

Modeling Crop Residue Yield Potential using Geographically Weighted Regression

Abstract

Ordinary Least Squares (OLS) regression often doesn't accurately model data with spatial nonstationarity. In this research work, a relatively new approach, Geographically Weighted Regression (GWR), is used for the modeling of such data. The model performance of OLS and GWR were compared in terms of coefficient of determination (R^2) and corrected Akaike Information Criterion (AICc). Moran's I and Geary's C were used to test the spatial autocorrelation of OLS and GWR residuals. The explanatory power of the models is assessed by approximate likelihood ratio test. Furthermore, the test of spatial heterogeneity of the GWR parameters was conducted by using data sets with small and large samples. The comparative study of R^2 and AICc between the models showed that all the GWR models performed better than the analogous OLS models. Test of spatial autocorrelation of residuals revealed that the OLS residuals had higher degrees of spatial autocorrelation than the GWR residuals indicating that GWR mitigates the spatial autocorrelation of residuals. Results of the approximate likelihood ratio test showed that GWR models performed better than the OLS models suggesting that the OLS relationship was not constant across the space of interest. More importantly, it was demonstrated that the data set would have to be large enough for the individual parameters of GWR models to be spatially heterogeneous. A real world data set with crop residue yield potential, temperature, and precipitation of the North-Central region (Illinois, Indiana, Iowa, Minnesota, Montana, Nebraska, North Dakota, South Dakota, Wisconsin, and Wyoming) of the USA was used as an application.

Keywords Geographically Weighted Regression, Spatial autocorrelation, spatial heterogeneity, residuals, crop residue yield potential

1. Introduction

A traditional modeling technique used in geographical analysis is based on Ordinary Least Squares (OLS). In this technique, it is assumed that patterns in the data are spatially constant, and therefore, the parameter estimates are the same for the whole study area. This technique does not account for location in the analysis of the relationship between the variables. Therefore, such parameter estimates can be considered as global statistics (Fotheringham et al. 2002). A simple form of the OLS model (in matrix form) is given by

$$Y = X\beta + \varepsilon \quad (1)$$

where Y is the $n \times 1$ vector of the dependent variable, X is the design matrix of independent (explanatory) variables, which includes a column of 1s for the intercept, β is the vector of regression coefficients (global parameter estimates for the independent variables) and ε is a random vector, where $\varepsilon \sim N(0, \sigma^2 I)$. The maximum likelihood estimate of β under such model is given by

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (2)$$

However, in practice, there may be situations when the pattern of the data is not spatially constant. This is referred to as spatial nonstationarity (Brunsdon et al. 1998). Under spatial nonstationarity, the measurement of a relationship depends partly on the location where the measurement is taken. Such a relationship which is nonstationary over space will not be well addressed by OLS regression. The violation of the assumptions of independence of residuals in spatial data creates additional problems, such as spatial autocorrelation. Spatial autocorrelation is a phenomenon in which the value of a given variable at a location is related to the values of the same variable in a nearby location (Zhang et al. 2004).

In an attempt to explore and explain the spatially varying relationships by allowing the model parameters to vary over space and thus to attempt to overcome some restrictive assumptions of OLS regression, the concept of Geographically Weighted Regression (GWR) was developed (Fotheringham et al. 2002). GWR is an extension of OLS regression in which the parameters are allowed to vary spatially. As such, spatial relationships between the variables can be examined and patterns can be identified. GWR is often referred to be a local model because it provides estimates to local statistics and hence it is more appropriate when the relationships vary spatially (Fotheringham et al. 2002). Brunsdon et al. 1998 define a GWR model as

$$Y_i = \sum_{j=1} X_{ij} \beta_j(p_i) + \varepsilon_i \quad (3)$$

where Y is again a $n \times 1$ vector of dependent variables, p_i is the coordinates for the observation i , and ε_i is the random error term for the i th observation. These $\beta_j(p_i)$ contain the local parameters to be estimated. The concept here is that these expressions would be substituted into the original model and the resulting expanded model (which may now be nonlinear) is calibrated (Brunsdon et al. 1998).

Tulbure et al. 2011 and Brown et al. 2012 have shown that GWR performs better than OLS regression model. Páez et al. 2011 mention that GWR is not recommended in situations with small sample sizes ($n \approx 160$ in their experiments). To the best of our knowledge, no work has been done in the literature to test the spatial heterogeneity of individual parameters of GWR model when applied to data of small and large samples. Therefore, the primary purpose of this research work was to test the spatial heterogeneity of individual parameters of GWR model

when applied to spatially varying data of small and large samples. Similarly, the secondary purpose of this research work was to test the performance of GWR and OLS regression models when applied to these data sets. In addition to comparing the performances of these two models, our third objective was to assess the influence of temperature and precipitation on crop residue yield potential and to assess the degree of spatial autocorrelation of OLS and GWR residuals using the outputs generated.

In this study, the spatial dependence of GWR residuals and OLS residuals were compared. The coefficient of determination (Adjusted R^2) and the Akaike Information Criterion (AIC and AICc) were used to compare the models. The widely accepted rule of thumb (Charlton et al. 2009) that a decrease of 3 in AIC is regarded as an improvement was also used. In addition to comparing the models with different independent variables, AIC_C can also be used to compare OLS and GWR models (as long as the response variable is the same in the models) and also in the software that determines the optimal value for the bandwidth; the bandwidth with lowest AIC_C value will be used to estimate the parameters (Charlton et al. 2009). An approximate likelihood ratio test based on the F test was used to check for significant improvement of GWR over OLS. A test of spatial heterogeneity of individual parameters was also used.

In Section 2 the weighting functions and kernel bandwidths for GWR are discussed. Section 3 presents a comparative study of GWR and OLS models. In addition, the test of spatial heterogeneity of individual parameters of GWR and the outputs from GWR are discussed. In section 4 the application of GWR and OLS regression models using crop residue yield potential data of South Dakota and the ten states of the North-Central region (Illinois, Indiana, Iowa, Minnesota, Montana, Nebraska, North Dakota, South Dakota, Wisconsin, and Wyoming) of the USA for the years 1970, 1980, and 2008 are discussed which is followed by results and discussion. The study is completed with a summary of the results in Section 5.

2. Weighting Functions and Kernel Bandwidths

Brunsdon et al. 1998 and Fotheringham et al. 2002 provide a very sophisticated explanation of the theoretical development of GWR. The basic methodology behind GWR is that a weighted distance decay function w_{ij} is used for model calibration. One of the methods for such calibration is to take a point p_i and to fit a weighted OLS regression with observations within the circle weighted as 1 and zero otherwise. That is, the weight w_{ij} for a given p_i assigned to observation j would be

$$w_{ij} = 1 \text{ if } d_{ij} < r, \text{ and } 0 \text{ otherwise} \quad (4)$$

where d_{ij} is the distance between locations of observations i and j and r is the radius of the circle.

A variety of weighting functions are available (see, for example, Brunsdon et al. 1998). In all of these, it is assumed that the weight decreases with an increase in distance between the observations. One of the weighting functions could be an exponential function given by

$$w_{ij} = \exp\left(-\frac{d_{ij}}{r}\right) \quad (5)$$

where d_{ij} is, as before, the distance between locations of observations i and j , and r is the kernel bandwidth.

Another weighting function commonly used in practice is to calculate the weight by using a Gaussian distance decay function (also referred to as the Continuous Weighting Function) given by

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2r^2}\right) \quad (6)$$

where the values of w_{ij} decay gradually with distance to the extent that when $d_{ij} = 3r$, the weight virtually decreases to zero (Brunsdon et al. 2001). Results from GWR are not very sensitive to the choice of the weighting function, provided that it is smooth and follows a distance decay property (Brunsdon et al. 1998). In this paper, the weights w_{ij} are calculated at each calibration location using the Gaussian distance decay function given by (6).

