ELSEVIER

A close look at the spatial structure implied by the CAR and SAR models

Melanie M. Wall[*],[1]

Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

Modeling spatial interactions that arise in spatially referenced data is commonly done by incorporating the spatial dependence into the covariance structure either explicitly or implicitly via an autoregressive model. In the case of lattice (regional summary) data, two common autoregressive models used are the conditional autoregressive model (CAR) and the simultaneously autoregressive model (SAR). Both of these models produce spatial dependence in the covariance structure as a function of a neighbor matrix W and often a fixed unknown spatial correlation parameter. This paper examines in detail the correlation structures implied by these models as applied to an irregular lattice in an attempt to demonstrate their many counterintuitive or impractical results. A data example is used for illustration where US statewide average SAT verbal scores are modeled and examined for spatial structure using different spatial models.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Spatial interaction; Lattice data; Spatial autoregression

1. Introduction

In many settings, averages or counts over geographically defined regions are observed and ecological regression analysis is performed. When the location of the geographic regions are known it is common to assume that observations on regions near each other may tend to have similar score on the omitted variables in the regression causing the error terms to be spatially autocorrelated. Therefore some underlying spatial process is often included in the model. Besides improving inference of regression coefficients, the

[*] Tel.: +1-612-6252138; fax: +1-612-6260660.

E-mail address: melanie@biostat.umn.edu (M.M. Wall).

model for the spatial process should be able to provide a clear picture of the residual spatial pattern thus providing insight into what omitted variables there may be. This paper examines the different spatial structures implied by using different models for the underlying spatial process on an irregular lattice (e.g. the lattice formed by the states of the US).

There are two fundamentally different ways to model the spatial structure underlying lattice data (i.e. regional summary data). They are both special cases of the general spatial process $\{Z(s): s \in D\}$ and their difference lies in what is assumed about the indexing set D. One method is to treat the lattice data as if it was observed on a continuous indexing set (i.e. geostatistical data, Cressie, 1993) instead of a discrete indexing set. When this method of modeling is employed, most commonly the summary data for each region are assumed to have been observed at the center or centroid of the region and distances between centroids are used to develop the spatial covariance structure through a variogram function. One of the most commonly cited problems with this technique is the arbitrariness of assigning the summary for the whole region to the centroid. Even if some thoughtfully chosen point in the region was used as the location (e.g. population weighted centroid), another conceptual problem with modeling lattice data in this way is that it is really not possible for the observations being modeled (i.e. regional averages) to occur continuously in the plane as the model would allow. On the other hand, the good thing about modeling in this way is that the spatial covariance function is modeled directly and thus its structure is usually straightforward to understand.

The other way of modeling the spatial structure underlying lattice data, and the method investigated in detail in this paper, does not ignore the discrete index nature of lattice data. This is done by defining a neighborhood structure based on the shape of the lattice. Thus instead of measuring distance between centroids of regions, a system is used that defines regions to be neighbors based on, for example, whether their borders touch or not. Once this neighborhood structure is defined, models resembling autoregressive models in time series are considered. Two very popular such models that incorporate this discrete neighbor information are known as the simultaneously and conditionally autoregressive models, i.e. SAR and CAR models. The SAR and CAR models were originally developed as models on the doubly infinite regular lattice beginning with Whittle (1954) for the SAR model and Besag (1974) for the CAR model. When used for modeling a doubly infinite regular lattice, these models are quite analogous to the well understood stationary autoregressive time series model defined on the integers. That is, the CAR is analogous in its Markov property, and the SAR in its functional form (Cressie 1993, Sections 6.3, 6.4). But, when these models are applied to irregular lattices, the effect that the neighborhood structure and the spatial correlation parameter have on the implied covariance structure is not well understood and has not been explicitly examined.

For the finite regular lattice, several authors have pointed out that the covariance structure implied by the SAR and CAR models yield non constant variances at each site as well as unequal covariances between regions that are the same number of neighbors apart, see, e.g. Haining (1990), Besag and Kooperberg (1995). In this paper we give more detailed description of the implied structure of these models and in particular

look at their structure on an irregular lattice. In Section 2 the SAR and CAR models are defined and their standard use discussed. Section 3 presents a spatial regression example on the irregular lattice of the United States where the spatial structure is compared using the SAR and CAR models as well as an exponential variogram and an i.i.d. model. The example demonstrates the difference in predictions obtained by these models and the differing correlations between 1st order neighbors that occur when using the SAR or CAR models. Section 4 looks at the SAR and CAR correlation structures for the U.S. lattice in general as a function of the "spatial dependence" parameter. Conclusions and discussion are in Section 5.

2. The SAR and CAR models

Let $\{Z(A_i): A_i \in (A_1 \dots A_n)\}$ be a Gaussian random process where $\{A_1 \dots A_n\}$ forms a lattice of D. We say the regions $\{A_1 \dots A_n\}$ form a lattice of D if $\{A_1 \dots A_n\}$ is a simple partition of D, i.e. $A_1 \cup A_2 \cup \cdots \cup A_n = D$ and $A_i \cap A_j = 0$ for all $i \neq j$.

