Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

A multivariate CAR model for mismatched lattices[☆]

Aaron T. Porter^{a,*}, Jacob J. Oleson^b

^a Colorado School of Mines, Department of Applied Mathematics and Statistics, United States

^b University of Iowa, Department of Biostatistics, United States

ARTICLE INFO

Article history:

Received 24 September 2013

Revised 29 May 2014

Accepted 25 August 2014

Available online 18 September 2014

Keywords:

Conditional autoregressive

Mismatched lattices

Infectious disease

American Community Survey

ABSTRACT

In this paper, we develop a multivariate Gaussian conditional autoregressive model for use on mismatched lattices. Most current multivariate CAR models are designed for each multivariate outcome to utilize the same lattice structure. In many applications, a change of basis will allow different lattices to be utilized, but this is not always the case, because a change of basis is not always desirable or even possible. Our multivariate CAR model allows each outcome to have a different neighborhood structure which can utilize different lattices for each structure. The model is applied in two real data analysis. The first is a Bayesian learning example in mapping the 2006 Iowa Mumps epidemic, which demonstrates the importance of utilizing multiple channels of infection flow in mapping infectious diseases. The second is a multivariate analysis of poverty levels and educational attainment in the American Community Survey.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Conditional autoregressive (CAR) models are common in the literature, as these models represent an intuitive way of incorporating spatial autocorrelation on lattices. The univariate CAR structure has been well developed, with major contributions in the field coming as early as 1974 (Besag, 1974). Excellent references on the topic include Cressie (1993) for the frequentist formulation of CAR models, and Banerjee et al. (2004) for the Bayesian formulation. These models are extremely flexible and have been used in a variety of fields, including noninfectious disease mapping (e.g., Clayton and Kaldor, 1987; Kim and Oleson, 2008; Sun et al., 2000), ecology (e.g., Lichstein

et al., 2002), and image processing (e.g., Chellappa and Chatterjee, 1985), to name just a few.

Since multivariate outcomes are often of interest, multivariate CAR models have also been developed. Early multivariate work was done by Mardia (1988), who proposed a very general multivariate CAR structure. Kim et al. (2001) proposed a two-fold model for mapping two correlated mortality rates. Additional work has been done by Sain and Cressie (2002), Gelfand and Vounatsou (2003) and Jin et al. (2005). A common theme among these models is that they utilize the same neighborhood structure for each of the multivariate outcomes. The model of Hossain et al. (2014) is of particular interest. It is a very general model based on the conditioning that contains many of the common multivariate CAR models in the literature as special cases. The model can be applied to data sets with more than two outcomes, but this is not demonstrated. Working directly with differing neighborhood structures is possible in their framework, but we do not know of applications of this model to differing neighborhood structures in the literature, potentially because the conditional structure makes interpretation of these results difficult.

[☆] This research was partially supported by the U.S. National Science Foundation (NSF) and the U.S. Census Bureau under NSF grant SES-1132031, funded through the NSF-Census Research Network (NCRN) program.

* Corresponding author. Tel.: +1 (303) 273-3677.

E-mail addresses: aporter@mines.edu (A.T. Porter), jacob-oleson@uiowa.edu (J.J. Oleson).

Work by [Sain et al. \(2011\)](#) has extended the univariate CAR structure to include different neighborhood structures for each variable. We prefer this model structure for modeling outcomes with respect to different neighborhood structures, as it allows asymmetric weighting between outcomes, which is one of the advantages of univariate CAR models. The application they utilize assumes the same lattice for every neighborhood structure, but the general full conditional mean structure they consider can be extended to mismatched lattices. Additionally, their formulation induces the potentially restrictive requirement that the variances of each outcome with a fixed lattice point differ by a known and fixed proportion. For many analyses, this constraint is reasonable. Expected counts or inverse weightings of the number of neighbors are often available and can be used in this regard. However, we will provide evidence that the variances are one of the most important features in this model, and we provide an extension of the model allowing the variances to be more flexibly estimated than is considered in their analyses.

In this paper, we develop a multivariate CAR model that extends the use of the concept of differing neighborhood structures found in [Sain et al. \(2011\)](#). We utilize a similar mean structure, but construct our model with the goal analyzing data where lattices may be mismatched. While a change of support is often possible for mismatched lattices, this technique is not always desirable, and is occasionally impossible, as in the case where measurements do not make sense over certain regions of an areal unit. In order to analyze mismatched lattices, we will relax the variance restrictions in [Sain et al. \(2011\)](#)'s formulation of the model. This will allow our model to perform unique inference even on matched lattices, and this is the second benefit of the structure we propose.

Specific examples of applications that would benefit from mismatched lattices or asymmetric cross correlation are easy to construct. We provide two motivating examples.

Our first motivating example considers infectious disease modeling. In 2006, there was a large Mumps epidemic in Iowa (described in detail in Section 3). As Iowa is predominantly rural, a standard CAR model which utilizes geographic borders to determine adjacency could be used to relate neighboring counties in the infectious disease process. However, it is clear from the data that cases of Mumps tended to spread along the highway system. This is an example of how multiple modes of travel and contact can spread an infectious agent. In a relatively rural state like Iowa, highway systems may serve as an important measure of adjacency for infectious disease spread, just as airline travel may be an important measure of adjacency in pandemic modeling on a larger geographical scale (e.g., [Hufnagel et al., 2004](#)).

While noninfectious disease mapping with CAR models is common, there is far less literature using CAR models for infectious disease mapping. One of the reasons is likely that many infectious diseases are spread via human to human transmission, and the geographical proximity of two locations may not be an accurate assessment of the overall proximity of the locations in terms of disease spread. This makes a reasonable neighborhood structure

difficult to construct. In the modern world, proximity is often defined by travel, rather than geographical measures. [Hufnagel et al. \(2004\)](#) showed that airline travel served as an accurate measure of proximity in the severe acute respiratory syndrome (SARS) pandemic of 2003. In general, it is likely that modern epidemics are spread via more than one causeway. For many epidemics, a realistic model may have to account for multiple modes of travel.