To fit a GWR model, at first the kernel bandwidth parameter r , is estimated from the data. At present, there are three methods for this: direct assignment of the bandwidth of number of nearest neighbors (McMillen 1996), cross validation process (an iterative process that finds the kernel bandwidth that minimizes the prediction error of all the response variables using a subset of data for prediction) (Brunsdon et al. 1996; Farber and Páez 2007), and a corrected Akaike Information Criterion (AIC_C) (Fotheringham et al. 2002). As of now, the most commonly used approach is cross validation, and this method is used in our analysis. The vector of estimated regression coefficients is given by

$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i Y \quad (7)$$

where W_i is a square matrix of weights relative to the position i and is given by

128

$$W_i = \begin{bmatrix} w_1 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & \dots & \dots \\ 0 & 0 & \dots & w_N \end{bmatrix}$$

129 and contains the geographical weights in its main diagonal and 0 in its off diagonal elements, N
130 is the number of observations, $X^T W_i X$ is the geographically weighted variance covariance
131 matrix (Charlton et al. 2009), and Y is the vector of values of the response variables.

132 At present, GWR can be run using GWR3.0 or R. In both of these, fixed and adaptive bandwidth
133 kernels can be chosen. The former computes a bandwidth which is held constant over space
134 while the latter adapts the bandwidth distance according to the density of the data (the denser
135 the data, the smaller the bandwidth and vice versa) (Fotheringham et al. 2002; Tu and Xia 2008;
136 Brown et al. 2012). In this study, an adaptive kernel bandwidth is chosen since sample densities
137 varied over the study area. The data analyses were done using the spgwr package (Bivand and
138 Yu 2013) for GWR in the statistical software package R, version 3.1 (R Development core team
139 2013).

140 3. Model Comparisons

141 The comparison between OLS and GWR models was performed by comparing their R^2 and AIC
142 /AICc using a small data set (crop residue data of 66 counties of South Dakota, U.S. for the
143 years 1970, 1980, and 2008) to those with a larger data set (crop residue yield potential data of
144 743 counties of ten states of the North-Central region (Illinois, Indiana, Iowa, Minnesota,
145 Montana, Nebraska, North Dakota, South Dakota, Wisconsin, and Wyoming) of the USA for the
146 years 1970, 1980, and 2008). The conventional interpretation of R^2 is that larger the R^2 ,
147 greater the explanatory power of the model. While comparing two models using AIC_C, the
148 model with lower AIC_C value is considered to be a better model. More precisely, if the AIC_C
149 values for two models differ by at least 3, then the model with the smaller AIC_C will be
150 considered to be the better model. Furthermore, the results of an approximate likelihood ratio
151 test (also called goodness of fit test) based on the F test (Fotheringham et al. 2002) was used to
152 assess the model improvement of GWR over OLS models. A statistically significant result from
153 the F test (at a given level of significance) indicates that the GWR model performs better than
154 the analogous OLS model.

155 Residual Analyses were conducted for both of the models, which include histograms of the OLS
156 and GWR residuals. Global Moran's I and Geary's C statistics (which are special cases of the
157 general cross product statistic applied to continuous data) were calculated to test the presence
158 of spatial autocorrelation of the residuals. These statistics are particularly useful with spatial

data whose covariance structures are defined by neighborhoods, a common situation for lattice data. Similar to the coefficient of correlation, the values of global Moran's I also range from -1 to 1 where the value of 1 and -1 indicate perfect positive spatial autocorrelation and perfect negative spatial autocorrelation, respectively. A global Moran's I with a value of 0 indicates no spatial autocorrelation. On the other hand, the values of Geary's C range from 0 to 2. It should be noted that Geary's C is interpreted much differently than Moran's I statistic. First, Geary's C has a mean of 1 under the null hypothesis of no spatial autocorrelation, and can never be negative. Second, values of Geary's C between 0 and 1 indicate positive autocorrelation, whereas values of C greater than 1 indicate negative spatial autocorrelation. The calculations of Moran's I and Geary's C were performed using the *spdep* package (Bivand 2013) in the statistical software package R, version 3.1 (R Development core team 2013). Spatial plots were used to detect the pattern of the OLS and GWR residuals.

3.1 Test of spatial heterogeneity of individual parameters of GWR

As explained by Charlton et al. 2006; Fotheringham et al. 2002; and Tulbure et al. 2011, individual parameter estimates of the GWR model are considered spatially heterogeneous if the inter-quartile range (difference between the third quartile and the first quartile) of GWR coefficients is greater than the range of values at ± 1 standard deviations of the respective global estimates i.e., $Q_3 - Q_1$ of GWR coefficients is greater than the range of $(b_1 - 1 S.D., b_1 + 1 S.D.)$ of OLS estimates which is simply twice the standard deviation ($S.D.$) of the OLS estimates.

3.2 Outputs from GWR

The spatial variation in parameter estimates, also considered as the main output from GWR, were used to see the variations in relationships revealed by these coefficients. In addition, histograms of the parameter estimates were used to detect the spatial variation revealed by these coefficients. As recommended by Wheeler and Tiefelsdorf 2005, scatterplots of estimated GWR coefficients for the pairs of predictor variables of the regression model were used to visualize the nature of dependence in the estimated coefficients.

4. Application

4.1 Data

The data set was obtained from the USDA National Agricultural Statistics Service (USDA-NASS, 2009). North-Central region of the USA was selected for this observational study which includes

the states of Illinois, Indiana, Iowa, Minnesota, Montana, Nebraska, North Dakota, South Dakota, Wisconsin, and Wyoming (as defined by the Sun Grant Research Initiative). This selected region is located between 37.0 and 49.0 degrees north latitude and 84.8 and 116.0 degrees west longitude (Chintala et al. 2013). The data set consists of the variables: County names, FIPS ID of the county, crop residue yield potential, temperature, and precipitation for the years 1970-2008. The county level dry crop residue yield potential (Mg ha^{-1}) was calculated using crop yield data collected from the USDA-NASS, 2009; temperature was the average temperature ($^{\circ}\text{C}$) during crop growing seasons (April-October), and precipitation was the annual precipitation (mm) derived from the monthly gridded PRISM weather data. The county level dry crop residue yield potential was considered as a dependent (response) variable and the climate variables temperature and precipitation were considered as independent (predictor) variables. For this analysis, the subset of the data for the three years 2008, 1980, and 1970 was considered. The response variable crop residue yield potential was closely normal and several transformations were attempted to normalize it. However, none of the transformations perfectly normalized it, and so it was kept as is for further analysis. Transformations were not attempted for the predictor variables. The data set had missing values of these variables for 56, 3, and 5 counties for 2008, 1980, and 1970 data respectively. Some examples of the counties with the missing data were Corson, Custer, Lawrence, Shannon, and Stanley. An imputation technique was used to estimate the missing values of crop residue yield potential, temperature, and precipitation. Thus, we had a complete data set for all 743 counties of the ten states of the North-Central region of the USA.

The original version of the data set with crop residue yield potential, temperature and precipitation for the years 1970 to 2008, was already been analyzed by averaging over time and then Conditional Autoregressive and Simultaneous Autoregressive models (Schabenberger and Gotway 2004) were fit (Chintala et al., in press). To the best of our knowledge, this dataset has not been analyzed using GWR yet.