One way to model this process is by the simultaneous autoregressive model (SAR)

$$Z(A_i) = \mu_i + \sum_{j=1}^{n} b_{ij}(Z(A_j) - \mu_j) + \varepsilon_i \tag{1}$$

where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)' \sim N(\mathbf{0}, \Lambda)$ with Λ diagonal, $E(Z(A_i)) = \mu_i$, and b_{ij} are known or unknown constants and $b_{ii} = 0$, $i = 1 \dots n$. This model is called *simultaneous* because in general the error terms ε_i will be correlated with $\{Z(A_j): j \neq i\}$. If n is finite, we can take $\mathbf{B} = (b_{ij})$ to be a matrix containing the b_{ij}. The joint distribution of $\mathbf{Z} = (Z(A_1), Z(A_2), \dots, Z(A_n))'$ is then

$$\mathbf{Z} \sim N(\boldsymbol{\mu}, (\mathbf{I}_n - \mathbf{B})^{-1}\Lambda(\mathbf{I}_n - \mathbf{B})^{-1'}). \tag{2}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$ and $\mathbf{I}_n$ is the n dimensional identity matrix.

Another way to model $\{Z(A_i): A_i \in (A_1 \dots A_n)\}$ is with the conditional autoregressive model (CAR)

$$Z(A_i)|Z(A_{(-i)}) \sim N\left(\mu_i + \sum_{j=1}^{n} c_{ij}(Z(A_j) - \mu_j), \tau_i^2\right) \tag{3}$$

where $Z(A_{(-i)}) = \{Z(A_j): j \neq i\}$, $E(Z(A_i)) = \mu_i$, τ_i^2 is the conditional variance, and c_{ij} are known or unknown constants, in particular $c_{ii} = 0$, $i = 1 \dots n$. If n is finite, we form the matrices $\mathbf{C} = (c_{ij})$ and $\mathbf{T} = \text{diag}\{\tau_1^2, \tau_2^2, \dots, \tau_n^2\}$ and by the factorization theorem (see, e.g., Besag, 1974)

$$\mathbf{Z} \sim N(\boldsymbol{\mu}, (\mathbf{I}_n - \mathbf{C})^{-1}\mathbf{T}). \tag{4}$$

The structure of $\mathbf{B}$ and $\mathbf{C}$ is usually mostly specified by the shape of the lattice. One common way to construct $\mathbf{B}$ or $\mathbf{C}$ is with a single parameter that scales a user defined neighborhood matrix $\mathbf{W}$ that indicates whether the regions are neighbors or not.

One common way to do this is to define $\boldsymbol{W} = (w_{ij})$ as

$$w_{ij} = \begin{cases} 1 & \text{if region } A_i \text{ shares a common edge or border with region } A_j \\ 0 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Thus, for the SAR model $\boldsymbol{B} = \rho_s \boldsymbol{W}$ and for the CAR model $\boldsymbol{C} = \rho_c \boldsymbol{W}$ where ρ_s and ρ_c are often referred to as "spatial correlation or spatial dependence" parameters and are left to be estimated. There are other ways to define the neighborhood structure $\boldsymbol{W}$, e.g. restricting rows of $\boldsymbol{W}$ to sum to 1 or creating more elaborate weights as functions of the length of borders. Clayton and Bernardinelli (1992) point out that the specification of $\boldsymbol{W}$ as above simply with 0 and 1's is not internally consistent in the case in which the number of neighbors varies (which is the case with most irregular lattices). They recommend a weighting scheme $\boldsymbol{W} = (w_{ij}^*)$ such that $w_{ij}^* = w_{ij}/w_{i+}$ so that the expected conditional means form an average rather than a sum. Note, for the CAR model, it is necessary that $\boldsymbol{W}$ and $\boldsymbol{T}$ satisfies the symmetry condition: $w_{ij}\tau_j^2 = w_{ji}\tau_i^2$. So if $\boldsymbol{W} = (w_{ij}^*)$ is used, the conditional variances τ_i should be proportional to $1/w_{i+}$.

The SAR and CAR models specified in this way with a single parameter times some weight matrix have been used extensively for modeling irregular lattices in the applied literature (in econometrics: e.g. Anselin and Florax (1995) and Kelejian and Prucha (1999); and in disease mapping: e.g. Clayton and Kaldor (1987), Mollie and Richardson (1991), Bell and Broemeling (2000), and Stern and Cressie (2000), and incorporated into software: e.g. S-Plus Spatial Stat and BUGS).

Besag et al. (1991) introduced the intrinsic conditional autoregressive model (ICAR) (extending Kunsch's (1987) terminology to irregular domains) which is also popular in disease mapping and image restoration literature. This model which can be considered a limiting case for the CAR yields an improper joint distribution for the $\boldsymbol{Z}$. Correlations do not exist for the ICAR process and it will not be considered here. The ICAR has been studied widely elsewhere, for example see Yasui et al. (2000) and references therein.