To test this hypothesis, one might consider modeling the final counts of Mumps in each county using four separate CAR models in a Poisson regression model in order to account for the spatial correlation between the counties. One model could utilize the standard CAR structure that considers any two counties that share a border as neighbors, while a second could consider any two counties that contain the same highway as neighbors. A third model could utilize both structures independently. This will allow for both lattice structures to be utilized directly.

Separating the spatial dependence in the disease map using two distinct structures certainly allows both measures of proximity to be modeled, but does not consider that the two latent effects are potentially correlated. A given county not only has correlation with other elements of its neighborhood structure, but across neighborhood structures as well. *A priori*, one might expect correlation between the modes of travel, as counties that have large numbers of people traveling along highways almost certainly have large numbers of people crossing county borders unassociated with highway travel, or that there may be negative correlation due to these being competing modes of travel. A model allowing for multivariate outcomes that have different spatial correlation parameters between multiple outcomes is desirable.

The aforementioned application is a Bayesian learning example in the sense that there is a single outcome informing two spatial structures, but the mismatching of the lattices will provide enough information to estimate the corresponding variance terms.

To further demonstrate the efficacy of this model, we provide a second example, this time modeling a multivariate outcome. We bivariate model the percent of the population in poverty and the percent of the population without a high school education in the Midwest as estimated by the American Community Survey (ACS) using our CAR model. In this analysis, poverty is collected on a county lattice, and high school education is collected on a state lattice. Variables such as these are often important covariates when modeling health outcomes, and this data set will act as a proof-of-concept of our model for health related outcomes. We will demonstrate that allowing cross correlation yields better model fit than traditional methods.

In Section 2 of this paper, we propose our multivariate CAR model that allows mismatched lattices to be utilized, and give conditions for consistency and propriety. We call this model the MMCAR, as it is designed with mismatched lattices in mind. We also discuss model properties, as well as giving a recommended parameterization guaranteeing consistency. Section 3 contains the results from the analysis of the Iowa Mumps epidemic. Section 4 contains an inferential analysis of the midwestern rates of poverty

and high school education as reported by the ACS. We close with a discussion in Section 5. Simulations studies and computational considerations may be found in the online web appendix.

2. Mismatched conditional autoregressive model (MMCAR Model)

2.1. Model definitions

Consider m outcomes, denoted by $\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_m)'$, where $\mathbf{X}'_k = (X_{k1}, \dots, X_{kn_k})'$, $k = 1, \dots, m$ and every X_{ki} coming from univariate normal distribution. Each $\mathbf{X}_k$ is collected on potentially different lattices, each with differing neighborhood structures, $C_1, \dots, C_m$. Within each neighborhood structure, we require positive correlation. Between neighborhood structures, we allow either positive or negative correlation. Consider potential spatial relationships between the elements of the two neighborhood structures C_k and C_l , $l \neq k$. The MMCAR model can be derived as follows:

Suppose

$$X_{ki} | \{X_{kj}, j \neq i, X_{lj}, l \neq k\} \sim N\left(\mu_{ki} + \sum_{kj \sim ki} \beta_{kij}(X_{kj} - \mu_{ki}) + \sum_l \sum_{lj \sim ki} \delta_{klj}(X_{lj} - \mu_{lj}), \sigma_{ki}^2\right), \quad (1)$$

where β_{kij} is the weight of the mean for X_{ki} associated with X_{kj} , δ_{klj} is the weight of the mean for X_{ki} associated with X_{lj} , μ_{ki} and μ_{lj} are the overall means of $\mathbf{X}_{ki}$ and $\mathbf{X}_{lj}$, respectively, and σ_{ki}^2 is the conditional variance of X_{ki} . We note here that some structure must be placed on μ_{ki} for the estimability of the model, such as $\mu_{ki} = \mu_k$ or $\mu_{ki} = x'_{ki}\beta$. As is common in the CAR modeling literature (e.g., Cressie, 1993, page 404), we assign $\beta_{kii} \equiv 0$ for every i and k . We use $ki \sim kj$ to denote that elements i and j are both units within neighborhood structure C_k and are neighbors. We use $ki \simeq lj$ to denote that element i of C_k is a neighbor of element j of C_l .

In this model, we have $\beta_{ki} = (\beta_{ki1}, \dots, \beta_{ki,n_k})'$ representing a vector of weights that define the contribution of members of C_k to the conditional mean of X_{ki} , and $\delta_{kli} = (\delta_{kli1}, \dots, \delta_{kli,l,n_l})'$ representing a vector of weights that define the contribution of members of C_l to the mean of X_{ki} .

We now derive necessary and sufficient conditions for which β and δ create a Markov Random Field. In order to derive the conditions under which the conditionals are self-consistent, we begin by using Brook's Lemma. Without loss of generality, assume $\mu_{ki} = \mathbf{0}$ for all k . Henceforth, we denote by $f(\mathbf{z})$ the value of the probability density function of $\mathbf{z}$ given the model parameters. To begin note that:

$$\begin{aligned} -2 \log(f(x_{ki}|x_{kj}, j \neq i, x_{lj}, l \neq k)) \\ = \frac{x_{ki}^2}{\sigma_{ki}^2} - 2 \frac{\sum_{kj \sim ki} \beta_{kij} x_{kj} x_{ki}}{\sigma_{ki}^2} - 2 \frac{\sum_l \sum_{lj \sim ki} \delta_{klj} x_{lj} x_{ki}}{\sigma_{ki}^2} + C, \end{aligned} \quad (2)$$

where C is an irrelevant constant. Brook's Lemma states

$$\frac{f(\mathbf{x})}{f(\mathbf{y})} = \prod_{i=1}^N \frac{f(x_i|x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_N)}{f(y_i|y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_N)}. \quad (3)$$