4.2 Results and Discussion

OLS and GWR regression models were fit for the crop residue yield potential as a function of two climate variables (temperature and precipitation) for small as well as large data sets (66 counties of South Dakota and 743 counties of ten states of the North-Central region of the USA) for the years 1970, 1980, and 2008. In all the cases, the values of Moran's I and Geary's C for residuals (Table 1) indicate that there is positive autocorrelation of OLS residuals. For instance, for the year 1970, the value of Moran's I for OLS residuals was 0.33 while that of Geary's C was 0.63 (both agreeing on positive autocorrelation of residuals). On the other hand, the values of Moran's I and Geary's C (Table 1) indicated that there was a negative autocorrelation of GWR residuals, for example, the value of Moran's I for GWR residuals was -0.12 and Geary's C was

1.11 (both agreeing on small negative autocorrelation of residuals). It should also be noted that in all the years, the degree of autocorrelation of OLS residuals was higher than that of GWR residuals indicating that GWR mitigated the autocorrelation of residuals. This was also supported by the spatial plot of OLS and GWR residuals (Figures 1 and 2). In addition, there was more variability of OLS residuals than GWR residuals, for instance, in 2008 data, the OLS residuals ranged from -5.286 to 7.1 while the GWR residuals ranged from -4.23 to 3.35. The variances of OLS and GWR were 3.92 and 0.84 respectively. These numbers helped us quantify the visible relationships between the two map patterns illustrated in Figures 1 and 2. The residuals under all the OLS models were non normal, while the residuals under all the GWR models were almost normal (based on histograms, Figure 3). Significant improvements in the fit of both the OLS and GWR models were observed for all models on comparing R^2 and AIC_C . In all cases, the values of R^2 under GWR models were higher than those of the corresponding OLS models (Table 1) indicating the higher explanatory power of the GWR models over the analogous OLS models. In addition, all the GWR models had smaller AIC_C values than the analogous OLS models (Table 1).

Table 1: Comparison of OLS and GWR

DATA	Year	Moran's I		Geary's C		Adjusted R^{2*}		AIC_C	
		OLS	GWR	OLS	GWR	OLS	GWR	OLS	GWR
66 counties	1970	0.33	-0.12	0.63	1.11	0.08	0.74	57.00	37.63
	1980	0.54	0.05	0.45	0.99	0.45	0.94	132.53	111.41
	2008	0.58	-0.01	0.4	0.95	0.01	0.74	287.48	237.5
743 counties	1970	0.62	-0.07	0.39	1.06	0.22	0.84	2645.44	1955.68
	1980	0.47	-0.05	0.53	1.06	0.46	0.87	2606.92	2047.67
	2008	0.43	0.21	0.56	0.96	0.39	0.87	3131.34	2412.15

*Note: * The adjusted R^2 for GWR models are the mean values of the corresponding R^2 values.*

The GWR models in all cases performed better than the corresponding OLS models, according to the approximate likelihood ratio test (Table 2) with extremely small p -values in each case suggesting that the ordinary least squares relationship between crop residue yield potential, temperature, and precipitation was not constant across the study area. Table 2 also shows the significant improvements in explained variance (reduced sum of squared errors) by the GWR models over the analogous OLS models.

Table 2: Goodness of fit test for improvement in model fit of GWR over OLS

DATA	Year	Source			<i>F</i> -Statistic	<i>P</i> -value
		OLS Residuals	GWR Residuals	GWR Improvement		
66 Counties	1970	8.12	2.40	5.73	3.54	6.7e-06
	1980	25.50	3.09	22.40	5.11	7.1e-08
	2008	266.76	72.65	194.11	7.14	1.0e-11
743 Counties	1970	1514.04	317.75	1196.29	7.53	< 2.2e-16
	1980	1437.55	336.52	1101.03	5.88	< 2.2e-16
	2008	2911.72	626.50	2285.22	8.27	< 2.2e-16

From Table 3, it is seen that for small data set with 66 counties of South Dakota for the years 1970 and 1980, the inter-quartile range (IQR) of GWR coefficients is greater than twice the standard deviations of the respective global estimates. However, for the small data set with 66 counties of South Dakota for the year 2008, the inter-quartile ranges of the intercept and precipitation of GWR coefficients are less than twice the standard deviations of their respective global estimates. On the other hand, for the large data set with 743 counties of ten states of the North-Central region of the USA for all the years, the inter-quartile range of all the GWR coefficients are greater than twice the standard deviations of their respective global estimates. This means that for the 2008 model, the parameters for intercept and precipitation for the small data set are not spatially heterogeneous. However, all the parameters for the large data set for all the models are spatially heterogeneous. Hence, a caution need to be exercised using GWR for the purpose of assessing spatial heterogeneity of parameters when analyzing small data sets.

Table 3: Test of spatial heterogeneity of individual parameters

DATA	Year	Variable	S.D. (Global)	2S.D. (Global)	IQR (GWR)	Results
66 Counties	1970	Intercept	0.74	0.148	2.134	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Temp	0.041	0.082	0.136	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.0011	0.0022	0.004	$Q_3 - Q_1 > 2 \text{ S.D.}$
	1980	Intercept	1.44	2.88	9.394	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Temp	0.08	0.16	0.649	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.0013	0.0026	0.006	$Q_3 - Q_1 > 2 \text{ S.D.}$
	2008	Intercept	5.6	11.2	8.31	$Q_3 - Q_1 < 2 \text{ S.D.}$
		Temp	0.335	0.670	0.691	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.007	0.014	0.011	$Q_3 - Q_1 < 2 \text{ S.D.}$
743 Counties	1970	Intercept	0.34	0.68	10.59	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Temp	0.027	0.054	0.67	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.0004	0.0008	0.006	$Q_3 - Q_1 > 2 \text{ S.D.}$
	1980	Intercept	0.36	0.72	11.58	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Temp	0.022	0.044	0.75	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.0003	0.0006	0.008	$Q_3 - Q_1 > 2 \text{ S.D.}$
	2008	Intercept	0.52	1.04	20.2	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Temp	0.04	0.08	1.43	$Q_3 - Q_1 > 2 \text{ S.D.}$
		Precpt	0.0005	0.001	0.012	$Q_3 - Q_1 > 2 \text{ S.D.}$

Note: Temp and Precpt stand for temperature and precipitation respectively.