3. An example

To provide an example of this type of ecological spatial regression using these models we consider state level summary data related to the SAT college entrance exam for the year 1999. These data were recorded from an article written in the Minneapolis Star Tribune on Wednesday September 1, 1999 which gave the College Board as its source. The data include statewide average verbal scores on the SAT as well as the percent of eligible students taking exam in the particular state. Fig. 1 shows a histogram of the verbal scores and a scatter plot of the verbal scores by the percent of eligible students taking the exam. Fig. 2 shows a choropleth map (by quartiles) of the scores by state. The map shows an indication that states in the Midwest have higher averages on the SAT verbal and the scatterplot shows a clear indication that a strong inverse
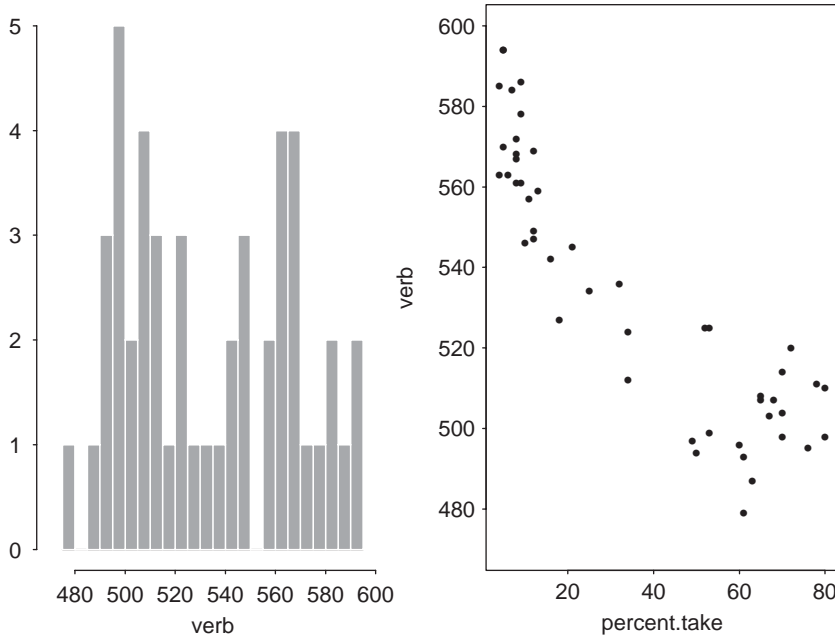
Fig. 1. Left: histogram of 48 contiguous state average SAT verbal scores for 1999, Right: scatterplot of state average SAT verbal scores against percent of students eligible who actually took the exam.
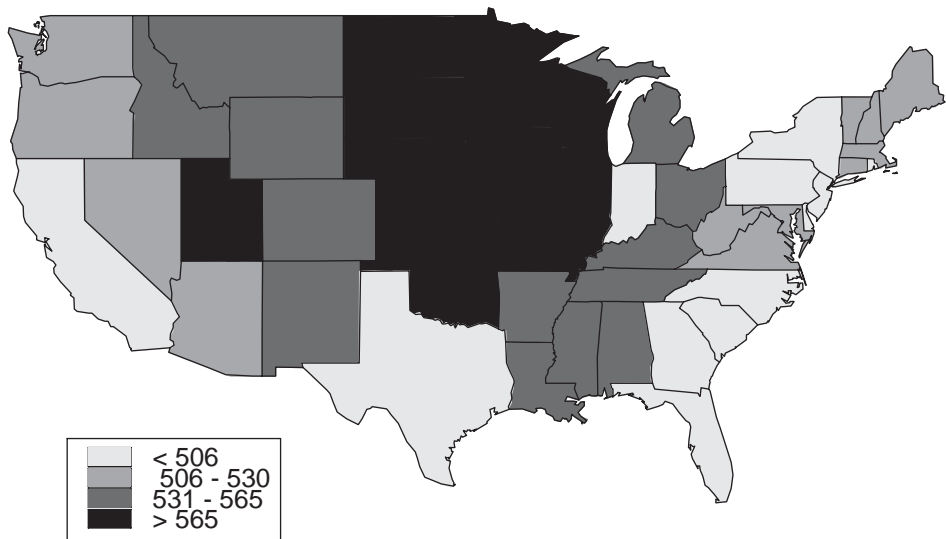

Fig. 2. Choropleth map of 48 contiguous state average SAT verbal scores for 1999.

Table 1
Results for fitting model 5 using 4 different models for $\boldsymbol{u}$

	SAR	CAR	Exponential variogram	IID
$\hat{\beta}_0$ (s.e.)	583.79 (4.89)	584.63 (4.86)	583.64 (6.49)	590.26 (3.93)
$\hat{\beta}_1$ (s.e.)	−2.42 (0.32)	−2.48 (0.32)	−2.19 (0.31)	−2.88 (0.30)
$\hat{\beta}_2$ (s.e.)	0.0183 (0.004)	0.0189 (0.004)	0.0146 (0.004)	0.0233 (0.004)
Spatial parameters	$\rho_s = 0.60$	$\rho_c = 0.83$	range $= 490$	N/A
Variance parameters	$\sigma_s^2 = 409.2$	$\sigma_c^2 = 408.5$	$\sigma_{g1}^2 = 11.4$, $\sigma_{g2}^2 = 148$	$\sigma^2 = 124.4$
Log likelihood	−179.9	−180.1	−177.2	−183.9

relationship exists between the average verbal score and the percent of students who took the exam.