Plugging (2) into (3), with $y_i = 0$ for all i yields

$$\begin{aligned} -2 \log(f(\mathbf{x})/f(\mathbf{0})) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \frac{x_{ki}^2}{\sigma_{ki}^2} - 2 \sum_{k=2}^m \sum_{i=2}^{n_k} \frac{\sum_{j < i} \beta_{kij} x_{kj} x_{ki}}{\sigma_{ki}^2} \\ &\quad - 2 \sum_{k=2}^m \sum_{i=2}^{n_k} \frac{\sum_{l < k} \sum_{j=2}^{n_l} \sum_{h < j} \delta_{kljh} x_{lh} x_{ki}}{\sigma_{ki}^2}. \end{aligned} \quad (4)$$

Now, for the conditionals to be self-consistent, any permutation in the variable order when applying Brook's Lemma must equal (3). A particularly informative permutation is the "forward" version of Brook's Lemma:

$$\frac{f(\mathbf{x})}{f(\mathbf{y})} = \prod_{i=1}^N \frac{f(x_i|y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_N)}{f(y_i|y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_N)}. \quad (5)$$

Plugging (2) into (4), and once again using $y_i = 0$ for all i yields

$$\begin{aligned} -2 * \log(f(\mathbf{x})/f(\mathbf{0})) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \frac{x_{ki}^2}{\sigma_{ki}^2} - 2 \sum_{k=1}^{m-1} \sum_{i=1}^{n_k-1} \frac{\sum_{j > i} \beta_{kij} x_{kj} x_{ki}}{\sigma_{ki}^2} \\ &\quad - 2 \sum_{k=1}^{m-1} \sum_{i=1}^{n_k-1} \frac{\sum_{l > k} \sum_{j=1}^{n_l-1} \sum_{h > j} \delta_{kljh} x_{lh} x_{ki}}{\sigma_{ki}^2}. \end{aligned} \quad (6)$$

In general, one sees that the likelihood will contain all of the terms $\frac{x_{ki}^2}{\sigma_{ki}^2}$, and all of the cross products between x_{ki} and x_{kj} , $j \sim i$, as well as all the cross products between x_{ki} and x_{lj} , regardless of the permutation of variables selected when applying Brook's Lemma.

Because expressions (3) and (5) must be equal for the conditionals to be self-consistent, we must have

$$\frac{\beta_{kij}}{\sigma_{ki}^2} = \frac{\beta_{kji}}{\sigma_{kj}^2} \quad \text{and} \quad \frac{\delta_{klj}}{\sigma_{ki}^2} = \frac{\delta_{lkj}}{\sigma_{lj}^2}.$$

Additionally, it is desirable for $\sum_{kj \sim ki} \beta_{kij} + \sum_{l \neq k} \sum_{lj \sim ki} \delta_{klj} = 1$, as this indicates the weights in determining that the mean of X_{ki} sum to one, a requirement of Intrinsic Gaussian Markov Random Fields (IGMRFs), such as the Intrinsic CAR of Besag et al. (1991). Given these weights, one can write the pdf of $\mathbf{X}$:

$$f(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{I} - \boldsymbol{\Omega})(\mathbf{x} - \boldsymbol{\mu})\right), \quad (7)$$

where $\boldsymbol{\mu}$ is the vector of overall means, $\boldsymbol{\Sigma}$ is a diagonal matrix with elements σ_{ki}^2 , and

$$\boldsymbol{\Omega} = \begin{bmatrix} \beta_1 & \delta_{12} & \delta_{13} & \cdots \\ \delta_{21} & \beta_2 & \delta_{23} & \cdots \\ \delta_{31} & \delta_{32} & \beta_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Note that β_k is defined as the weight matrix for neighborhood structure k , and that δ_{kli} is an n_k by n_l matrix with entries of δ_{klj} as element (i, j) , and zeroes elsewhere.

Algebra will confirm that the pdf in (6) will yield the same expression as Eqs. (3) and (5). We also see that post-multiplying $(\mathbf{I} - \boldsymbol{\Omega})$ by the one vector yields the zero vector, indicating that the covariance matrix in (6) is

singular. Therefore, we present sufficient conditions for propriety.

In order to create a proper distribution, one can multiply Ω by a propriety parameter α . If $\alpha \in \left(\frac{1}{\lambda_{(1)}}, \frac{1}{\lambda_{(n)}}\right)$, where $\lambda_{(1)}$ and $\lambda_{(n)}$ are the minimum and maximum eigenvalues of $\Sigma^{-1/2}\Omega\Sigma^{1/2}$, then $\Sigma^{-1}(\mathbf{I} - \Omega)$ will be positive definite, and $\mathbf{X} \sim N(\boldsymbol{\mu}, (\mathbf{I} - \alpha\Omega)^{-1}\Sigma)$. We note here that $\lambda_{(1)}$ can be shown to be negative, and it is generally undesirable to allow α to take negative values. Therefore, we propose α be constrained to be in the set $(0, 1/\lambda_{(n)})$.

