probabilities for each explanatory variable coefficient, the mean overall correlation between the variable coefficients, the mean local coefficient correlation at each location, the average RMSE of the coefficients, and RMSE of the response.

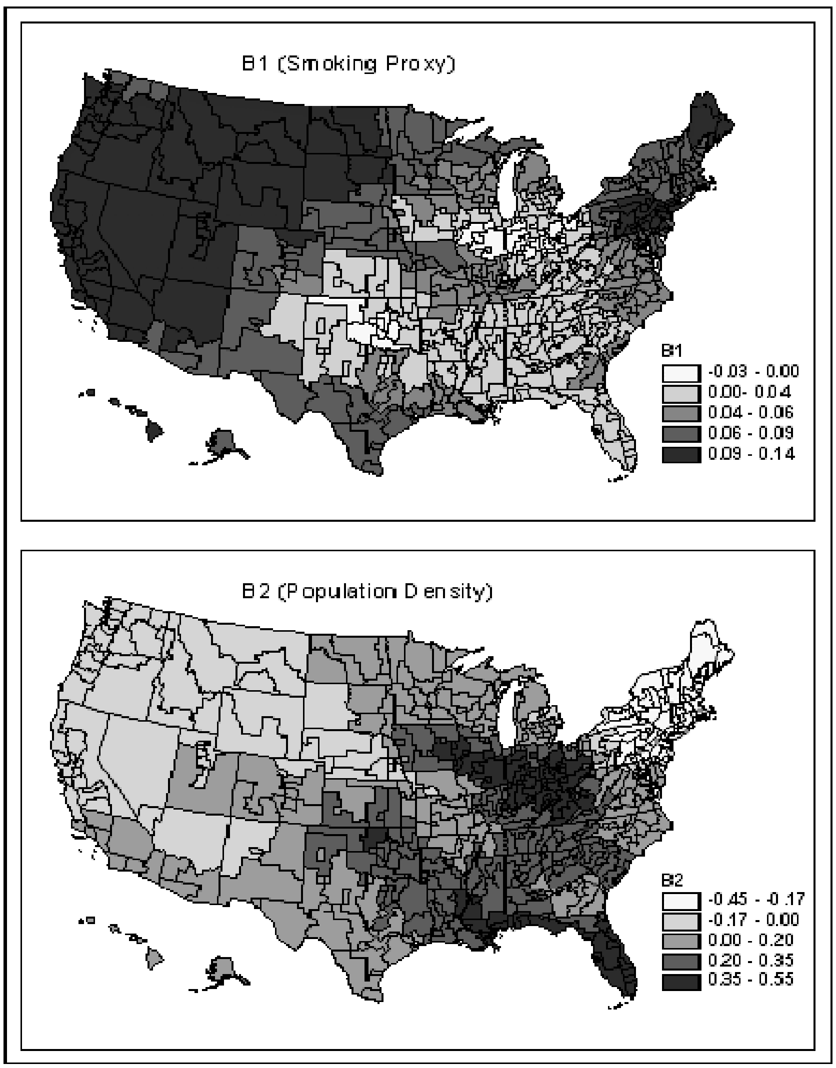

Figure 1. Estimated coefficients for smoking proxy (top) and population density (bottom) for the GWR model

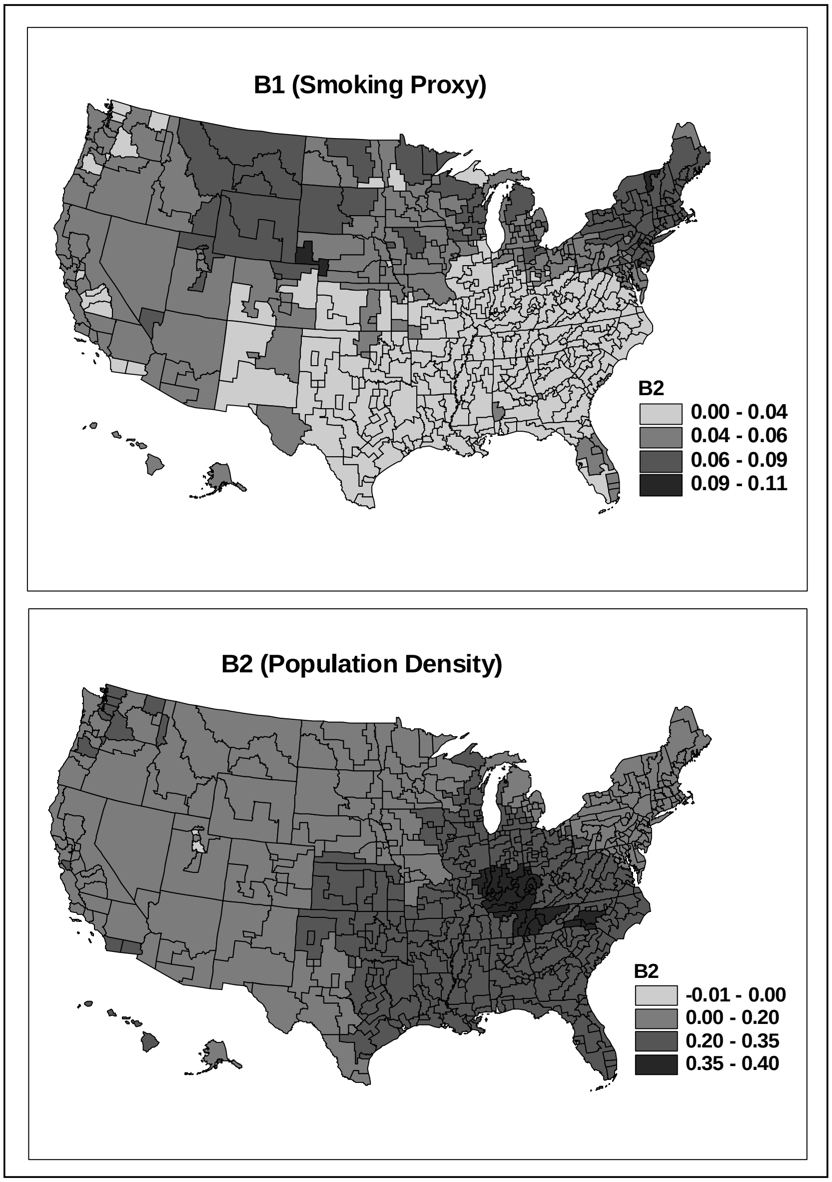

Figure 2. Estimated coefficients for smoking proxy (top) and population density (bottom) for the SVCP model