UCLA
Research Reports

Title
Heteroscedastic CAR models for areally referenced temporal processes for analyzing California asthma hospitalization data

Permalink
https://escholarship.org/uc/item/4rj1t11c

Authors
Quick, Harrison
Carlin, Bradley P.
Banerjee, Sudipto

Publication Date
2015-01-28

Peer reviewed

eScholarship.org

Powered by the California Digital Library
University of California

Heteroscedastic CAR models for areally referenced temporal processes for analyzing California asthma hospitalization data

Harrison Quick*[1], **Bradley P. Carlin**[2], **and Sudipto Banerjee**[3]

[1] Division of Heart Disease and Stroke Prevention, Centers for Disease Control and Prevention, Atlanta, GA 30329

[2] Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota 55455

[3] Department of Biostatistics, University of California, Los Angeles, CA 90095-1772.

* Corresponding author *email:* HQuick@cdc.gov

Summary. Often in regionally aggregated spatiotemporal models, a single variance parameter is used to capture variability in the spatial structure of the model, ignoring the impact that spatially-varying factors may have on the variability in the underlying process. We extend existing methodologies to allow for region-specific variance components in our analysis of monthly asthma hospitalization rates in California counties, introducing a heteroscedastic CAR model that can greatly improve the fit of our spatiotemporal process. After demonstrating the effectiveness of our new model via simulation, we reanalyze the asthma hospitalization data and note a number of important findings.

Key words: Bayesian methods, Gaussian process, Gradients, Markov chain Monte Carlo, Spatial process models

1 Introduction

Space-time models can be classified as considering one of the following four settings: (a) space is viewed as continuous, but time is taken to be discrete, (b) space and time are both continuous, (c) space and time are both discrete, and (d) space is viewed as discrete, but time is taken to be continuous. Almost exclusively, the existing literature considers the first three settings; see the recent books by Banerjee et al. (2015) or Cressie and Wikle (2011) and references therein for an excellent review of existing approaches falling within the first three categories. MacNab and Gustafson (2007) and Ugarte et al. (2010) have considered spline-based methods for discrete-space, continuous time data but have not focused upon temporal predictions at fine resolutions. Baladandayuthapani et al. (2008) considered spatially correlated functional data modeling but treating space as continuous. We agree with Delicado et al. (2010) that spatially associated functional modeling of time has received little attention, especially for regionally aggregated data.

Recently, Quick et al. (2013) departed from the three common settings (a)–(c) by proffering Bayesian space-time models in case (d), based upon a dynamic Markov random field (MRF, or more specifically a conditional autoregression, or CAR) model that evolves continuously over time while treating space as a finite set of regions. In addition to permitting inference at a temporal resolution finer than that at which the data were sampled, they developed a methodology for conducting inference on infinitesimal rates of change (i.e., temporal gradients) at arbitrary time points for each county. Such inference provides an understanding of the local effects of temporal impact in a way that is precluded by discrete-time models.

Despite recent advances, the size of space-time datasets encountered today often presents computational challenges. For a dataset with N observations, estimating models specified by general (non-sparse) space-time covariance matrices will involve storing and factorizing or decomposing $N \times N$ matrices, thereby rendering them computationally infeasible. How we model dependencies, therefore, is critical for practical implementation of space-time models. By using a simple *separable* model that separates the spatial and temporal associations, one can achieve computational gains from the resulting Kronecker product form of the space-time covariance matrix. This structure, however, limits the model to only one variance parameter for controlling both spatial and temporal variability. Unfortunately, if the true underlying process is less smooth in some regions, or if there are spatial outliers, such a model may both oversmooth *and* undersmooth the space-time effects; i.e., we may oversmooth regions with extreme values while allowing too much variability in the more moderate regions.

Our current manuscript enriches continuous-time dynamic CAR models with region-specific variance components for a more thorough analysis of the asthma hospitalization dataset described in Section 2. Section 3 then provides background on dynamic CAR models and the methodological motivation for the heteroscedastic CAR (HCAR) model we propose in Section 4. We then conduct two simulation studies in Section 5, in which we investigate the ability of our model to estimate the parent and gradient processes and important model parameters. Our analysis of the California asthma hospitalization rate data using the HCAR follows in Section 6, in which our model sheds light on features previously overlooked. Finally, Section 7 summarizes our findings and concludes.

2 Asthma Hospitalization Data

In this paper, we reanalyze the data from Quick et al. (2013). These data consist of $N_t = 216$ monthly asthma hospitalization rates collected by the California Health and Human Services Agency from 1991 to 2008 in the $N_s = 58$ counties of California. The hospitalization rates were based on all hospital discharges where asthma was listed as the primary diagnosis, and rates per 100,000 residents were computed. Summary maps of these data can be found in

Figure 1. More information regarding these data can be found in Delamater et al. (2012).

The explanatory variables used in our current analysis include ozone level, population density, percent of the population that was black, and the percent of the population that was under the age of 18. While the demographic covariates are based on the 2000 Census and only vary spatially, our ozone covariate represents the number of days each month with average ozone levels above 0.07 ppm over 8 consecutive hours, the state standard. Unfortunately, these measurements are aggregated at the *air basin* level, which cover large areas with similar weather and geographic conditions; as a result, air basins often span across multiple counties. As asthma rates are known to vary seasonally, we will nest the effect of ozone within month and include monthly fixed effects.