This result follows by completing the multiplication $\Sigma^{-1}(\mathbf{I} - \alpha\Omega) = \Sigma^{-1} - \alpha\Sigma^{-1}\Omega$ which is clearly symmetric. Pre and post-multiplying by $\Sigma^{1/2}$ yields $\mathbf{I} - \alpha\Sigma^{-1/2}\Omega\Sigma^{1/2}$. Consider the spectral decomposition of $\mathbf{I} - \alpha\Sigma^{-1/2}\Omega\Sigma^{1/2}$, which can be written as $P(\mathbf{I} - \alpha\Lambda)P'$, where P is an orthogonal matrix with columns being the eigenvectors of $\Sigma^{1/2}\Omega\Sigma^{-1/2}$ and Λ is a diagonal matrix of eigenvalues. Constraining $\alpha \in (\lambda_{(1)}, \lambda_{(n)})$ where $\lambda_{(1)} < \dots < \lambda_{(n)}$ are the ordered eigenvalues of $\Sigma^{-1/2}\Omega\Sigma^{1/2}$, makes the matrix $\mathbf{I} - \alpha\Lambda$ a strictly positive diagonal matrix. This indicates that $\Sigma^{-1}(\mathbf{I} - \alpha\Omega)$ is positive definite for these constraints on α .

This model can be used as an intrinsic model when $\min(\delta) \geq 0$, and this may be preferred by some practitioners. When $\min(\delta) < 0$, $\Sigma^{-1}(\mathbf{I} - \Omega)$ cannot be guaranteed to nonnegative definite, and a propriety parameter should be used. Additionally, if one assumes $\min(\delta) \geq 0$, the Ω matrix is row stochastic. Because the minimum and maximum eigenvalues of row stochastic matrices are greater than -1 and less than 1, respectively, any $\alpha \in (-1, 1)$ is sufficient to guarantee the existence of the inverse of $\Sigma^{-1}(\mathbf{I} - \alpha\Omega)$. This property can be exploited to simplify the sampling of α . If $\min(\delta) < 0$, this result does not necessarily hold. Additionally, we require the use of proper prior distributions for the vector δ . If improper prior distributions are used for any element of this parameter vector, an improper posterior distribution will result in the event that $\alpha = 0$.

2.2. Model properties

A set of natural parametric forms satisfying the conditions on σ^2 (the diagonal of Σ), β , and δ would be

$$\sigma_{ki}^2 = \frac{\sigma_k^2}{\left(\frac{w_{ki+}}{1 - \delta_{k++}}\right)}, \quad \beta_{kij} = \frac{w_{kij}}{\left(\frac{w_{ki+}}{1 - \delta_{k++}}\right)}, \quad \text{and} \quad \delta_{klj} = \frac{\sigma_{ki}^2}{\sigma_{lj}^2} \delta_{lkji}$$

where w_{kij} is an indicator that elements i and j are neighbors in neighborhood structure k , and the “+” indicates summing over the subscript. Additionally, this induces the constraint $\sum_i \sum_{j \neq ki} \delta_{klj} < 1$ for all i and k , in order to ensure that the conditional variances are all positive.

We consider these forms natural as they extend the weights of the Intrinsic CAR (ICAR) model (Besag et al., 1991). The requirement $\delta_{klj} = (\sigma_{ki}^2 / \sigma_{lj}^2) \delta_{lkji}$ is required by the model in order to yield a symmetric covariance matrix, but one will note the similarity between the parameterization σ_{ki}^2 and β_{kij} with the common ICAR structure

$$\sigma_{ki}^2 = \frac{\sigma_k^2}{w_{ki+}} \quad \text{and} \quad \beta_{kij} = \frac{w_{kij}}{w_{ki+}}.$$

It is clear that we have adjusted the ICAR weights by the contribution to the outcome of the other neighborhood structures. We have found this dimension reduction helpful for the MMCAR model. Allowing freedom in both the variances and weights is often too flexible, and may result in identifiability issues. When inference is of interest to a practitioner, this parameterization is useful.

Of course, these are not the only constraints that work for this model, but they do provide a natural interpretation and make sampling more efficient. It is also important to note that the structure on δ allows for information to be drawn asymmetrically across the different neighborhood structures. The constraints yield the desirable result that outcomes with higher variances will draw less from their neighboring outcomes and more from neighborhood structures with lower variances. In the case where δ_{k++} is positive, the division by $1 - \delta_{k++}$ implies that a fraction of the outcome's own neighborhood structure will be used in determining the outcome's mean. Therefore, the smoothing performed within the neighborhood structure is reduced, and information is drawn across neighborhood structures to perform additional smoothing. This is anticipated to be the typical case, but the positivity of δ_{k++} is not required by the model. If δ_{k++} is negative, the division by $1 - \delta_{k++}$ increases the smoothing performed within the neighborhood structure, resulting in spatial oversmoothing. The form of σ_{ki}^2 is a natural way to reduce the number of parameters in the model while still satisfying model constraints. From this point forward, we use this particular parametrization scheme.

The MMCAR structure is very flexible in modeling the spatial patterns across multiple outcomes. Consider the simple case of two outcomes and a single δ parameter. There are four cases of interest.

1. With α positive and δ near one, we have little spatial correlation within outcomes and strong positive correlation within each areal unit.
2. With α positive and δ zero, we have independent CAR models.
3. With α positive and δ negative, we have negative correlation within areal units and spatial over-smoothing.
4. With α zero, we have purely independent latent spatial effects.

It is worth noting the contribution of Wall (2004) when considering the propriety parameter α . While α is not required to be positive by the model, CAR models on irregular lattices may behave in unexpected ways for negative α values. Additionally, as α varies through $(0, 1)$, the CAR model does not preserve the ordering of correlations between pairs of outcomes. All of these results were obtained in the case where all the weights in the proximity matrix were positive. It is unclear how these results may change if the proximity matrix is allowed to have negative entries. Model properties are demonstrated via simulation in the [online supplementary materials](#).