Let $Z(A_i)$ represent the average SAT verbal score in state A_i and $X(A_i)$ the percent of eligible students who actually took the exam in state A_i, $i = 1 \ldots 48$. We consider the model

$$Z(A_i) = \beta_0 + \beta_1 X(A_i) + \beta_2 (X(A_i))^2 + u(A_i), \tag{5}$$

where $\boldsymbol{u} = (u(A_1), u(A_2), \ldots, u(A_{48}))'$ is assumed to be normally distributed with mean zero. We consider four different covariance structures for $\boldsymbol{u}$. First we consider the SAR and CAR spatial structure using the $\boldsymbol{W} = (w_{ij}^*) = (w_{ij}/w_{i+})$ neighbor structure defined in Section 2 that takes all states touching each other with an edge to be nonzero. Specifics for the neighborhood structure of the US lattice are given in Appendix A. For the CAR model we take $\boldsymbol{T} = \sigma_c^2 \operatorname{diag}(1/w_{i+})$ and for the SAR model we take $\boldsymbol{\Lambda} = \sigma_s^2 \operatorname{diag}(1/w_{i+})$. While the non-constant specification for $\boldsymbol{T}$ is necessary for the CAR to satisfy the symmetry condition, the reason for choosing similar $\boldsymbol{\Lambda}$ in the SAR model is only for comparison, i.e. it is not necessary. We then consider an isotropic exponential variogram structure for the $\boldsymbol{u}$ where the centroid of each state is used to calculate distances between states (i.e. $|d_{ij}|$ is the distance between the centroids of state A_i and A_j) and covariances are defined by $\operatorname{Cov}(u(A_i), u(A_j)) = \sigma_{g1}^2 + \sigma_{g2}^2 * \exp(-|d_{ij}|/\text{range})$. Finally we consider i.i.d. structure for $\boldsymbol{u}$ where $\operatorname{Var}(\boldsymbol{u}) = \sigma^2 \mathbf{I}_{48}$. The results are shown in Table 1.

As can be seen from Table 1, the estimated β_0, β_1, and β_2 are similar in all four cases—95% confidence intervals around them would all overlap. Notice that as might be expected, the i.i.d. model yields the smallest standard errors because it is assuming that there are 48 independent observations. Using Moran's I we test for the existence of spatial structure in the residuals from the i.i.d. model and find a significant p-value indicating that there is spatial correlation missed by simply fitting the large scale trend. The SAR and CAR models yield spatial parameters of $\rho_s = 0.60$ and $\rho_c = 0.83$ and the range parameter for the exponential variogram is 490 miles. The variance parameters in the different models are not directly comparable, but it is possible to make some conclusions. Recall that $\boldsymbol{T} = \sigma_c^2 \operatorname{diag}(1/w_{i+})$ and $\boldsymbol{\Lambda} = \sigma_s^2 \operatorname{diag}(1/w_{i+})$ for the CAR and SAR models and note that the average number of neighbors the states have is approximately 4.45. So roughly, the average variance in $\boldsymbol{T}$ or $\boldsymbol{\Lambda}$ is $\sigma_c^2/4.45 = 91.80$ or $\sigma_s^2/4.45 = 91.96$
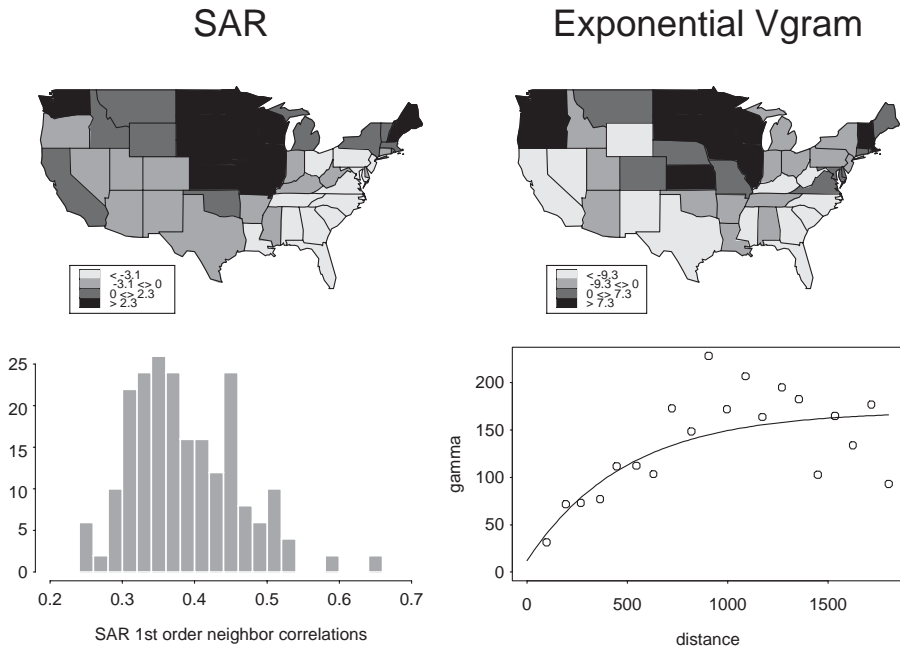
Fig. 3. Top: Predicted small-scale spatial structure using the SAR and exponential variogram, Bottom: (Left) histogram of the 1st order neighbor correlations implied by the SAR model, (right) correlation structure implied by the exponential variogram model, fitted variogram overlayed empirical variogram.

respectively. The nugget effect, σ_{g1}^2, representing the non-spatial random error in the exponential covariogram model is much smaller than the similar variances T and Λ in the CAR and SAR models implying that the exponential variogram model will fit the data more closely than CAR or SAR. Finally, we present the log-likelihood values for each model but note these should be compared with caution. None of the models are nested within any of the other models. Note the i.i.d. model fit here is not nested within the SAR and CAR models because a constant variance $\sigma^2 \mathbf{I}$ is assumed for the i.i.d. model and even when $\rho_s = 0$ or $\rho_c = 0$, the SAR and CAR models still have non constant variance because of the specification for Λ and T.