3 Dynamic CAR models

Consider a map comprising N_s regions that are delineated by well-defined boundaries, and let $Y_i(t)$ be the outcome arising from region i at time t. While the region-specific outcome $Y_i(t)$ is conceptualized as a continuous function of time, the observations are collected at a finite collection of distinct time points $\mathcal{T} = \{t_1, t_2, \ldots, t_{N_t}\}$. In this work, we will assume that the data come from the same set of time points in $\mathcal{T}$ for each region — an assumption that is not necessary for further development but will facilitate the notation.

The observation from region i at time t_j is modeled by a space-time regression model

$$Y_i(t_j) = \mu_i(t_j) + Z_i(t_j) + \epsilon_i(t_j), \quad \epsilon_i(t_j) \overset{ind}{\sim} N(0, \tau_i^2), \tag{1}$$

for $i = 1, 2, \ldots, N_s$ and $j = 1, 2, \ldots, N_t$, where $\mu_i(t_j) = \mathbf{x}_i(t_j)^T \boldsymbol{\beta}$ captures large scale variation or trends, $\mathbf{x}_i(t_j)$ is a vector of explanatory variables observed at the county level for each timepoint, $\boldsymbol{\beta}$ is the corresponding vector of regression slopes, $Z_i(t_j)$ is the space-time random effect arising from an areally-referenced stochastic process over time, $Z(t)$, that captures smaller-scale variations in the time scale while also accommodating spatial associations, and τ_i^2 captures any residual variation not captured by the other components for region i.

The process $Z_i(t)$ can be looked upon as a random function of time that specifies the probability distribution of correlated space-time random effects. We seek a specification that permits the functions $Z_i(t)$ and $Z_k(t)$ from neighboring regions to be more similar than those from non-neighbors. We achieve this by constructing a joint distribution for the entire collection of $Z_i(t_j)$'s. If we collect the $Z_i(t_j)$'s for all the regions into an $N_s \times 1$ vector function $\mathbf{Z}(t_j)$ for any timepoint t_j and then stack them into an $N_sN_t \times 1$ column vector $\mathbf{Z} = (\mathbf{Z}(t_1)^T, \mathbf{Z}(t_2)^T, \ldots, \mathbf{Z}(t_{N_t})^T)^T$, then the distribution of $\mathbf{Z}$ is given by

$$\mathbf{Z} \sim N(\mathbf{0}, R(\phi) \otimes \sigma^2(D - \alpha W)^{-1}) \; , \tag{2}$$

where $\sigma^2(D - \alpha W)^{-1}$ is the covariance matrix for the proper CAR model with a 0/1 adjacency matrix W, and a diagonal matrix D having ith diagonal element equal to the number of neighbors for the ith region, n_i. A sufficient condition for the precision matrix to be invertible is $\alpha \in (0, 1)$, which ensures a proper distribution for $\mathbf{Z}(t)$ at each timepoint t. The matrix $R(\phi)$ is an $N_t \times N_t$ correlation matrix with (j, j')th element

$$\rho(t_j, t_{j'}; \phi) = (1 + \phi|t_{j'} - t_j|) \times \exp(-\phi|t_{j'} - t_j|) \; , \tag{3}$$

which corresponds to the correlation matrix for a temporal Gaussian process with a Matérn correlation function with smoothness parameter 3/2, denoted Matérn(3/2). The correlation function in (3) ensures that the $Z_i(t)$'s are mean-square differentiable functions of t, legitimizing inference on infinitesimal rates of temporal change (Quick et al., 2013). Henceforth, we refer to the spatiotemporal CAR model induced by (2) and (3) as the CAR$_{ST}$ model.

A limitation of the CAR$_{ST}$ model in (2) is the presence of a single variance parameter σ^2 to capture the scale of temporal and spatial variations. The diagonal elements in $(D - \alpha W)^{-1}$ depend upon the adjacencies in the map and adjust σ^2 accordingly to offer region-specific marginal variances. In some settings, however, such as when a particular county exhibits a trend which is markedly different from its neighbors, this model can prove to be too restrictive. In this paper, we propose to relax this assumption by extending the CAR$_{ST}$ to allow for region-specific variance parameters, σ_i^2. In doing so, we permit these so-called "outlying" counties the flexibility they need to break free from their neighbors.

The issue of differing levels of variability in areal spatial settings is a topic that has been addressed before. Two primary examples include Lawson and Clark (2002), who split the spatial process into a mixture of a CAR component (L_2) and an L_1 process, and Brewer and Nolan (2007), who developed an Empirical Bayes, pairwise additive approach of the form

$$\pi(Z_i(t) \,|\, Z_{(i)}(t)) \propto \exp\left[-\frac{1}{2}\sum_{k\sim i}\frac{\left(Z_i(t) - Z_k(t)\right)^2}{\sigma_{ik}^2}\right] \tag{4}$$

where the pairwise-defined $\sigma_{ik}^2 = \sigma_i^2 + \sigma_k^2$ have replaced σ^2 in the standard CAR model and $Z_{(i)}(t)$ denotes the vector $\mathbf{Z}(t)$ with the ith element removed.