The differences between the structure of our model and the model of Sain et al. (2011) are worth noting. Most notably, we have allowed more freedom in the variance terms. This was done with the thought of mismatched lattices in

mind. In the case of overlapping lattices, their requirement of variances that differ by a fixed and known quantity may be restrictive. Our model can handle matched lattices with this requirement removed, which may be useful when a researcher is not sure how the variances of the multiple outcomes may be related, or does not want to impose rigid structure on this particular set of parameters. Sain et al. (2011) do not require that $\sum_{kj \sim ki} \beta_{kij} + \sum_{l \neq k} \sum_{lj \sim ki} \delta_{klj} = 1$, instead implicitly requiring that $\sum_{kj \sim ki} \beta_{kij} + \sum_{l \neq k} \sum_{lj \sim ki} \delta_{klj} < 1$ due to a diagonal dominance requirement. This allows their model additional flexibility in the estimation of the cross-covariance structures. Our requirement is in one sense restrictive in that there is more structure imposed on the within and between neighborhood weights, but allows simple bounds to be found for the weights to guarantee a positive definite covariance structure, which some practitioners may prefer. It also develops the MMCAR in the structure of IGMRFs.

3. Data analysis I: Bayesian learning for infectious disease mapping

During the 2006 Mumps outbreak in Iowa, 77 counties reported a total of 1558 suspected cases of Mumps. A map of these data can be found in Fig. 1, along with a map showing which counties contain highways. The relationship between highways and cases in the western half of the state is less clear from this map than a spatio-temporal plot, but we intend to use this model as an exploratory tool. The MMCAR is computationally less expensive than typical infectious diseases models, such as Susceptible, Infectious, Removed (SIR) models. A major reason to perform an infectious disease map is to explore possible channels of disease spread before building these more computational models.

We use ten of the most traveled highways in Iowa as the generating structure (Interstates 80, 380, 29 and 35, as well as State Routes and US Highways 20, 30, 151, 60, 27, 61, and 34). These highways are only considered as part of the structure when they are multilane divided highways, as those stretches contain the most traffic. When two highways intersect, the longer highway is broken into two highway segments, with the argument being that

shorter stretches of highways are more likely to be accurately modeled with the same random effect, and allows an analysis of which stretches of highway carried the highest risk of infection. It will also allow for differences in risk to be associated with interstates versus state routes, which often carry different types of traffic. This structure then consists of every multilane highway in Iowa that passes through more than two counties, yielding a highway lattice consisting of 12 stretches of highway that pass through 53 counties. The structure in this example is mismatched in two ways. First, there are 46 counties which do not have a multilane highway, and these counties are deleted from the highway lattice structure entirely. Second, Each stretch of highway is associated with two or more counties. This implies that a single highway random effect will be associated with multiple counties in our model. Specifically, there are 12 random effects associated with highways, and these 12 random effects are each associated with one of the 53 counties that contain a multilane highway.

A previous generalized linear mixed model was constructed for these data, in part, to determine how the disease spread spatially and how that spread related to college towns. The analysis utilized county centroids to define spatial correlation, along with a concept developed therein defined as “zones of risk” (Polgreen et al., 2010). We wish to additionally consider the effect of highways in this process, as early cases seem to appear along highways as a primary means of spread. To assess this effect, we utilize the MMCAR structure when analyzing the final counts of Mumps in each county. This MMCAR analysis is then compared to standard CAR analyses via DIC (Spiegelhalter et al., 2002). Four Poisson regression models were constructed. Model One considers two counties to be neighbors if and only if a border was shared. Because counties with initial cases (Dubuque County and Johnson County) and counties with major universities (Black Hawk County, Johnson County, and Story County) tend to have larger infection counts, these counties are modeled using a different log-mean parameter. The data model we consider is:

$$Y_i = \exp(\mu_1 + \omega_i),$$

if county i does not contain an initial infection or a major university, and

$$Y_i = \exp(\mu_2 + \omega_i),$$

Fig. 1. Maps of the total county cases of Mumps (left) and the Iowa highway system (right).

if county i contains an initial infection or a major university.

Here, Y_i is the total number of confirmed cases, μ_1 and μ_2 are the log-means of the counts for each county type, and ω_i is a latent effect to account for the border structure. In our model, ω is given a $CAR(\alpha, \sigma^2)$ prior, with σ^2 given an $IG(0.01, 0.01)$ prior. Additionally, μ_1 and μ_2 were assigned $N(0, 10^2)$ priors for $i = 1, 2$, which are weakly informative at the scale of the parameters. To create a fair comparison to the MMCAR model, a propriety parameter was used in this model, and estimated. We assume a strong spatial relationship within these outcomes, so α assigned a $Beta(18, 2)$ prior for all of the CAR analyses, a prior which has successfully been used in the univariate CAR model (Carlin and Banerjee, 2003). A discussion of prior specifications for α can be found in the online supplement.

Model Two considers the highway structure as generating the neighborhood structure. The data model we consider is:

$$Y_i = \exp(\mu_1),$$

if county i does not contain an initial infection or a major university and additionally does not contain a highway,

$$Y_i = \exp(\mu_1 + \omega_j),$$

if county i does not contain an initial infection or a major university but does contain major highway j , and

$$Y_i = \exp(\mu_2 + \omega_j),$$

if county i contains an initial infection or a major university. The priors are the same as in Model One.

Model Three is the MMCAR model. Two neighborhood structures are defined. The first is the same used in Model One, defined by shared borders. The second considered the highway structure in the state of Iowa, where

$$Y_i = \exp(\mu_1 + \omega_{1,i} + \omega_{2,j}),$$

if county i does not contain an initial infection or a major university, but contains the major highway j ,

$$Y_i = \exp(\mu_1 + \omega_{1,i}),$$

if county i does not contain an initial infection or a major university and does not contain a major highway, and

$$Y_i = \exp(\mu_2 + \omega_{1,i} + \omega_{2,i}),$$

if county i contains an initial infection or a major university (all of which contain major highways). Again, Y_i is the total number of confirmed cases, and μ_1 and μ_2 are the log-means of the counts for each county type. In the MMCAR model, $\omega_{1,i}$ corresponds to a latent spatial effect generated by the border structure, and $\omega_{2,j}$ corresponds to a latent spatial effect generated by the highway structure, and associated with highway j .