Now we focus attention on the prediction of the small-scale spatial structure, i.e., the predicted values for the $\boldsymbol{u}$ under the SAR, CAR, and exponential variogram model. We want to examine the spatial structure remaining after taking out the large scale trend due to the effect of the percent of students taking the exam. We find that the CAR and SAR models give very similar predicted value for $\boldsymbol{u}$ with a correlation between them of 0.998 but the CAR predictions have larger variability than the SAR predictions (Top left, Fig. 4). The predicted values for $\boldsymbol{u}$ using the exponential variogram are different and have correlation 0.612 with the SAR and CAR values. In the top of Fig. 3 we present choropleth maps of the predicted small-scale spatial structure using the SAR and exponential variogram (CAR is not shown because it is nearly identical to SAR).
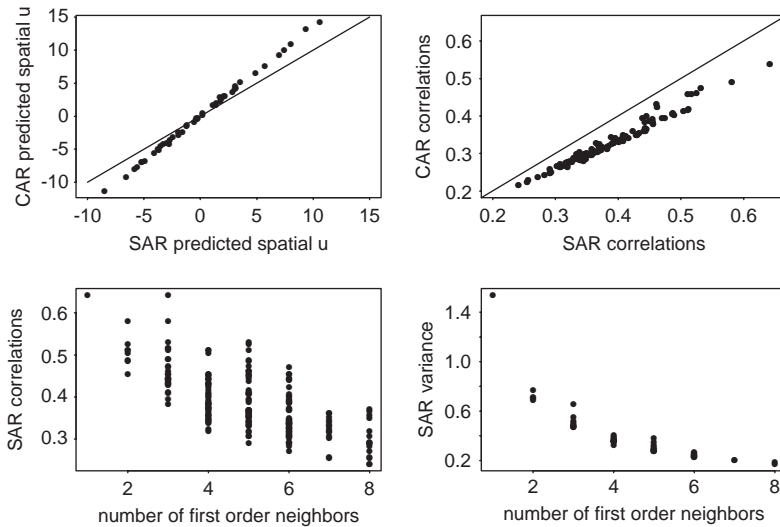
Fig. 4. Top: Comparison of SAR and CAR results for predicted spatial **u** (top left) and first order correlations (top right); Bottom: Stratified by the number of first order neighbors, correlation (bottom left) and variances (bottom right) are given for the SAR model.

The graduated color has been meaningfully separated at zero since zero represents the value where the predicted verbal score is exactly given by the large scale model $\beta_0 + \beta_1 X(A_i) + \beta_2 (X(A_i))^2$. Values greater than zero imply that the small-scale spatial structure model predicts the state average to be higher than what would have been predicted by the large scale model alone, and visa versa for negative values.

The SAR model results in a "smoother" picture than the exponential variogram which perhaps is part of the reason for its popularity in practice. But what does the SAR model say about the spatial correlation? The bottom left plot in Fig. 3 shows a histogram of all the first order neighbor correlations implied by the SAR model. The smallest correlation is 0.24 which occurs between Missouri and Tennessee and the largest correlation, equal to 0.64, occurs between Maine and New Hampshire. For these extreme cases we note that Maine is the only state with just one neighbor (i.e. New Hampshire) and that Tennessee and Missouri are the only two states with eight neighbors (the largest number of neighbors) and they are neighbors of each other. So it seems that the implied correlation might simply be related to the number of neighbors each region has. To the contrary, the bottom left plot of Fig. 4 shows this relationship is not simple. We further emphasize the point in Table 2, showing that both Tennessee and Missouri have eight different correlations for each of their eight different first order neighbors. The structure of the neighbor matrix in the SAR model implies that Tennessee is more correlated with Alabama than it is with Mississippi and that Missouri is more like Kansas than Iowa. Is this reasonable? Because there is no systematic structure to the SAR (or CAR) covariance model, there is not a good way to examine whether it is a reasonable model for describing the spatial structure of the

Table 2
Implied correlation between Tennessee and its first order neighbors and between Missouri and its first order neighbors for the SAR and CAR Models (numbers in parenthesis are labels from Fig. 6)

Tennessee (40)			Missouri (23)		
1st order neighbors	SAR	CAR	1st order neighbors	SAR	CAR
Alabama (1)	0.371	0.324	Arkansas (3)	0.272	0.238
Arkansas (3)	0.291	0.257	Illinois (11)	0.291	0.247
Georgia (9)	0.365	0.327	Iowa (13)	0.282	0.244
Kentucky (15)	0.256	0.229	Kansas (14)	0.319	0.263
Mississippi (22)	0.349	0.300	Kentucky (15)	0.255	0.223
Missouri (23)	0.241	0.216	Nebraska (25)	0.291	0.248
North Carolina (31)	0.358	0.312	Oklahoma (34)	0.293	0.251
Virginia (44)	0.306	0.265	Tennessee (40)	0.241	0.216

data. On the other hand, for the exponential variogram (because we can completely and succinctly describe the implied covariance structure) we at least have a way to check if its fit to the residual spatial structure is reasonable by examining the fitted model through the empirical variogram, bottom right of Fig. 3.