Reich and Hodges (2008) take a similar approach to Brewer and Nolan (2007) with their fully Bayesian, spatially adaptive CAR (SACAR) model, in which they assume $\sigma_{ik}^2 = \sigma_i\sigma_k$, where $\sigma_i^2 = \exp(s_0 + s_i)$ and $\mathbf{s} = (s_1, \ldots, s_{N_s})' \sim \text{CAR}(\lambda)$ in order to ensure the identifiability of the parameters; that is, the prior distribution of the σ_i^2 *assumes* a spatial structure. In this approach, the s_i are constrained such that $\sum_i s_i = 0$ and s_0 is an intercept term to be estimated. These techniques are based on the improper CAR model (which sets $\alpha = 1$) and designed for purely spatial models, but extending (2) to allow for the proper analogs of these structures appears straightforward.

4 The heteroscedastic CAR (HCAR) model

To remedy the situation discussed in Section 3, we now allocate a different variance component σ_i^2 to each region. Rather than assume $\mathbf{Z}(t) \sim \mathrm{CAR}(\sigma^2)$ for each time point, we let $Z_i(t) = \sum_{k=1}^{N_s} a_{ik}(\alpha) v_k(t)$, where the $v_k(t)$'s are a set of N_s i.i.d. temporal Gaussian processes and $a_{ik}(\alpha)$'s are the associated coefficients. Specifically, if we let $v_k(t) \sim GP(0, \sigma_j^2 \rho(\cdot; \phi))$ — admitting a different variance σ_k^2 for each county — and define $a_{ik}(\alpha)$ to be the (i,k)-element of the matrix $A(\alpha)$ where $A(\alpha)A(\alpha)^T = (D - \alpha W)^{-1}$ (i.e., we define $A(\alpha)$ as the Cholesky decomposition of $(D - \alpha W)^{-1}$), it can be shown that this yields

$$\mathbf{Z} \sim N(\mathbf{0}, R(\phi) \otimes \Sigma_S (D - \alpha W)^{-1} \Sigma_S) \ , \tag{5}$$

where Σ_S is an $N_s \times N_s$ diagonal matrix with σ_i as its ith diagonal element. In contrast to the aforementioned approaches, the conditional distribution for a given t takes the form

$$\pi(Z_i(t) \,|\, Z_{(i)}(t)) \propto \exp\left[-\frac{1}{2} \sum_{k \sim i} \left(\frac{Z_i(t)}{\sigma_i} - \alpha \frac{Z_k(t)}{\sigma_k} \right)^2 \right] , \tag{6}$$

which we denote $\mathbf{Z}(t) \sim \mathrm{HCAR}(\alpha, \boldsymbol{\sigma})$, letting $\mathbf{Z} \sim \mathrm{HCAR}_{ST}(\alpha, \phi, \sigma_1, \ldots, \sigma_{N_t})$ denote the expression in (5). In this structure, the σ_i can be directly viewed as scaling parameters for their respective Z_i, rather than as components of a pairwise structure. A comparison between the CAR, the SACAR, and the HCAR is presented in Table 1.

When it comes to modeling the σ_i, we use an approach similar to that used by Reich and Hodges (2008). We let $\sigma_i = \exp(u_0 + u_i)$ where u_0 represents a baseline for our spatiotemporal variance and the u_i are region-specific adjustments with the constraint that $\sum_i u_i = 0$. Unlike the SACAR model, however, we do not assume a spatial correlation structure on the u_i, as doing so would restrict our model's flexibility for fitting outlying regions. Instead, we assume

$u_i \sim N(0, \gamma^2)$; we then place inverse gamma priors on both $\sigma_0^2 = \exp(2u_0)$ and γ^2. Based on our experience, updating σ_0^2 separately in this fashion speeds up the convergence and increases the stability of our Markov chain Monte Carlo (MCMC) algorithm.

For the remaining parameters, more conventional priors are suitable. Specifically, we place non-informative, conjugate priors on the regression coefficients $\boldsymbol{\beta}$, and assume the τ_i^2 follow inverse gamma priors with mean 1 and infinite variance. Conjugate priors do not exist for our spatial association parameter α or our temporal association parameter ϕ; as such, we have chosen a beta prior with mean 0.9 and an infinite peak at 1 for α, and a Uniform$(3/(N_t - 1), 10)$ prior for ϕ. These bounds for ϕ are intended for N_t equally-spaced time points and are based on an exponential correlation function. Specifically, this lower bound restricts the effective range of temporal correlation to the length of the study period (i.e., $\text{Cov}_{\exp}\left(Z_i(1), Z_i(N_t)\right) = \exp\left[-\phi(N_t - 1)\right] \leq 0.05 \approx \exp[-3]$), and this upper bound for ϕ corresponds to near 0 correlation at $|t_{j+1} - t_j| = 1$ using (3); typically, these bounds will be much wider than necessary. For the HCAR$_{ST}$ model presented in (5), the joint posterior distribution for our parameters is given by:

$$\begin{aligned} p(\boldsymbol{\theta}, \mathbf{Z} \,|\, \mathbf{Y}) \propto & N(\boldsymbol{\beta} \,|\, \mu_\beta, \Sigma_\beta) \times \prod_{i=1}^{N_s} \left[N(u_i \,|\, 0, \gamma^2) \times IG(\tau_i^2 \,|\, a_\tau, b_\tau) \right] \\ & \times IG(\sigma_0^2 \,|\, a_\sigma, b_\sigma) \times IG(\gamma^2 \,|\, a_\gamma, b_\gamma) \times U(\phi \,|\, a_\phi, b_\phi) \\ & \times Beta(\alpha \,|\, a_\alpha, b_\alpha) \times \text{HCAR}_{ST}(\alpha, \phi, \sigma_1, \ldots, \sigma_{N_t}) \\ & \times \prod_{j=1}^{N_t} \prod_{i=1}^{N_s} N(Y_i(t_j) \,|\, \mathbf{x}_i(t_j)^T \boldsymbol{\beta} + Z_i(t_j), \tau_i^2), \end{aligned} \tag{7}$$