In this model $(\omega_1', \omega_2')'$ is given an $MMCAR(\sigma_1^2, \sigma_2^2, \delta, \alpha)$ prior, with σ_1^2 corresponding to the border structure and given an $IG(0.01, 0.01)$ prior, σ_2^2 corresponding to the highway structure and given an $IG(0.01, 0.01)$ prior, μ_1 and μ_2 given $N(0, 10^2)$ priors for $i = 1, 2$, α given a $Unif(0, 1)$ prior, and a single δ parameter. This δ parameter is chosen in such a way that the association from the border structure to the highway structure is constant for every highway

segment, an assumption chosen to facilitate modeling δ . The prior on δ is $Uniform(-2, 0.125)$, where 0.125 is selected to yield the full positive support of δ given the model structure. This support follows because we assume that δ is constant in its relationship from county to highway coupled with the fact that the model requires $\sum_i \sum_{j \sim k_i} \delta_{klij} < 1$ and the longest stretch of highway covers eight counties. Therefore, for at least one stretch of highway, $\delta \geq 1/8$ will violate the assumption $\sum_i \sum_{j \sim k_i} \delta_{klij} < 1$. The lower bound of -2 has been chosen to guarantee the propriety of the posterior of δ and is believed to be small enough that the model is unlikely to select the strong spatial over-smoothing associated with such small δ values. For the MMCAR model, α is given a $Unif(0, 1)$ prior distribution in order to allow more flexibility in estimating δ .

Model Four is a model similar to the MMCAR structure, but uses two independent CAR structures. The prior specifications are identical to Model Three, but there is no δ parameter present in these data.

One might question whether this example is suitable to a multivariate analysis, given that it is a Bayesian learning example. From the data, it is suspected that there are two latent spatial structures at work. Thinking of Mumps in terms of physical transfer, certainly Mumps can travel from county to county across borders. However, the data demonstrate longer range dependence when graphed over time. The MMCAR model can be used to test the importance of hypothesized channels of spread before undertaking the building of more complex and computationally expensive infectious disease models. Because the lattices here are mismatched, there will be multiple counties associated with each highway random effect. In this way there is sufficient information to distinguish between the two structures, and so this modeling framework will take advantage of both lattice structures.

For each model, a single chain was run following the sampling scheme in the online appendix. A burn-in period of 50,000 samples was discarded and then 350,000 samples were collected from Model One. A burn-in period of 50,000 iterations was also used for Model Two, with 250,000 iterations being collected after burn-in for analysis. Model Three required a burn-in of 250,000 iterations and 550,000 iterations were collected from the converged chains. Model Four required a 300,000 iteration burn in period and then 900,000 iterations were selected from the converged chain. The differences in the lengths of the chains are due to differences in model complexity. Convergence was assessed via the Geweke diagnostic (Geweke, 1992), with p -values greater than 0.05 considered indicating convergence for the risk surface and variance parameters, which are the quantities of interest for inference. As is common in CAR modeling, mixing was slow between the model intercept and the mean of the random effects (Rue and Held, 2005). However, we do not wish to draw inference on these quantities, and all quantities of interest pass convergence diagnostics. Table 1 provides parameter posteriors for all model parameters. We see that the means and variances are very similar for Models One and Three, with the propriety parameter being lower for model Three. Model Four provides similar means, but higher latent effect variances. The propriety parameters of Model Three are more akin to those

Table 1

Parameter medians and 95% credible intervals and DIC for Model 1 (CAR model only accounting for only the border structure), Model 2 (CAR model only accounting for the highway structure), Model 3 (MMCAR model accounting for both the border and highway structures), and Model 4 (two independent CAR models accounting for both the border and highway structures).

Parameter	Model 1	Model 2	Model 3	Model 4
μ_1	1.480 (1.372, 1.585)	1.637 (1.561, 1.710)	1.089 (0.516, 1.821)	1.210 (1.083, 1.336)
μ_2	4.573 (4.352, 4.778)	4.374 (4.307, 4.441)	3.992 (2.780, 4.796)	4.153 (3.595, 4.685)
σ_1^2	3.961 (2.768, 5.587)	N/A	7.630 (4.826, 13.044)	8.833 (5.442, 11.095)
σ_2^2	N/A	6.432 (1.974, 19.357)	0.738 (0.245, 2.810)	3.474 (1.311, 8.479)
α_1	0.897 (0.758, 0.982)	0.891 (0.726, 0.985)	0.837 (0.468, 0.979)	0.900 (0.739, 0.987)
α_2	N/A	N/A	N/A	0.900 (0.739, 0.986)
δ	N/A	N/A	0.010 (-0.052, 0.080)	N/A
DIC	554.361	1222.961	471.600	465.325

Fig. 2. Posteriors for the covariance parameters in the Mumps analysis.

of Model One. Fig. 2 provides posterior densities for the MMCAR model and illuminates the issue of the estimation of the propriety parameter. The model suggests an independence of the highway and border effects.

The DICs of both models containing both sets of latent effects are lower than the standard CAR model used for chronic disease mapping, which is much lower than that of the model containing only a highway effect. This indicates that the model is able to determine the importance of both channels of infection in mapping the spread of Mumps. Additionally, it lends credence to the idea of using the MMCAR here, despite having only one outcome informing two sets of latent effects. However, the DIC of Model Four is lower than that of Model Three, the MMCAR

model. This may be due to independence in the highway and border structures. It is worth noting that previous work has demonstrated that DIC does not always select the best model in terms of risk surface prediction when performing disease mapping (e.g., Best et al., 2005; Hossain et al., 2014). For our current example, we additionally compared Models Three and Four using Binder's Loss (Binder, 1978), and found that the selections agreed with the selections of DIC for this dataset.