4. Relation between ρ_s and ρ_c and the implied spatial correlation

In this section we examine the correlation structure for the SAR and CAR models in general as a function of their "spatial dependence" parameters ρ_s and ρ_c. Note that when the SAR and CAR covariance matrices, $(\mathbf{I}_n - \rho_s \mathbf{W})^{-1} \boldsymbol{\Lambda} (\mathbf{I}_n - \rho_s \mathbf{W})^{-1\prime}$ and $(\mathbf{I}_n - \rho_c \mathbf{W})^{-1} \boldsymbol{T}$ where $\boldsymbol{\Lambda} = \sigma_s^2 \operatorname{diag}(1/w_{i+})$ and $\boldsymbol{T} = \sigma_c^2 \operatorname{diag}(1/w_{i+})$ are standardized to be correlation matrices, they are functions of only $\mathbf{W}$ and ρ_s or ρ_c. For demonstration we consider the US lattice neighbor matrix $\mathbf{W} = (w_{ij}^*)$ and focus on how the model correlations behave as functions of the true parameters ρ_s and ρ_c (i.e. irrespective of data).

The usual restrictions on the parameter spaces of ρ_s and ρ_c are given as $\{\rho_s: \rho_s \omega_i < 1\}$ and $\{\rho_c: \rho_c \omega_i < 1\}$ for $i = 1 \ldots n$ (see, e.g. Haining 1990, p. 82) where ω_i are the eigenvalues of $\mathbf{W}$. For the CAR model where the covariance is $(\mathbf{I}_n - \rho_c \mathbf{W})^{-1} \boldsymbol{T}$, the restriction on ρ_c is a necessary condition to ensure positive definiteness. For the SAR model the condition for ρ_s is too strong (Kelejian and Robinson, 1995) since it is only necessary that $(\mathbf{I}_n - \rho_s \mathbf{W})$ is nonsingular. This is satisfied by requiring ρ_s such that $\rho_s \neq 1/\omega_i$ for $i = 1 \ldots n$. Whether this much broader parameter space is of any use is questionable because the interpretation of ρ_s becomes extremely difficult or impossible when it is outside of the commonly considered region $\{\rho_s: \rho_s \omega_i < 1\}$ for all $i = 1 \ldots n$. Based on the eigenvalues of $\mathbf{W}$ for the US lattice, a restriction for the parameter space of ρ_c and ρ_s is $(-1.392, 1)$. (The upper limit of 1 is a result of the fact that the rows of $\mathbf{W}$ were taken to sum to one.)

Here we consider how the implied correlations between all the first order neighbors based on these two models change as a function of ρ_s or ρ_c. On the US lattice there
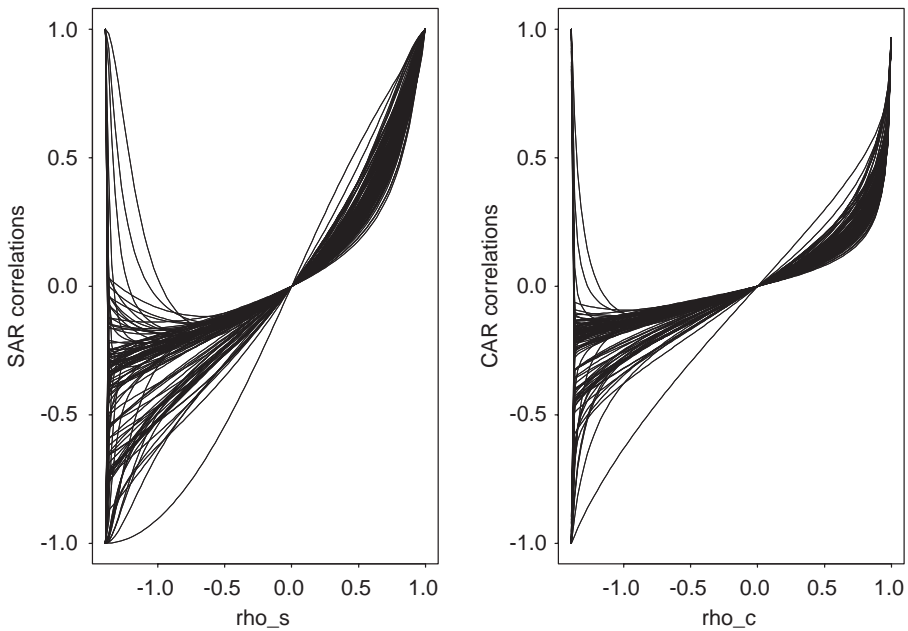
Fig. 5. Lines in both plots represent the implied model correlations between first order neighbors in the US lattice based on the SAR model (left) and CAR model (right) as functions of the respective parameters ρ_s and ρ_c.

are 107 pairs of states that are considered first order neighbors, that is, their borders touch. Fig. 5 shows the implied correlations between these neighbors as a function of ρ_s and ρ_c. Immediately we see in Fig. 5 that for any given ρ_s or ρ_c, there is variability in the correlations among all the first order neighbors. This variability in the correlations changes as a function of ρ_s and ρ_c, for example, correlations range from 0.03 to 0.19 when $\rho_s = 0.1$ while the range is much larger for $\rho_s = 0.6$ where the correlations range from 0.24 to 0.64. It is also clear from Fig. 5 that the first order neighbor correlations increase at a slower rate as a function of ρ_c in the CAR model than for ρ_s in the SAR model. A few things about Fig. 5 are intuitively pleasing. As the "spatial correlation" parameter (ρ_s or ρ_c) increases from zero to the upper end of the parameter space, the implied correlations between all sites monotonically increase. This matches our intuition from autoregressive models in time series that says: as the autoregressive parameter increases from zero, the correlation between times increases. Another is that as the "spatial correlation" parameter (ρ_s or ρ_c) reaches the endpoints of the parameter space, the implied correlations between all the pairs of sites tend toward 1 or −1. (Although it is difficult to see in Fig. 5, as ρ_s and ρ_c reach the lower end of interval −1.392, all of the correlations do approach 1 or −1).