where we are letting $\mu_i(t) = \mathbf{x}_i(t)'\boldsymbol{\beta}$ and $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \phi, \alpha, \sigma_0^2, u_1, \ldots, u_{N_s}, \gamma^2, \tau_1^2, \ldots, \tau_{N_s}^2\}$. We will use MCMC to evaluate (7), using Metropolis steps for updating $u_1, \ldots, u_{N_s}$, ϕ, and α, and Gibbs steps for all other parameters.

4.1 Modeling temporal gradients

In addition to providing a good fit to the data, our model has been specified in order to permit inference on the temporal gradient process. First developed for two-dimensional continuous space case in Banerjee et al. (2003), the gradient process can be used to detect sudden changes in the residual surface. Using real estate prices in Baton Rouge, LA, the authors show that significant gradients can be used to indicate important predictors (such as proximity to popular shopping centers) still missing from the mean model. Here, we are interested in temporal gradients, which may correspond to features such as sudden temporal changes in weather or public health policy that can affect asthma hospitalization rates.

The development of our temporal gradient process is similar to that in Quick et al. (2013). As the temporal correlation structure of our model remains a Matérn(3/2), inference for the temporal gradient process, $\mathbf{Z}'$, simply requires substituting $\Sigma_S(D-\alpha W)^{-1}\Sigma_S$ for $\sigma^2(D-\alpha W)^{-1}$ in the covariance structure for $\mathbf{Z}$. The result is $\Sigma_Z = R(\phi)\otimes\Sigma_S(D-\alpha W)^{-1}\Sigma_S$ and the expressions for the first and second order derivatives for the matrix-valued covariance function can be derived as

$$K'_Z(\Delta) = -\phi^2\Delta\exp(-\phi|\Delta|)\left(\Sigma_S(D-\alpha W)^{-1}\Sigma_S\right) \tag{8}$$

and $K''_Z(0) = -\phi^2\left(\Sigma_S(D-\alpha W)^{-1}\Sigma_S\right)$, respectively, where Δ is some temporal distance, say $(t_{j'}-t_j)$. Then the conditional distribution for the gradients, $\mathbf{Z}'$, are found to be multivariate normal with mean and variance-covariance matrix given by

$$\begin{aligned} \boldsymbol{\mu}_{Z'\,|\,Z,\theta} &= \text{cov}(\mathbf{Z}'(t_0),\mathbf{Z})\text{var}(\mathbf{Z})^{-1}\mathbf{Z} = -(K'_Z)^T\Sigma_Z^{-1}\mathbf{Z} \\ \text{and } \Sigma_{Z'\,|\,Z,\theta} &= -K''_Z(0) - (K'_Z)^T\Sigma_Z^{-1}(K'_Z)\ , \end{aligned}$$

where $(K'_Z)^T$ is an $N_s \times N_s N_t$ block matrix whose j-th block is given by the $N_s \times N_s$ matrix $K'_Z(\Delta_{0j})$, with $\Delta_{0j} = t_j - t_0$. Note that Σ_Z is an $N_s N_t \times N_s N_t$ matrix, but we can use the properties of the MRF to restrict to inverting only $N_t \times N_t$ matrices.

4.2 Diagnostic for Determining Spatial Outliers

Another approach for identifying potentially important covariates missing from our model is through spatial outliers, i.e., spatial regions which are different than their neighbors. While many methods focus on the nature of the response surface itself, such as through the use of a Bayesian p-value (Gelman et al., 1996) or via leave-one-out analyses (e.g., Stern and Cressie, 2000), these are typically designed to assess the fit of the entire model. Furthermore, a leave-one-out analysis for our model would depend heavily on the prior specifications for σ_i and τ_i^2 (i.e., removing the ith county prohibits the posterior learning about σ_i and τ_i^2), and thus is impractical here. Our work is thus more in line with that of Lu and Carlin (2005), who developed areal wombling methods for regional boundary analysis. In our case, geographical features such as mountains are known to affect factors such as air quality, and access to preventive care may also differ between counties.

Here, we have devised a diagnostic which can be used to identify regions that are potential spatial outliers using the model for $\mathbf{Z}$ in (5). Extending the expressions listed in Table 1 from a single time point to the general space-time case, we find

$$\mathbf{Z}_i \mid \mathbf{Z}_{(i)} \sim N_{N_t}\left(\alpha \frac{\sigma_i}{n_i} \sum_{k \sim i} \frac{\mathbf{Z}_k}{\sigma_k},\ \frac{\sigma_i^2}{n_i} R(\phi)\right). \tag{9}$$

Using this, we have constructed the following diagnostic:

$$Q_i(\mathbf{Z}_i \,|\, \cdot) = \left(\frac{\sqrt{n_i}}{\sigma_i}\mathbf{Z}_i - \alpha\sqrt{n_i}\frac{\sigma_i}{n_i}\sum_{k \sim i}\frac{\mathbf{Z}_k}{\sigma_k}\right)^T R(\phi)^{-1} \left(\frac{\sqrt{n_i}}{\sigma_i}\mathbf{Z}_i - \alpha\sqrt{n_i}\frac{\sigma_i}{n_i}\sum_{k \sim i}\frac{\mathbf{Z}_k}{\sigma_k}\right). \tag{10}$$