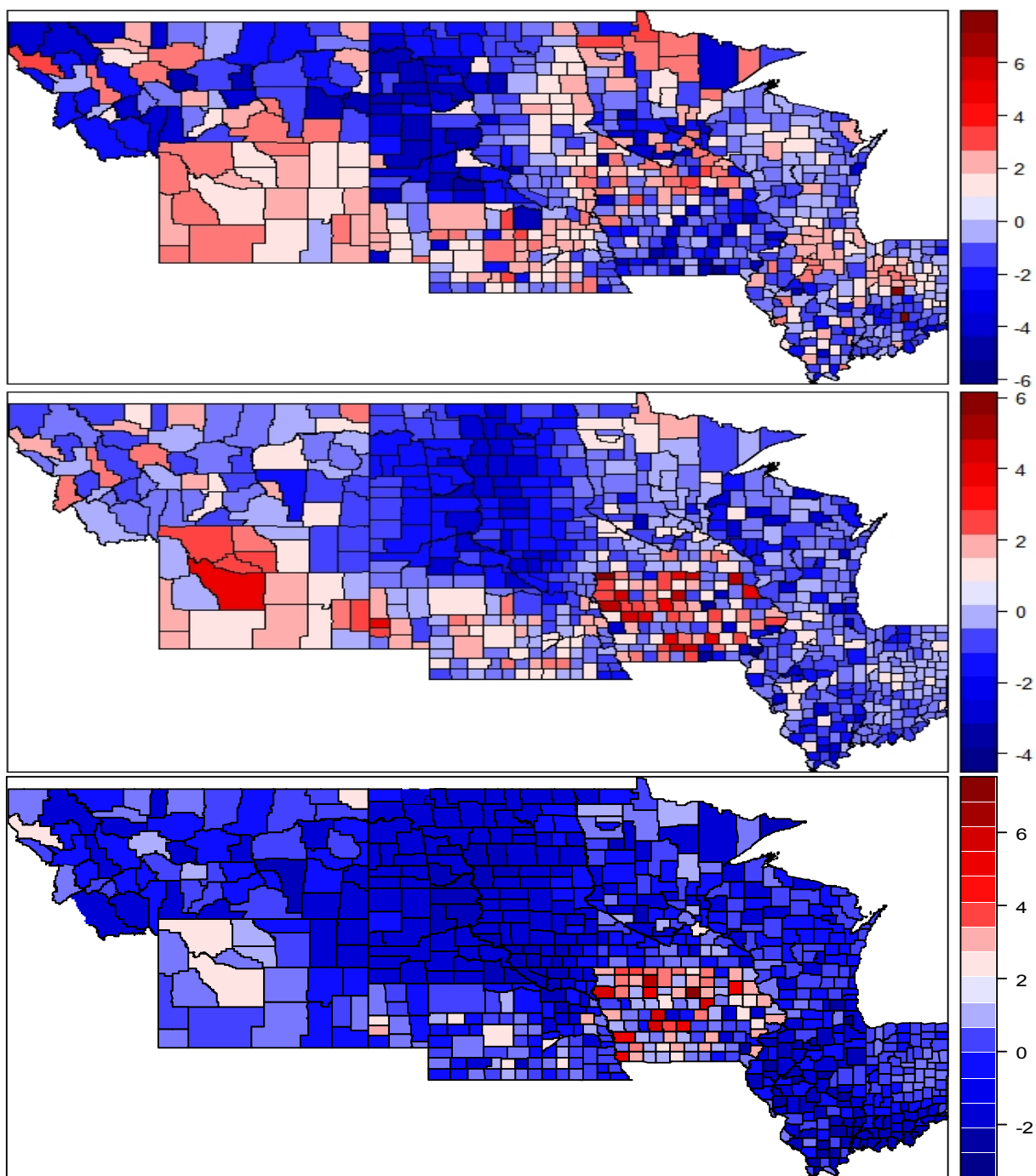


Fig. 1. Spatial distribution of OLS residuals for the ten mid western states of the US. *Top to bottom:* 2008, 1980, and 1970

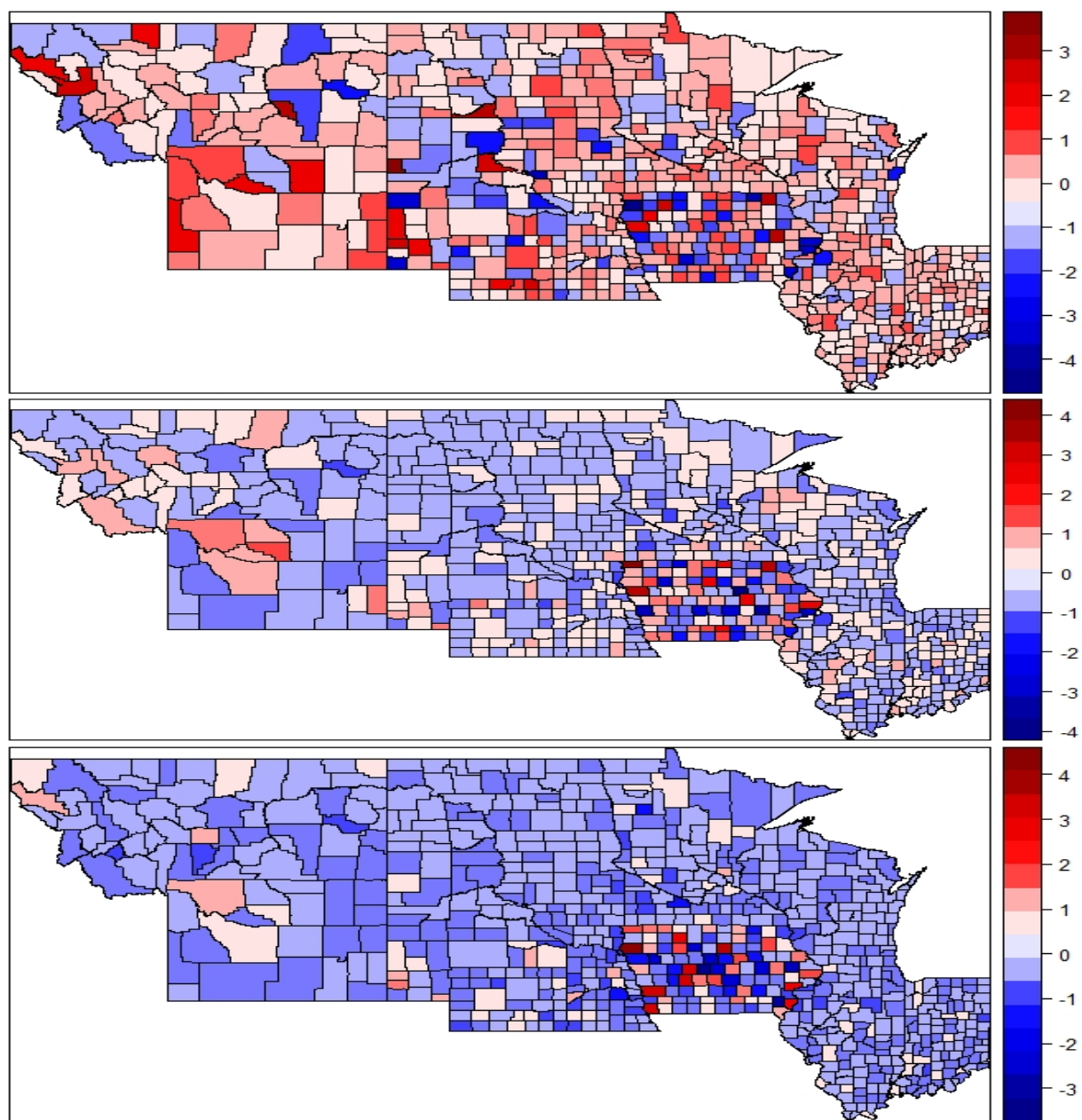


Fig. 2. Spatial distribution of GWR residuals for the ten mid western states of the US. *Top to bottom: 2008, 1980, and 1970*

The normality of OLS and GWR residuals for all the models was compared using histograms (Figure 3). It was found that the GWR residuals for all the models were more closely normally distributed than OLS residuals. It supported our previous conclusion that GWR had the ability to better model spatially varying data. In addition, the non-normality of OLS residuals implied that the assumptions of OLS regression models were not satisfied, raising a question about the validity of these models. When the assumptions of OLS models are violated, the results of such models could be misleading, and statistical

inferences provided by such models could be spurious. In addition, the strong autocorrelation of OLS residuals suggests that the OLS relationship between crop residue yield potential, temperature, and precipitation was not stationary. On the other hand, the lower degree of spatial autocorrelation of GWR residuals implied that the GWR models were more trustworthy. Hence, the non-normality and spatial autocorrelation of the OLS residuals and lower values of the coefficients of determination of OLS models support the assertion that GWR models the spatially varying data more efficiently than the analogous OLS models.