Despite the two CARs model being preferred to the MMCAR model, we have demonstrated that the MMCAR can be effectively used to model data on mismatched lattices. Additionally, the development of infectious disease mapping across multiple channels is an important one.

Table 2

Parameter medians and 95% central credible intervals for the MMCAR model (Model 1), and the two independent CARs model (Model 2) of poverty and high school graduation (and equivalency) data.

Parameter	Model 1	Model 2
μ_1	95.85 (94.90, 96.78)	95.86 (94.93, 96.80)
μ_2	144.72 (1367.29, 151.59)	143.22 (139.41, 147.06)
σ_1^2	6256.44 (5235.37, 7700.85)	4981.85 (4433.85, 5629.52)
σ_2^2	768.33 (363.71, 2079.95)	1314.74 (513.41, 3668.39)
σ_μ^2	8907.68 (2658.92, 59509.26)	8836.10 (2711.23, 56540.77)
δ	0.0075 (0.0062, 0.0081)	N/A
DIC	8953.03	8965.51

4. Data analysis II: The MMCAR model for a multivariate outcome in the ACS

The American Community Survey (ACS) is a survey conducted by the United States Census Bureau in order to collect demographic information at intervals finer than the decennial census. The ACS provides one-year and three-year period estimates for locations with populations of at least 65,000 and 20,000 individuals, and provides five-year period estimates at every scale of geography in the United States, down to the census tract level. The demographic information contained within the survey is often useful in disease mapping, and an analysis of these data will show

the broad applicability of the MMCAR model. We demonstrate the ability of this model to detect multiple cross correlation parameters by performing a multivariate analysis of poverty and educational attainment in the Midwest. Specifically, we are interested in accurately modeling the percent of families in poverty at the county-level and the percent of individuals 25 and older who do not have a high school education or equivalent in North Dakota, South Dakota, Kansas, Iowa, Missouri, Illinois, Minnesota, Indiana, Ohio, Wisconsin, and Michigan. Both outcomes represent five-year period estimates ending in 2012.

Accurate county-level estimates of poverty are important in a wide range of public policy and public health applications, and state-level estimates of high school education rates are highly variable in the ACS even at the state-level. Using these two related outcomes, collected on different lattices, we are able to refine the estimates of both for use in public policy and public health policy making. We do this by means of a regression model similar to the celebrated Fay-Herriot model (Fay and Herriot, 1979) for small area estimation, but here we consider a multivariate analog without auxiliary information. The actual model we consider is

$$Y_{1i} = \mu_1 + u_{1i} + \epsilon_{1i},$$

$$Y_{2j} = \mu_2 + u_{2j} + \epsilon_{2j}.$$

Here, Y_{1i} is the percentage of families in poverty in county i of the 983 county lattice (multiplied by 10 for numerical stability) and Y_{2j} is the percentage of individuals without

Fig. 3. Plots of the parameter posteriors for the variance components of the ACS example. Each plot is based on 10,000 iterations after 1,000 iteration burn-in period.

a high school education in state j of the 11 state lattice (multiplied by 10 for numerical stability). The $\epsilon_{1i} \sim N(0, \sigma_{1i}^2)$, where σ_{1i}^2 is the known sampling variance (multiplied by 100) of the ACS estimate of the percent of families in poverty in county i , and $\epsilon_{2j} \sim N(0, \sigma_{2j}^2)$ is the known sampling variance (multiplied by 100) of the ACS estimate of the percent of individuals without a high school education in state j . Finally $(u_{1,1}, \dots, u_{1,983}, u_{2,1}, \dots, u_{2,11})'$ is given an Intrinsic MMCAR prior MMCAR $(\sigma_1^2, \sigma_2^2, \delta)$ where we use a single δ parameter, assumed constant from the county to the state level for computational efficiency. We use an intrinsic MMCAR because a proper MMCAR provided estimates of α that were approximately one, leading to numerical instability.

The priors selected for the analysis were $\sigma_1^2 \sim IG(0.01, 0.01)$, $\sigma_2^2 \sim IG(0.01, 0.01)$, $\mu_1 \sim N(0, \sigma_\mu^2)$, $\mu_2 \sim N(0, \sigma_\mu^2)$, $\sigma_\mu^2 \sim IG(0.01, 0.01)$ and $\delta \sim Unif(0, 1/115)$. The prior for δ must have a lower bound of zero due to the intrinsic nature of the MMCAR model selected, and has an upper bound of $1/115$ because the state of Missouri has 115 counties, which is the maximum of all of the selected states, and the structure of the model requires $115\delta < 1$. This may seem to be a restrictive prior, but the posterior distribution of δ is quite different from the uniform prior selected, and suggests strong between-lattice association. We compare the model to an equivalent model with independent intrinsic conditional autoregressive (ICAR) priors on the vectors $(u_{1,1}, \dots, u_{1,983})' \sim ICAR(\sigma_1^2)$ and $(u_{2,1}, \dots, u_{2,11})' \sim ICAR(\sigma_2^2)$. The priors on the parameters of this model are equivalent to those in the MMCAR model.

The parameter posteriors summaries can be found in Table 2 and the densities of the variance components are plotted in Fig. 3. The summary statistics are based on 10,000 MCMC iterations after 1000 iterations of burn-in, and convergence is again assessed by the Geweke diagnostic. We see that the mean structure of the two models is comparable, but that the variance structure induced by the random effects is very different between the two models. Notably, δ is concentrated near the maximum of its range, indicating strong between-lattice association. The posteriors would suggest that the random effects associated with the poverty estimates account for approximately 40–90% of the estimates of the state-level high school education estimates, depending on the state and the MCMC iteration, and that the random effects associated with high school education estimates account for approximately 6% of the random effects for each county in the lattice. Finally, the DICs of both the MMCAR and the two CARs model clearly demonstrate improved model fit when employing the MMCAR. We conclude that the intrinsic MMCAR provides a superior fit than independent ICAR models for these data.