For the sake of simplicity, we suppress the conditional notation and denote (10) as $Q_i(\mathbf{Z}_i)$. In essence, $Q_i(\mathbf{Z}_i)$ is a measure of posterior learning, as large values correspond to large departures from the conditional mean of $\mathbf{Z}_i$ given by the prior. In this case, however, our prior distribution for $\mathbf{Z}$ assumes a particular degree of spatial smoothing, so large values of $Q_i(\mathbf{Z}_i)$ also indicate potential outlying regions. In practice, we obtain $Q_i(\mathbf{Z}_i^{(\ell)})$ from the ℓth iteration of the sampler and obtain a posterior distribution for $Q_i(\mathbf{Z}_i)$. Note that *a priori* the expression in (10) follows a $\chi^2_{N_t}$ distribution. As such, we recommend using the upper 95th-percentile of a $\chi^2_{N_t}$ (denoted $\chi^2_{N_t}(0.95)$) to diagnose whether or not the ith county is a spatial outlier, based on the posterior distribution for $Q_i(\mathbf{Z}_i)$. Specifically, we compute $P\left(Q_i(\mathbf{Z}_i) > \chi^2_{N_t}(0.95)\right)$, the posterior probability that $Q_i(\mathbf{Z}_i)$ exceeds $\chi^2_{N_t}(0.95)$. While it may suffice to let values greater than 0.5 suggest potential outlying regions, we recommend a more conservative approach of focusing on regions whose probability exceeds 0.95.

5 Simulation Studies

To verify the effectiveness of our model, we have conducted two simulation studies. The first example — as presented in Quick et al. (2013) — demonstrates the ability of our model to accurately capture the underlying process, $\mathbf{Z}$, and its gradient process, $\mathbf{Z}'$, as well as identify counties where the flexibility of the HCAR$_{ST}$ model is required. Then, using the posterior estimates for σ_i and τ_i^2 as our true values, we generate data for our second example directly from the HCAR$_{ST}$ model in order to ensure that our parameter estimates are accurate and meaningful. Both simulation studies are illustrated here using the $N_s = 58$ counties of California as our spatial grid, $N_t = 50$ evenly spaced timepoints, and 100 datasets. Each dataset is analyzed using a single chain run for 3,000 iterations using both the CAR$_{ST}$ and HCAR$_{ST}$ models. In addition, we also considered a spatiotemporal CAR model with an

AR(1) structure in place of the Matérn(3/2), but this yielded similar results to the CAR$_{ST}$ without permitting gradient estimation; as such, results from this model fit are suppressed for brevity. Parameter estimates will be given as posterior medians (with 95% credible intervals; 95% CI), and estimates will be evaluated via coverage—the percent of datasets in which the 95% CI contains the true value—and bias and mean squared error (MSE). A reference map highlighting a few important counties is provided in Figure 3(c).

5.1 Estimating $\mathbf{Z}$ and $\mathbf{Z}'$

In our first simulation study, we fit our models to 100 datasets arising from

$$Y_i(t_j) \overset{\text{ind}}{\sim} N\left(10\left[x_{i1} * \sin\left(\frac{t_j}{2}\right) + x_{i2} * \cos\left(\frac{t_j}{2}\right)\right], \tau_i^2\right), \tag{11}$$

where we induced spatiotemporal clustering by letting x_{i1} and x_{i2} be the ith county's % black and April 1991 ozone level, respectively, and where τ_i^2 is generated from an inverse gamma distribution yielding values of $\tau_i^2 \approx 3$. Figure 2 displays the posterior bands for $\mathbf{Z}_i$ from both the CAR$_{ST}$ and HCAR$_{ST}$ models for two particular counties: Alameda and Lassen. Here, our posterior bands are compared to the true underlying curve (black line) and the mean of the adjacent regions' true underlying curve (red line).

In addition to demonstrating the ability of the HCAR$_{ST}$ to accurately capture the trends in these two counties, these figures illustrate two distinct patterns that may be difficult for a standard spatiotemporal model to capture. First and foremost, the time trend for Alameda County appears to have a shape which is similar to that of its neighbors, but with values which are much more extreme. If we were to model these data using the CAR$_{ST}$ model, however, we would completely miss these extremes (see Figure 2(a)), as the single σ^2 is restricting Alameda County to behave like its neighbors. By using a model with region-specific σ_i, we are able to reduce the MSE for Alameda County from 237.8 to just 7.8 (compared to the average MSE of 5.5). Coverage for this county has also improved dramatically, increasing from just 34.3% to 100% without unnecessarily large credible bands.

The case of Lassen County is a bit more subtle. To begin with, both the HCAR_{ST} (Figure 2(d)) and the CAR_{ST} (Figure 2(b)) appear to estimate its time trend quite well. What is perhaps a bit disturbing here is that these models are masking the fact that Lassen County's time trend is almost the exact *opposite* of its neighboring counties, causing researchers to miss out on what may be a very important feature of these data. This deviation is brought into focus, however, when we look at $Q_i(\mathbf{Z}_i)$. As shown in Figure 4(a), our diagnostic indicates that Lassen County is markedly different than its neighbors, with a $Q_i(\mathbf{Z}_i)$ estimated to be 95.7 (72.6, 125.9), dwarfing the suggested cutoff of $\chi^2_{50}(0.95) = 67.5$.