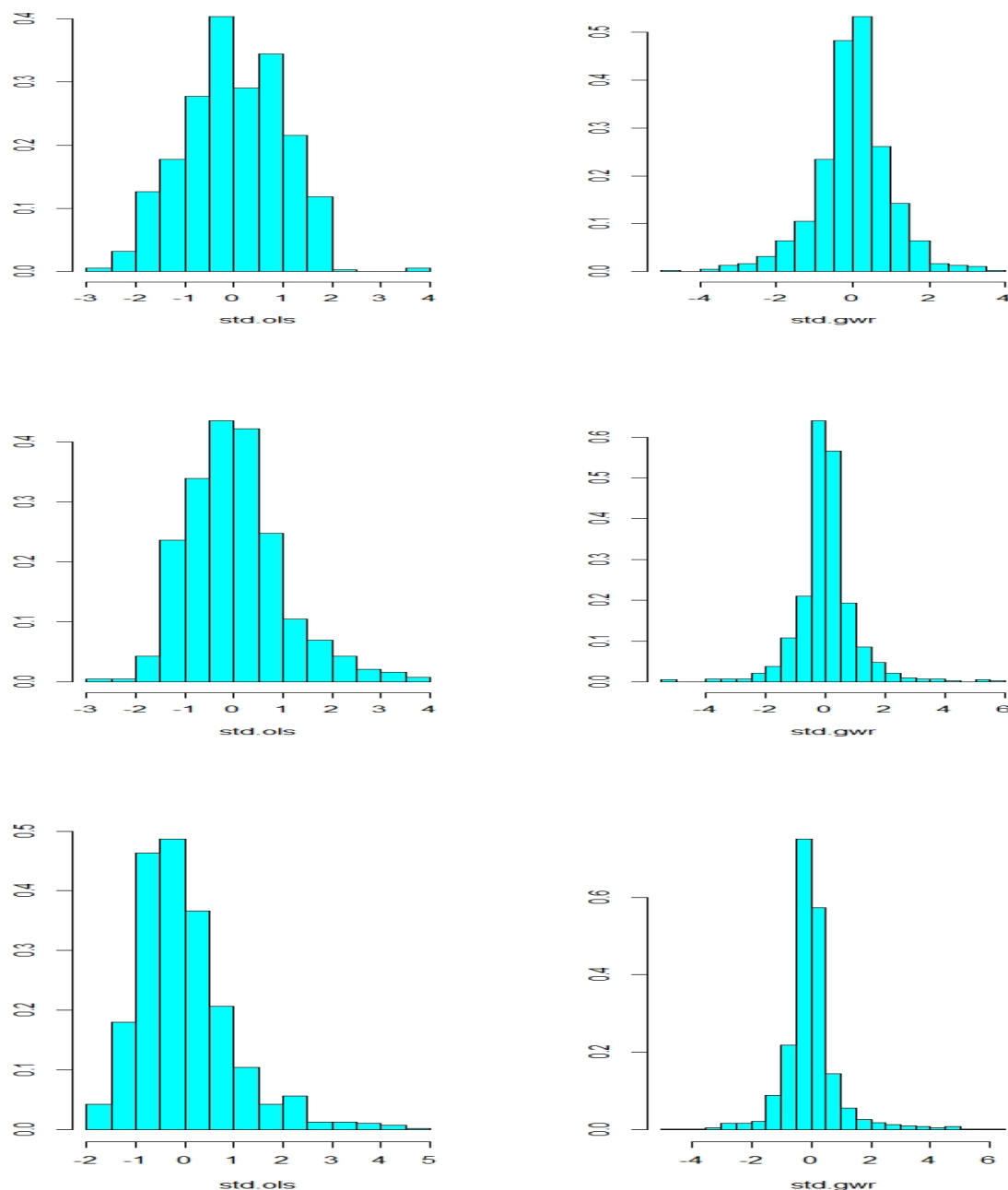
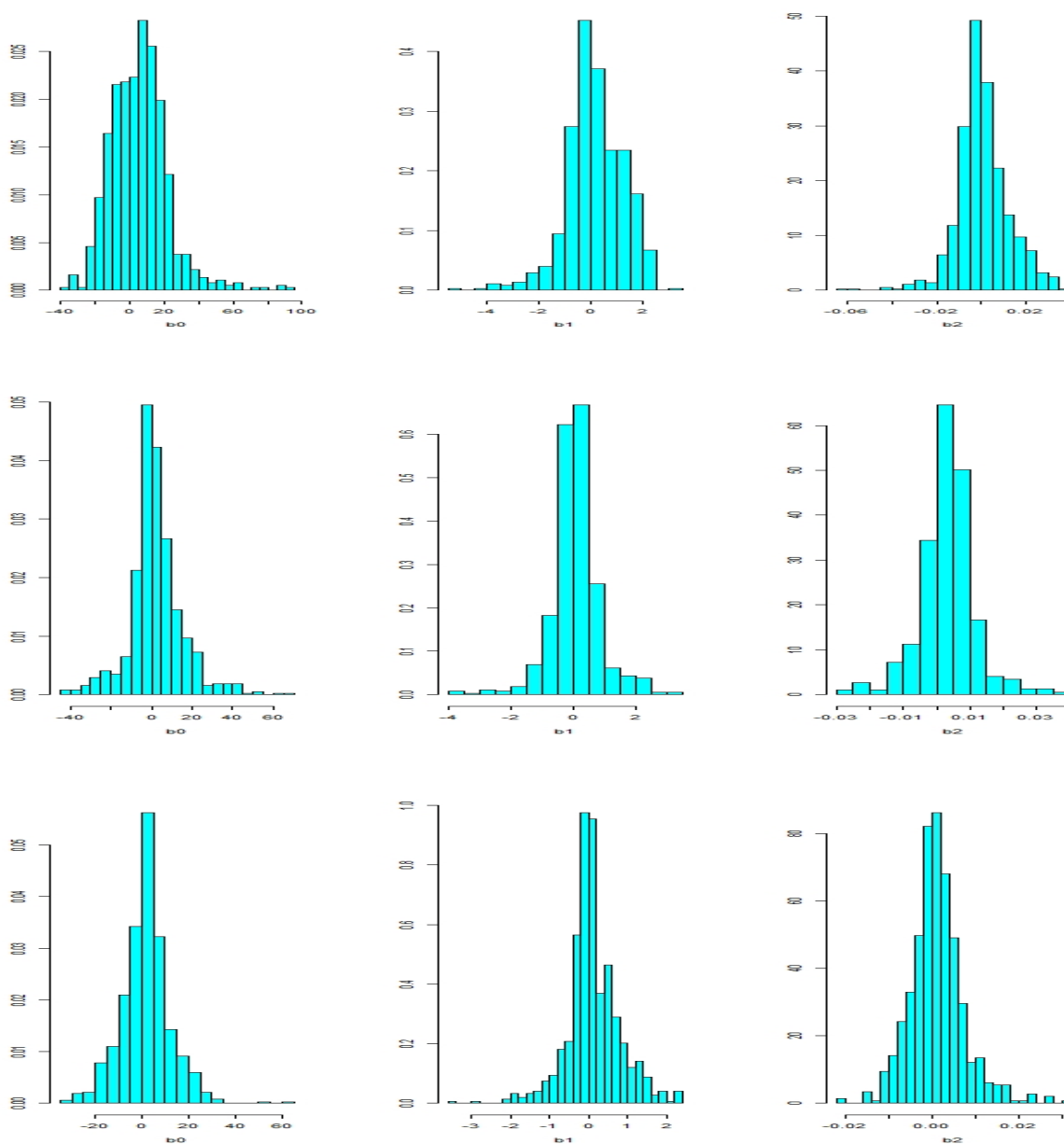


Fig. 3. Histogram of residuals. Top to bottom: 2008, 1980, and 1970; left to right: OLS, and GWR

311 The histograms of estimated GWR coefficients for the years 2008, 1980, and 1970 are given in
 312 Figure 4. These histograms allow the distributions of the coefficients to be easily compared for
 313 the three years. The distributions for intercept are right skewed, which indicate that a few
 314 counties have much larger intercepts than other counties. The distributions for temperature
 315 coefficients are left skewed, which indicate that a few counties have larger negative coefficients
 316 for temperature than other counties. The distributions for precipitation coefficients are left
 317 skewed for 2008, somewhat symmetrical for 1980, and right skewed for 1970. This indicates
 318 that the shape of the distribution of coefficients for precipitation has changed over time.