We now point out probably the most displeasing result of these models. That is, the ranking of the implied correlations from largest to smallest is not consistent as ρ_s and

ρ_c change. In other words, there are lines in Fig. 5 that cross each other. For example, when $\rho_c = 0.49$ the Corr(Alabama, Florida) $= 0.20$ while the Corr(Alabama, Georgia) $=$ 0.16. But, when $\rho_c = 0.975$ the correlation between Alabama and Georgia is greater than the correlation between Alabama and Florida, i.e. Corr(Alabama, Florida) $= 0.65$ while the Corr(Alabama, Georgia) $= 0.67$. Thus the already difficult to interpret correlations become even more difficult to understand when we realize their relation to one another can change depending on the "spatial correlation" parameter.

Further non-intuitive behavior is seen in Fig. 5 when ρ_s or ρ_c is negative. The implied correlation between some first order neighbors can be positive but that depends on what value of ρ_s or ρ_c is being considered. These particular first order neighbor correlations are negative when ρ_s or ρ_c is slightly negative but as ρ_s or ρ_c become more negative, the implied correlations become positive. For example, when $\rho_s = -0.716$ the Corr(Maryland, Penn.) $= -0.18$ and Corr(Vermont, Mass) $= -0.15$ but when rho $= -1.32$, Corr(Maryland, Penn.) $= 0.085$ and Corr(Vermont, Mass) $= 0.94$. Out of 107 first order neighbor pairs in the US lattice, 34 end up having positive correlations when ρ_s or ρ_c is negative. Note the pairs are the same for both the SAR and CAR models. It is unclear what distinguishes these pairs of sites from the others. We have tried looking for similarities among the neighbor patterns of these state pairs but found nothing. For example, they do not all have even or odd numbers of neighbors and they are not located in any particular region of the graph.

5. Summary

It has been demonstrated that the implied spatial correlation between the different states using the SAR and CAR models does not seem to follow an intuitive or practical scheme. For instance, there does not appear to be any reason in general why a researcher would want to fit a spatial model that insists on Missouri and Tennessee being the least spatially correlated states in the land. And why should Missouri be more correlated with Kansas than with Iowa? This is what the SAR and CAR models with the $\boldsymbol{W} = (w_{ij}^*)$ neighbor matrix imply. It is also noted that similar nonintuitive spatial structure occurs when the weight matrix is simply taken to be the 0,1 matrix, i.e. $\boldsymbol{W} = (w_{ij})$.

Cressie (1993) refers to $\boldsymbol{B}$ and $\boldsymbol{C}$ in (2) and (4) as the "spatial-dependence matrix" in the model, and, Ord (1975) says that the (i, j)th element of these matrices "represents the degree of possible interaction of location j on location i". These descriptions are misleading because they seem to imply that one can examine the structure of $\boldsymbol{B}$ and $\boldsymbol{C}$ directly to understand the spatial correlation being modeled for $\{Z(A_i): i = 1 \ldots n\}$. This is not the case, since it is the inverses $(\mathbf{I}_n - \boldsymbol{B})^{-1}$ and $(\mathbf{I}_n - \boldsymbol{C})^{-1}$, respectively, that actually explain the spatial structure. And as we have seen in this paper, although these covariances are clearly just functions of $\boldsymbol{B}$ or $\boldsymbol{C}$, in general there is no obvious intuitive connection between them and the resulting spatial correlations.

From this discussion it seems that other ways of modeling lattice data which directly model the covariance structure such as geostatistical models should be considered especially when there is interest in understanding the spatial structure. In an attempt to

alleviate the undesirable properties implied by the CAR model, Besag and Kooperberg (1995) considered a method which is a "partial synthesis of standard geostatistical and Gaussian Markov random field formulations".

Despite their popularity, these SAR and CAR models have been fit over and over again without much emphasis placed on trying to decipher what they mean. This may be due to the fact that often the primary interest in analyses that incorporate them is determining significant predictors in a regression rather than understanding the spatial structure itself. However, if there is any chance of determining whether the SAR or CAR model provide a good fit for the data, it seems prudent to first understand the SAR and CAR models. The focus of this paper has been to make transparent and point out possible problems with the way these models incorporate the geographic structure of the lattice into the spatial covariance structure. The hope is that clarification of these problems may lead to advances in their solution.

Appendix A

The matrix W contains many zeros and a simple description of the neighborhood structure is given below. This method lists row-by-row each of the 48 states followed by that state's neighbors. The function *read.neighbor* in the S+Spatial Stat package reads this neighborhood structure and by default creates a (0,1) W matrix which can then be scales so that the non-zero elements equal $1/(w_{i+})$.