What's worth noting here is that while our diagnostic is highlighting Lassen County as a potential outlier, it is *not* doing this for Alameda County. In this example, the features of the time trends from (11) are primarily being driven by x_{i1}, the percent of the county's population which is black. Here, the fact that Lassen County has a large black population and is surrounded by counties with small black populations is causing it to have a different time trend than its neighbors. Because of this, the model is forced to give Lassen County a large σ_i (as shown in Figure 3(a)) in order to allow it deviate from its neighbors, resulting in a large $Q_i(\mathbf{Z}_i)$. Alameda County, on the other hand, has the highest % black in the state and is surrounded by counties that are also higher than the state average. Here, the model simply attributes a large σ_i to Alameda County, yielding a similar trend on a different scale.

We now move on to the more "global" properties of the model. Overall, our model performed quite well, with 96% of the $Z_i(t_j)$ from the underlying curves being covered by their respective 95% credible intervals. Furthermore, the coverages for the $\mathbf{Z}_i$ are now more tightly concentrated around the desired 95%, with the variability of our coverages only 1/4 of that from the CAR$_{ST}$ model. Furthermore, in addition to achieving improved performance in counties such as Alameda County — where coverages were much lower than desired — we've also reduced undersmoothing in a number of regions by allowing them to have a small value of σ_i^2. Lastly, this improvement in the estimation of $\mathbf{Z}$ allows us to more accurately estimate the temporal gradients, $\mathbf{Z}'$, although a large amount of variability exists, leading to extremely high coverage.

5.2 Parameter Estimation

Our second simulation study focuses on parameter estimation; specifically, our goal is to verify that our estimates of σ_i are reasonable. We generated data from

$$Y_i(t_j) = \beta_0 + x_i(t_j)\beta_1 + Z_i(t_j) + \epsilon_i(t_j),\ i = 1, \ldots, N_s,\ j = 1, \ldots, N_t \tag{12}$$

where $x_i(t_j)$ is a continuously-varying covariate generated from a standard normal distribution, $\mathbf{Z} \sim \text{HCAR}_{ST}(\alpha, \phi, \boldsymbol{\sigma})$, and $\epsilon_i(t_j) \sim N(0, \tau_i^2)$, using the posterior medians of σ_i and τ_i from Section 5.1 and letting $\beta_0 = 3$, $\beta_1 = 2$, $\alpha = 0.85$, and $\phi = 0.75$. Here, our model performs again performs quite well, with an average of 94.2% coverage for $Z_i(t_j)$, 95% for $\boldsymbol{\beta}$, and 91.6% for our τ_i. As demonstrated in Quick et al. (2013), both ϕ and α remain difficult to estimate accurately. In particular, our coverage for α is 74%, while our coverage for ϕ is 85%, with biases for both α (-3.8%) and ϕ (+3.2%) in favor of less spatial and temporal association, respectively. Finally, the spatiotemporal variance parameters, σ_i, are generally well estimated, with 94.7% of their credible intervals containing the true value. We have also considered alternative values for α and ϕ and have obtained similar results.

Given that these data sets were generated directly from our model, it is not surprising that none of our estimated Q_i indicate spatial outliers in these data. Throughout all of our datasets, our estimated $Q_i(\mathbf{Z}_i)$ closely resemble samples from a χ^2_{50} distribution, suggesting that this distribution should serve as a good guide for which to diagnose outlying counties.

6 Data Analysis

As mentioned in Section 1, we model the asthma hospitalization rates using our covariates — ozone level, population density, percent under the age of 18, and percent black — and fixed effects for each of the 12 months of the year. Convergence of the MCMC algorithm was rapid, but we ran our sampler for 10,000 iterations to ensure the stability of our estimates. Using the deviance information criterion (DIC; Spiegelhalter et al., 2002), our HCAR$_{ST}$ model improves upon that of (2) by 2087 units despite being more complex, with p_D indicating that our model contains 1285 more effective parameters. As seen in our first simulation study, much of this improvement in DIC is attributable to a small number of counties, with four counties accounting for 40% of this change. While a number of these counties simply have more accurate fits, one of these counties is a strong candidate for being a spatial outlier, and we will revisit this county in our discussion of Q_i. In addition to assessing fit, we also assessed the sensitivity of our results to the informative priors used for α and γ^2 and also considered using a CAR structure for our u_i's, and obtained nearly identical results.

In Table 2, we present the posterior estimates for our HCAR$_{ST}$ model parameters $\boldsymbol{\beta}$, α and ϕ compared to those from the CAR$_{ST}$. Here, we find that the median for ϕ has increased 33%, indicating less temporal association in our model; this may be due to an increase in spatial association, as evidenced by the increase in α. A number of our regression coefficients have also changed. For instance, the median for β_2, the effect for percent black, has increased 44%, while the monthly effects for the summer have reduced in magnitude — the change in β_2 will be discussed in more detail later. With regard to the monthly effects, it seems as though our random effects have absorbed much of this temporal change, which will become more apparent when we discuss our temporal gradient analysis. Finally, while counter-intuitive, the negative effects of ozone are common to all standard statistical approaches (e.g., an OLS fit) for these data.