322 Fig. 4. Histogram of estimated GWR coefficients. *Top to bottom: 2008, 1980, and 1970; left to right:*
 323 *coefficients of intercept, temperature, and precipitation*

As recommended by Wheeler and Tiefelsdorf 2005, scatterplots of estimated GWR coefficients for the pairs of predictor variables of the regression model were used to visualize the nature of dependence in the estimated coefficients. Figure 5 shows the scatter plots for the pairs of the three regression terms:

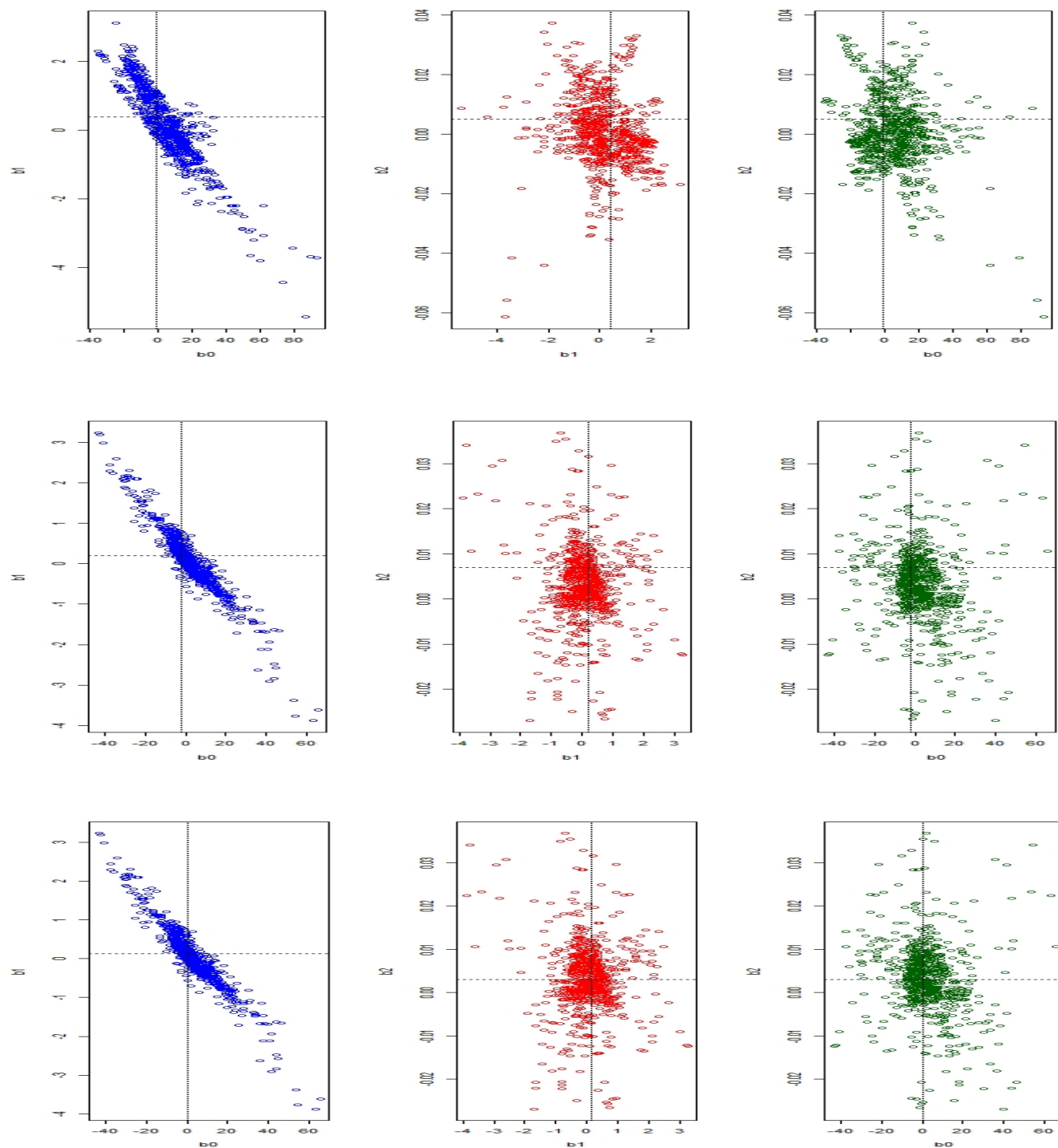


Fig. 5. Scatter plots of estimated GWR coefficients. *Top to bottom: 2008, 1980, and 1970; left to right: coefficients of intercept, temperature, and precipitation*

intercept, temperature, and precipitation. The vertical and the horizontal dashed reference lines denote the levels of analogous global parameter estimates (Table 4 has the parameter estimates for the OLS regression models for 2008, 1980, and 1970 models). The Pearson product moment correlation coefficients for the pairs of the regression terms are given in Table 5. These results strongly agree with the scatter plots from Figure 5. There is negative correlation between the coefficients of correlations (with the exception of the coefficients between intercept and precipitation for 1970). The degree of correlation between intercept and temperature is very strong. The correlations between the other pairs of regression terms are comparatively weaker. The scatter plots show that there is large variation around the OLS estimates for intercept; however, there is small variation around the OLS estimates for

Table 4: Parameter estimates for the OLS regression models

Year	Intercept	Temperature	Precipitation
2008	-1.16	0.393	0.005
1980	-1.888	0.195	0.007
1970	0.145	0.126	0.003

temperature and precipitation. Furthermore, the estimated GWR coefficients do not center on their analogous OLS estimates. The strong correlation between the coefficients of intercept and temperature could signal the presence of local collinearity in the GWR model which could lead to complications on the statistical inference drawn from these GWR coefficients.

Table 5: Correlation coefficients between the estimated GWR coefficients

Year	cor(b0,b1)	cor(b1,b2)	cor(b0,b2)
2008	-0.911	-0.144	-0.234
1980	-0.955	-0.097	-0.152
1970	-0.951	-0.333	0.0734

The spatial variation of the estimated GWR coefficients for the years 2008, 1980 and 1970 for intercept, temperature, and precipitation are shown in Figure 6, 7, and 8 respectively. The spatial variations in relationship shown by these distributions are interesting. A comparative study of Figures 6 and 7 shows a clear inverse map pattern between intercept and temperature (as one would expect, since the correlation coefficients between intercept and temperature from table 5 are large and negative in all the cases): in general, when the intercept parameter is high, the temperature parameter is low and vice versa. On the other hand, there is no clear pattern between the maps of intercept and precipitation and temperature and precipitation (Figures 7 and 8, as the coefficients of correlations between both the pairs are small in both the cases). Thus, the correlation coefficients given in Table 5 help us quantify the three map patterns illustrated in Figures 6, 7, and 8. For instance, the coefficients of correlations between temperature and precipitation in the years 2008 and 1980 are small. Due to this, there is not much difference in the patterns of the plots of temperature and precipitations for these years.