5. Discussion

In this paper, we have developed a Gaussian conditional autoregressive model allowing for correlated multivariate outcomes on mismatched lattices. We have extended the structure of Sain et al. (2011), that allows for asymmetric correlation information to be used between lattices by allowing mismatched lattices and by considering a formulation which allows more flexibility in the relative

variances of outcomes, a desirable extension when latent spatial effects are considered. To reduce variability in the parameter posteriors, we have provided a recommended parameterization and sampling scheme for our model.

This model will prove useful in modeling infectious disease counts, especially in cases where there are multiple modes of travel or contact. The model is computationally less expensive than spatial SIR models, and so it can be used to test the significance of different channels of disease spread before their inclusion in more complex models. We have demonstrated a highway effect in the disease counts of the Iowa epidemic, and suggest this reiterates the importance of considering multiple contact pathways when modeling infectious diseases. It seems that the highway structure may be an important proxy for disease spread in rural areas. Further studies may confirm this.

The MMCAR was developed with mismatched lattices in mind, but in developing this model, we have relaxed assumptions of the multivariate CAR models already in the literature. The MMCAR allows for new inferences to be performed even on matched lattices. In particular, the ability to have unique variances for each outcome while still allowing some flexibility in the cross correlations is new and an important contribution to the multivariate CAR literature. Current models allow either flexibility in the variances, or flexibility in the cross correlations, but not both. The ACS example we provide is just one case in which the type of flexibility provided by our MMCAR parameterization is desirable.

The MMCAR prior provides improved model fit over two independent ICAR priors in the ACS example, indicating the importance of flexible cross correlation for this data example. Spatial statistics in the field of surveys and public policy is a rapidly growing field, and public health studies that are concerned with intervention can benefit from improved fitting of data like the data we consider here.

Acknowledgements

We wish to thank two anonymous referees for their constructive comments that have led to a better manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.sste.2014.08.001>.

References

- Banerjee S, Carlin BP, Gelfand AE. Hierarchical modeling and analysis for spatial data. Chapman & Hall; 2004.
- Besag J. Spatial interaction and the statistical analysis of lattice systems. J R Stat Soc Ser B 1974;36(2):192–236.
- Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics (with discussion). Ann Inst Stat Math 1991;43:1–59.
- Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. Stat Methods Med Res 2005;14:35–59.
- Binder DA. Bayesian cluster analysis. Biometrika 1978;65:31–8.
- Carlin BP, Banerjee S. Hierarchical multivariate CAR models for spatio-temporally correlated survival data. Bayesian Stat 2003;7:45–64.

- Chellappa R, Chatterjee S. Classification of textures using Gaussian markov random fields. *IEEE Trans Acoust Speech Signal Process* 1985;33(4):959–63.
- Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987;43(3):671–81.
- Cressie NAC. *Statistics for spatial data*. Wiley; 1993 [Revised Edition].
- Fay R, Herriot R. Estimates of income for small places: an application of James-Stein procedures to census data. *J Am Stat Assoc* 1979;74:269–77.
- Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* 2003;4(1):11–25.
- Geweke J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. *Bayesian Statistics 4*. Oxford, UK: Clarendon Press; 1992.
- Hossain MM, Lawson AB, Cai B, Choi J, Liu J, Kirby RS. Space-time areal mixture model: relabeling algorithm and model selection issues. *Environmetrics* 2014. <http://dx.doi.org/10.1002/env.2265>.
- Hufnagel L, Brockmann D, Geisel T, May R. Forecast and control of epidemics in a globalized world. *PNAS* 2004;101(42):15124–9.
- Jin X, Carlin B, Banerjee S. Generalized hierarchical multivariate car models for areal data. *Biometrics* 2005;61:950–61.
- Kim H, Oleson JJ. A bayesian dynamic spatio-temporal interaction model: an application to prostate cancer incidence. *Geogr Anal* 2008;40:77–96.
- Kim H, Sun D, Tsutakawa RK. A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model. *J Am Stat Assoc* 2001;96(456):1506–22.
- Lichstein JW, Simons TR, Shiner SA, Franzreb KE. Spatial autocorrelation and autoregressive models in ecology. *Ecol Monogr* 2002;72(3):445–63.
- Mardia K. Multi-dimensional multivariate gaussian markov random fields with application to image processing. *J Multivar Anal* 1988;24:29–56.
- Polgreen P, Bohnett L, Yang M, Pentella M, Cavanaugh J. A spatial analysis of the spread of mumps: the importance of college students and their spring break associated travel. *Epidemiol Infect* 2010;138:434–41.
- Rue H, Held L. *Gaussian Markov random fields: theory and applications*. CRC Press; 2005.
- Sain SR, Cressie N. Multivariate lattice models for spatial environmental data. *Proc. Statist. Environ. Sect. Am. Statist. Ass* 2002;2820–5.
- Sain SR, Furrer R, Cressie N. A spatial analysis of multivariate output from regional climate models. *Ann Appl Stat* 2011;5(1):150–75.
- Spiegelhalter D, Best NG, Carlin B, Van der Linde A. Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Ser B* 2002;64:583–616.
- Sun D, Tsutakawa R, Kim H, He Z. Spatio-temporal interaction with disease mapping. *Stat Med* 2000;19:2015–35.
- Wall MM. A close look at the spatial structure implied by the CAR and SAR models. *J Stat Plan Infer* 2004;121(2):311–24.