Figure 3(b) displays our estimated $HCAR_{ST}$ variance parameters, σ_i. As shown in our first simulation study, large values of σ_i can indicate regions with high within-county variability (e.g., Alameda County) or high between-county variability (e.g., Lassen County). In this case, the cluster of large σ_i in the northern counties likely indicates the latter explanation, a claim supported by the lack of spatial smoothness in the raw data (Figure 1). The error variance parameters, τ_i, are strongly negatively correlated with population (i.e., counties with low population have a higher error variance), which coincides with the variance of the rate estimate from a normal approximation to the binomial distribution.

Figure 4(b) displays the estimated values of our outlier diagnostic, Q_i. Here, the darker the shade of red, the stronger the evidence that region i is an outlier. This figure highlights Imperial County, whose $Q_i(\mathbf{Z}_i)$ has median (95% CI) $Q_i = 335$ (306, 416). This is significantly larger than $\chi^2_{216}(0.95) = 258$, indicating that this county is quite different than its neighbors. Giving credence to this result is a recent article in the *LA Times* focused on Imperial County:

> "Imperial County is different because it leads the state for asthmatic children going to the ER and being hospitalized, but experts are unable to pinpoint the cause. Doctors and public health officials said that a combination of whipping winds, pesticide-tinged farmland dust and large numbers of low-income families lacking health insurance contribute to high rates of asthma hospitalizations and ER visits." – *Gorman (2012)*

While this article proposes a number of possible explanations for Imperial County's abnormally high rates of asthma hospitalization, our analysis can help public health officials narrow this list even further by identifying risk factors in Imperial County that are not applicable to its neighbors. For instance, its geographical location is such that the two counties it neighbors — San Diego and Riverside — have their most densely populated areas near the Pacific Ocean, rather than in the desert valley. Also, as mentioned in that article, its close proximity to an industrial portion of northern Mexico may be magnifying these effects.

Turning our attention to temporal gradients, Figure 5 shows striking results. Previous research has shown that asthma hospitalization rates are lowest during the summer and highest during the winter months. Here, we map the average month-to-month temporal gradients. These results indicate that the underlying random effects achieve their apex during the late fall and early winter (Figures 5(d) and 5(e)), decline for most of the spring (Figures 5(a) and 5(b)), remain steady through July and August (Figure 5(c)), and then increase heavily during August and September, a cycle that may coincide with agricultural activity in some of the more rural areas of the state, particularly those in the Imperial and San Joaquin Valleys, where dry, windy conditions and stagnant air can lead to prolonged periods of poor air quality. Compared to the results using the CAR_{ST} model in equation (2), our results contain more extreme estimated gradients. This is not surprising, as (i) the single variance model led to oversmoothing and thus more gradual rates of change and (ii) the spatiotemporal process under the HCAR_{ST} has absorbed more of the seasonal trend that was previously contained in the monthly fixed effects.

In addition to affecting the monthly fixed effects, the HCAR_{ST} model also had a quite dramatic effect on β_2, the coefficient for % black, as first seen in Table 2. Upon further investigation, it appears that many of the counties with large σ_i are *influential* with regards to β_2; i.e., if you fit an ordinary least squares model to these data, remove one or more of these counties from the analysis, and then refit the model, you'll find that β_2 increases. This is largely due to the fact that many of the counties with large σ_i have the smallest black populations, yet are typically underestimated, and increasing β_2 to better accommodate the majority of the data leads to further underestimation in these counties. This suggests that the σ_i in the HCAR also serve to counter influential observations by allowing enough flexibility in $\mathbf{Z}_i$ to compensate for poorer fits in the regression component of the model.

7 Discussion

In this paper, we have provided a brief overview of dynamic CAR models, and proposed a novel heteroscedastic CAR (HCAR) method for permitting region-specific variance parameters in a spatiotemporal process, motivated by efforts to analyze asthma hospitalization data from California. We also examined the validity of our model via simulation, and illustrated its use by reanalyzing the California dataset. Not only did our model produce a better fit to the real data and dramatically alter the estimated effect of race, it was also able to correctly single out a known outlying region. Coupled with improved gradient estimation (by virtue of the improved fit of $\mathbf{Z}$), these results demonstrate the ability of our model to highlight features of spatiotemporal data that can be used by researchers to identify important predictors and risk factors missing from their statistical model.

One point of discussion is the interpretation of the σ_i. In our first simulation study, we discovered that large values of σ_i indicate one of two possibilities — (a) the process $\mathbf{Z}_i$ has a large temporal variance or (b) the process $\mathbf{Z}_i$ is a potential spatial outlier — and in our analysis of the asthma data, we suggest that large σ_i may also correspond to influential counties. Because of this, one must be careful not to interpret σ_i^2 as the variance of the temporal process $Z_i(t)$, but rather the variance of the temporal process $v_i(t)$. A more accurate interpretation of σ_i is as a scale parameter. For instance, we can use the σ_i to construct a spatiotemporal process $\mathbf{U} = (I_{N_t} \otimes {\Sigma_S}^{-1})\mathbf{Z}$ with unit variance, where $U_i(t_j) = Z_i(t_j)/\sigma_i$, suggesting that our model smooths $\mathbf{Z}_i/\sigma_i$ rather than directly smoothing $\mathbf{Z}_i$. Similarly, these roles of σ_i suggest that the σ_i themselves are likely to be spatially associated with outlying values corresponding to outlying $\mathbf{Z}_i$. It was for this reason that in Section 5 we chose to use posterior estimates from our first simulation to generate data for our second simulation, rather than generating true values for σ_i randomly (say, from their prior distributions).