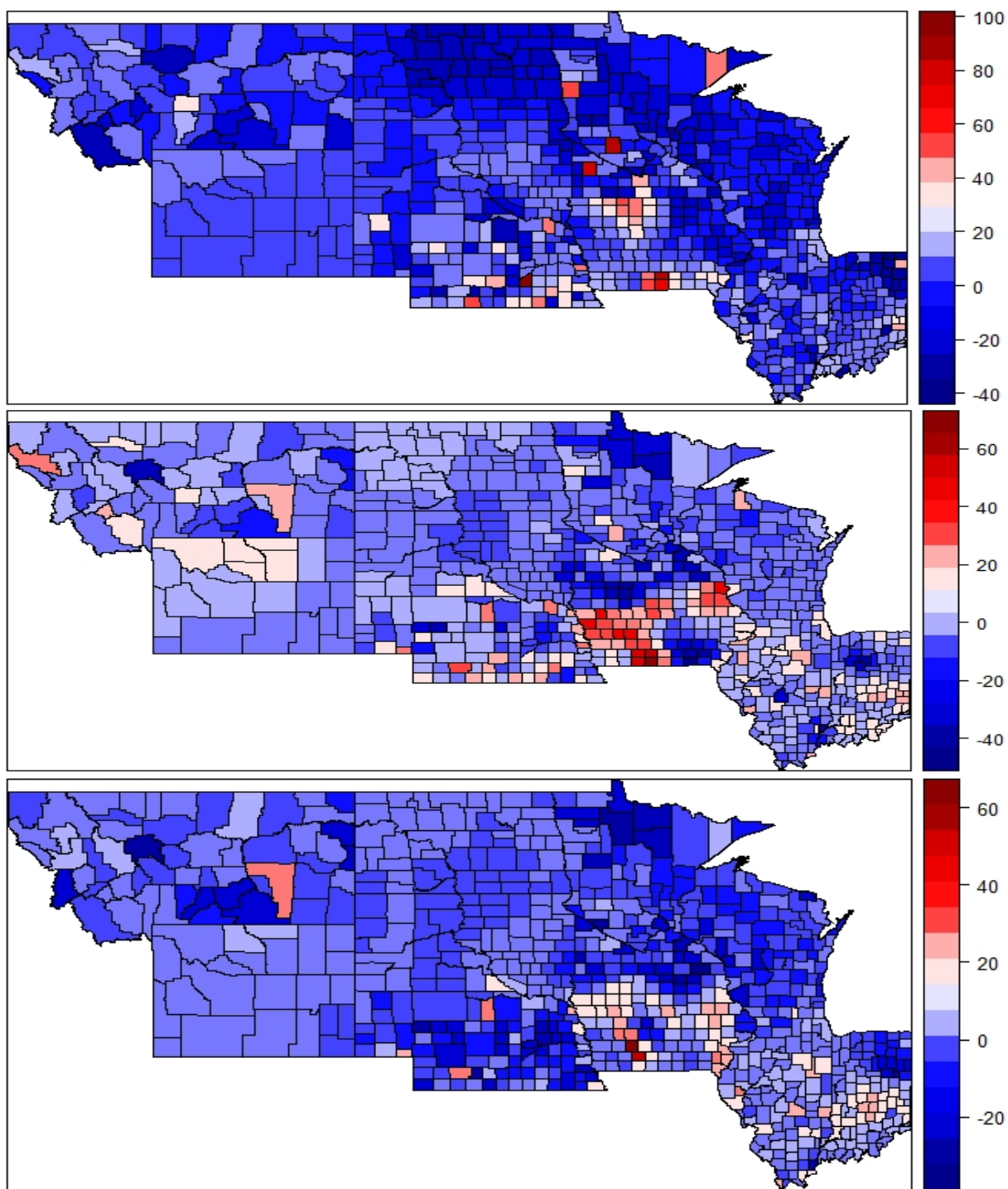


Fig. 6. Spatial distribution of estimated GWR coefficients for intercept for the ten mid western states of the US. *Top to bottom: 2008, 1980, and 1970*

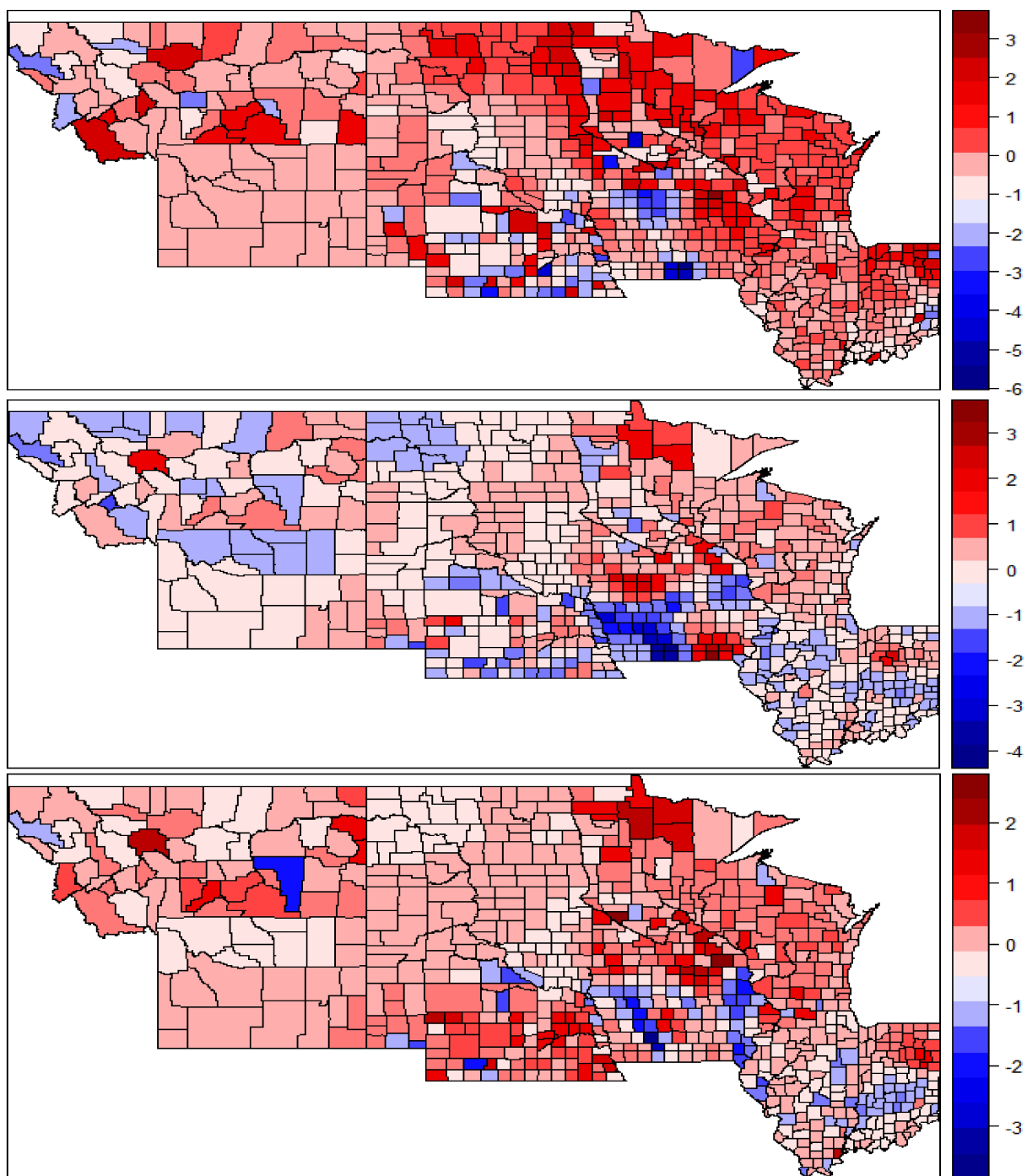


Fig. 7. Spatial distribution of estimated GWR coefficients for temperature for the ten mid western states of the US. *Top to bottom: 2008, 1980, and 1970*