The list below indicates the neighborhood structure for the 48 contiguous states. Note they are numbered in alphabetical order (the numbers correspond to the map in Fig. 6).
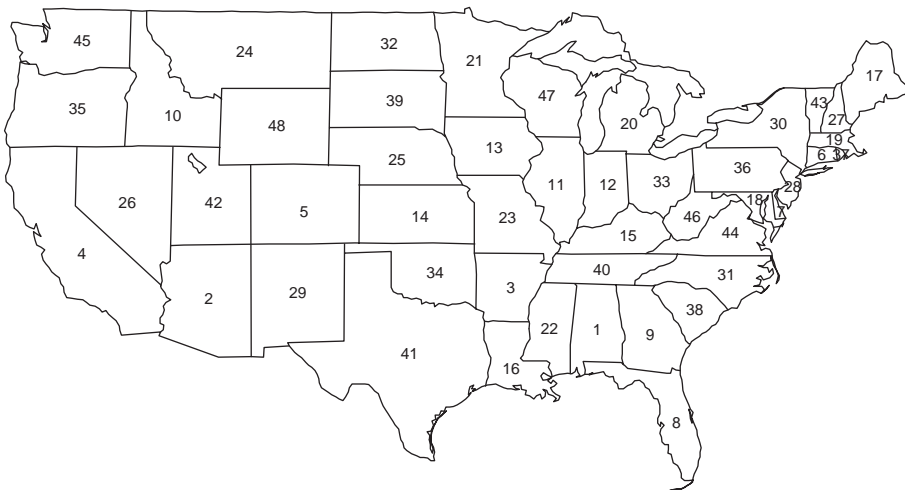

Fig. 6. US map with states numbered.

```
1 8 9 40 22
2 4 26 42 5 29
3 23 40 22 16 41 34
4 35 26 2
5 48 25 14 34 29 2 42
6 30 19 37
7 28 18 36
8 9 1
9 1 8 38 31 40
10 45 35 26 42 48 24
11 47 12 15 23 13
12 11 20 33 15
13 21 47 11 23 25 39
14 25 23 34 5
15 23 11 12 33 46 44 40
16 41 3 22
17 27
18 36 7 44 46
19 43 27 37 6 30
20 12 33 47
21 32 39 13 47
22 16 3 40 1
23 13 11 15 40 3 34 14 25
24 10 48 39 32
25 39 13 23 14 5 48
26 35 10 42 2 4
27 43 19 17
28 7 30 36
29 2 42 5 34 41
30 43 19 6 28 36
31 44 40 9 38
32 24 39 21
33 20 12 15 46 36
34 14 23 3 41 29 5
35 45 10 26 4
36 30 28 7 18 46 33
37 6 19
38 31 9
39 32 24 48 25 13 21
40 15 44 31 9 1 22 3 23
41 29 34 3 16
42 10 48 5 29 2 26
43 30 19 27
44 46 18 31 40 15
45 35 10
46 33 36 18 44 15
47 20 11 13 21
48 24 39 25 5 42 10
```

References

Anselin, L., Florax, R. (Eds.), 1995. New directions in Spatial Econometrics. Springer, Berlin.

Bell, B.S., Broemeling, L., 2000. A Bayesian analysis for spatial processes with application to disease mapping. Statist. Med. 19, 957–974.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion). J. Roy. Statist. Soc. Ser. B 36, 192–236.

Besag, J., Kooperberg, C., 1995. On conditional and intrinsic autoregression. Biometrika 82, 733–746.

Besag, J., York, J., Mollie, A., 1991. Bayesian image restoration with two application in spatial statistics. Ann. Inst. Statist. Math. 43 (1), 1–59.

Clayton, Bernardinelli, 1992. Bayesian methods for mapping disease risk. In: Elliott, P., Cuzick, J., English, D., Stern, R. (Eds.), Geographical and Environmental Epidemiology: Methods for Small-Area Studies, pp. 205–220.

Clayton, D., Kaldor, J., 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 43, 671–681.

Cressie, N., 1993. Statistics for Spatial Data. Wiley, New York.

Haining, R., 1990. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press, Cambridge.

Kelejian, H.H., Prucha, I.R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. Internat. Econom. Rev. 40 (2), 509–533.

Kelejian, H.H., Robinson, D.P., 1995. A suggested alternative to the autoregressive model. In: Anselin, L., Florax, R. (Eds.), New Directions in Spatial Econometrics. Springer, Berlin.

Kunsch, H.R., 1987. Intrinsic autoregressions and related models on the two dimensional lattice. Biometrika 74, 517–524.

Mollie, A., Richardson, S., 1991. Empirical bayes estimates of cancer mortality rates using spatial models. Statist. Med. 10, 95–112.

Ord, K., 1975. Estimation methods for models of spatial interaction. J. Acoust. Soc. Amer. 70, 120–126.

Stern, H., Cressie, N., 2000. Posterior predictive model checks for disease mapping models. Statist. Med. 19, 2377–2397.

Whittle, P., 1954. On stationary process in the plane. Biometrika 41, 434–449.

Yasui, Y., Liu, H., Benach, J., Winget, M., 2000. An empirical evaluations of various priors in the empirical Bayes estimation of small area disease risks. Statist. Med. 19, 2409–2420.