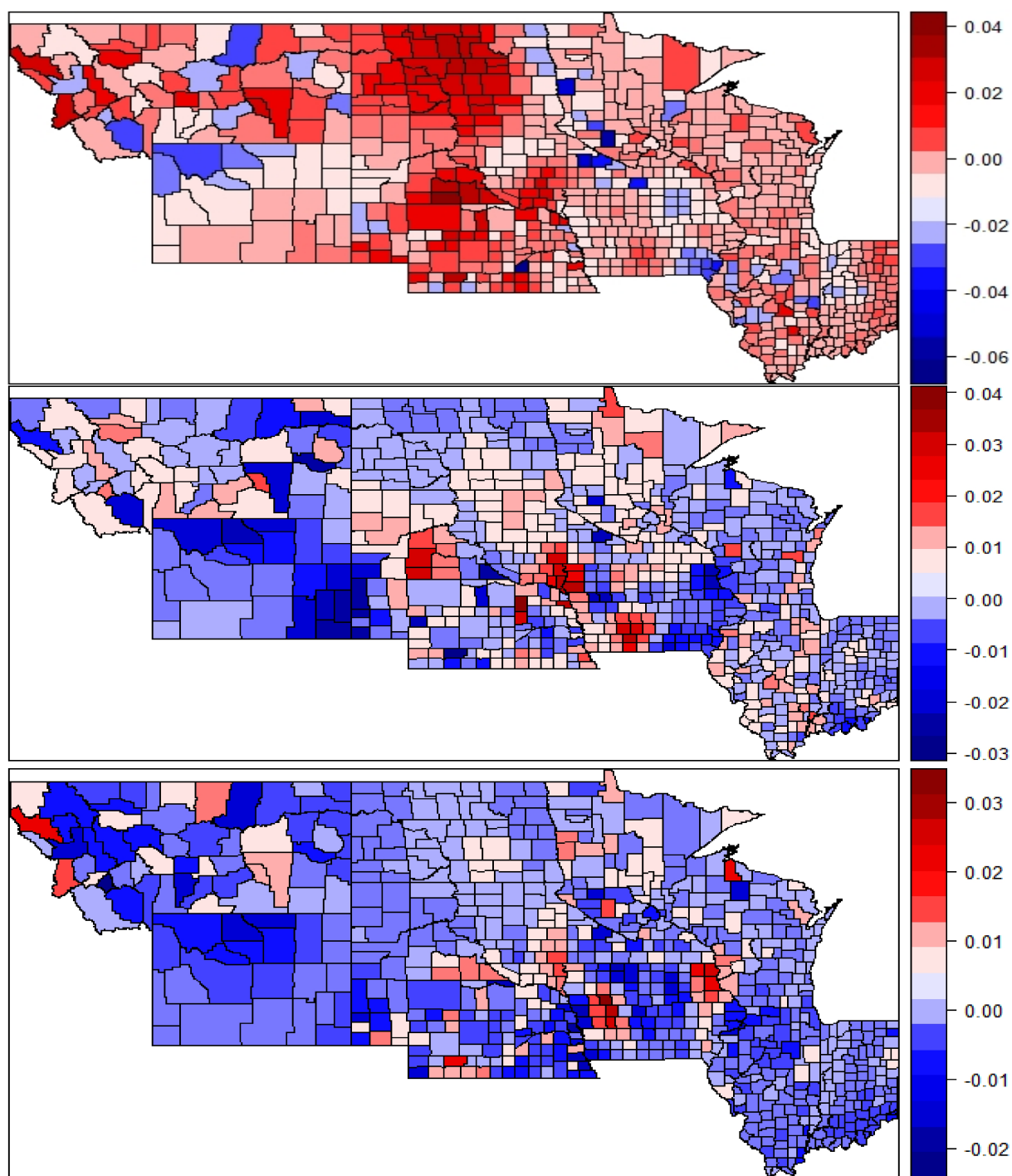


Fig. 8. Spatial distribution of estimated GWR coefficients for precipitation for the ten mid western states of the US. *Top to bottom: 2008, 1980, and 1970*

5. Summary

In this research work, OLS and GWR models were fit for crop residue yield potential of South Dakota, United States for the years 2008, 1980, and 1970 and the performance of these models were compared. The same analyses were repeated with a larger data set (crop residue yield potential data of ten states of the North-Central region of the USA). The approximate likelihood ratio test (also called the goodness of fit test) suggested that all the GWR models performed better than the analogous OLS regression models indicating that OLS relationship between crop residue yield potential, temperature, and precipitation was not constant across the study area. The coefficients of determination in all the GWR models were higher than in the analogous OLS models, indicating that GWR models had higher power at explaining spatially varying relationships than the analogous OLS models. All the GWR models had smaller AICC values than those of the analogous OLS models by at least 3 indicating that GWR models performed better and provided more parsimonious results than the analogous OLS models. In this study, the residuals of all the GWR models were approximately normally distributed while those of OLS models were not normally distributed. Moran's I and Geary's C statistics showed that the degree of autocorrelation of OLS residuals were higher than that of GWR residuals indicating that GWR mitigated the autocorrelation of residuals.

Out of the three cases of our analysis, two of the three parameters of GWR model (using small data set for 2008) were not spatially heterogeneous. This could be due to the small sample size, which increases the possibility of spurious correlations between local coefficients (Páez et al. 2011). In the other two cases, i.e. 1980 and 1970; all the parameters of the GWR model were spatially heterogeneous for small and large data sets. However, as suggested by Páez et al. 2011, the use of GWR should not be practiced for data with small sample size; and serious caution should be exercised when interpreting the results from such outputs.

In our case, there was strong (negative) correlation between the estimated GWR coefficients of intercept and temperature. However, the correlations between the other pairs of coefficients were weaker. The strong correlation between the GWR coefficients of intercept and temperature could be the result of local collinearity in the model. This could be investigated using maps of approximate local regression coefficient correlations (Wheeler and Tiefelsdorf 2005); local variance inflation factors (VIFs), variance-decomposition proportions and the associated condition indices (Belsley 1991; Wheeler 2007) as diagnostic tools for collinearity when estimating GWR coefficients.

Competing Interests

Authors have declared that no competing interests exist.

References

1. Belsley, D. (1991) Conditioning diagnostics: collinearity and weak data in regression. Wiley, New York
2. Bivand, R. (2013). spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-60. <http://CRAN.R-project.org/package=spdep>
3. Bivand, R. and Yu, D. (2013). spgwr: Geographically weighted regression. R package version 0.6-22. <http://CRAN.R-project.org/package=spgwr>
4. Brown, S., Versace, V. L., Laurenson, L., Ierodiaconou, D., Fawcett, J., & Salzman, S. (2012). Assessment of spatiotemporal varying relationships between rainfall, land cover and surface water area using geographically weighted regression. *Environmental Modeling & Assessment*, 17(3), 241-254.
5. Brunsdon, C., Fotheringham, S., & Charlton, M. (1998) Geographically weighted regression-modeling spatial non-stationarity. *The Statistician* 47:431-443
6. Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, 28(4), 281-298.
7. Brunsdon, C., McClatchey, J., & Unwin, D. J. (2001). Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. *International Journal of Climatology*, 21(4), 455-466.
8. Charlton, M., Fotheringham, S., & Brunsdon, C. (2009). Geographically weighted regression. White paper. National Centre for Geocomputation. National University of Ireland Maynooth.
9. Charlton, M., Fotheringham, S., & Brunsdon, C. (2006). Geographically Weighted Regression, ESRC National Center for Research Methods
10. Chintala, R., Wimberly, M. C., Djira, G. D., & Tulbure, M. G. (2013). Interannual variability of crop residue potential in the north central region of the United States. *Biomass and Bioenergy*, 49, 231-238.
11. Chintala, R., Djira, G., & Devkota, M. Modeling the Effect of Climate Variables on Crop Residue Potential for the North Central Region of the United States (in press)
12. Farber, S., & Páez, A. (2007). A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems*, 9(4), 371-396.
13. Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). Geographically weighted regression. New York: Wiley.
14. McMillen, D. P. (1996). One hundred fifty years of land values in Chicago: a nonparametric approach. *Journal of Urban Economics*, 40(1), 100-124.
15. Páez, A., Farber, S., & Wheeler, D. (2011). A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environment and Planning-Part A*, 43(12), 2992.
16. R Development Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

17. Schabenberger, O., & Gotway, C. A. (2004). Statistical methods for spatial data analysis (Vol. 64). Chapman and Hall/CRC.
18. Tu, J., & Xia, Z. G. (2008). Examining spatially varying relationships between land use and water quality using geographically weighted regression I: model design and evaluation. *Science of the total environment*, 407(1), 358-378.
19. Tulbure, M. G., Wimberly, M. C., Roy, D. P., & Henebry, G. M. (2011). Spatial and temporal heterogeneity of agricultural fires in the central United States in relation to land cover and land use. *Landscape Ecology*, 26(2), 211-224.
20. Wheeler, D. C. (2007). Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A*, 39(10), 2464-2481.
21. Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, 7(2), 161-187.
22. Zhang, L., Bi, H., Cheng, P., & Davis, C. J. (2004). Modeling spatial variation in tree diameter–height relationships. *Forest Ecology and Management*, 189(1), 317-329.