On the subject of region-specific variances, one may wonder how the HCAR$_{ST}$ performs with $\tau_i^2 = \tau^2$, i.e., using a single error variance parameter. In the case of the asthma hospitalization data, removing the region-specific τ_i^2 from the model actually leads to our model *overfitting* the data. Here, the estimated posterior medians of the τ_i^2 in our model range from 0.13 to 250, reflecting the vast differences in the populations between the counties of California. When we require the counties to share a single error variance parameter, the model is forced to absorb the variability due to population into the σ_i, and our estimate for τ^2 reduces toward 0. In situations where the error variances are not substantially different from one-another, as in the case of our simulation study, this problem does not arise. In practice, however, we recommend always including the region-specific τ_i^2 in the model, as their inclusion does not lead to any noticeable increases in computational burden.

Another feature introduced in this paper is the outlier diagnostic, $Q_i(\mathbf{Z}_i)$, which we have demonstrated can be a powerful tool for identifying important covariates missing from the model. While this diagnostic has been presented under the HCAR$_{ST}$ framework, it is also worth noting that $Q_i(\mathbf{Z}_i)$ can be estimated for any spatiotemporal process. For instance, calculating $Q_i(\mathbf{Z}_i)$ under the CAR$_{ST}$ model in (2) yields similar results; however, due to the potential to both over- and undersmooth the $\mathbf{Z}_i$, this leads to more extreme values of $Q_i(\mathbf{Z}_i)$ on both ends of the spectrum.

As illustrated here, one strength of the modeling framework used in Quick et al. (2013) — which itself is related to the third case of order-free multivariate CAR distributions in Jin et al. (2007) — is its flexibility. An obvious generalization of their work would be to replace $R(\phi)$ with other temporal correlation structures, as many applications will not require mean-square differentiable temporal processes. In this paper, we extended their approach to allow for region-specific variance parameters in situations with a large number of temporal observations. Should there be a limited number of temporal observations, however, our methods here could easily be adapted to ensure the stability of our σ_i estimates; for instance, we could induce spatial association in the σ_i as in Reich and Hodges (2008). One may also wish to embed this model structure into a generalized linear model; such a model would still permit inference on the gradient process $\mathbf{Z}'$ but would require Metropolis updates for $\mathbf{Z}$, increasing the necessary computational burden.

Finally, we remark about future explorations with non-separable versions of HCAR$_{ST}$. The separable covariance structures considered here impose a common temporal decay term for each county. This assumption can be relaxed by formulating non-separable models. A rich class for continuous time areal models has been described in the supplemental material for Quick et al. (2013). For instance, we can replace the specification of $v_j(t)$ in Section 4 with one of the form $v_j(t) \sim GP(0, \sigma_j^2 \rho(\cdot; \phi_j))$, which admits both a different variance σ_j^2 for each county (like the HCAR$_{ST}$) *and* also a different ϕ_j for each county (unlike the separable HCAR$_{ST}$). While this may impact the gradient process — where gradients are estimated directly from the model and the estimated $\mathbf{Z}$ — we do not anticipate substantial changes in model fit, where the flexibility of the HCAR$_{ST}$ model, coupled with information from the data, will lead to $\mathbf{Z}$ which are nonseparable *a posteriori*.

Acknowledgements

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

Baladandayuthapani, V., Mallick, B., Hong, M., Lupton, J., Turner, N., and Carroll, R. (2008). “Bayesian hierarchical spatially correlated functional data analysis with application to colon carcinoginesis.” *Biometrics*, 64, 64–73.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2015). *Hierarchical Modeling and Analysis of Spatial Data*. Chapman & Hall/CRC.

Banerjee, S., Gelfand, A. E., and Sirmans, C. F. (2003). “Directional rates of change under spatial process models.” *Journal of the American Statistical Association*, 98, 946–954.

Brewer, M. and Nolan, A. (2007). “Variable smoothing in Bayesian intrinsic autoregressions.” *Environmetrics*, 18, 841–857.

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. Hoboken, NJ: Wiley.

Delamater, P. L., Finley, A. O., and Banerjee, S. (2012). “An analysis of asthma hospitalizations, air pollution, and weather conditions in Los Angeles County, California.” *Science of the Total Environment*, 425, 110–118.

Delicado, P., Giraldo, R., Comas, C., and Mateu, J. (2010). “Statistics for spatial functional data: some recent contributions.” *Environmetrics*, 21, 224–239.

Gelman, A., Meng, X.-L., and Stern, H. (1996). “Posterior predictive assessment of model fitness via realized discrepancies.” *Statistica Sinica*, 6, 733–807.

Gorman, A. (2012). “Imperial County leads state in treatment of children with asthma.” *LA Times*. `http://articles.latimes.com/2012/jul/16/local/la-me-imperial-county-asthma-20120716`.

Jin, X., Banerjee, S., and Carlin, B. P. (2007). "Order-free co-regionalized areal data models with application to multiple-disease mapping." *Journal of the Royal Statistical Society, Series B*, 69, 817–838.

Lawson, A. and Clark, A. (2002). "Spatial mixture relative risk models applied to disease mapping." *Statistics in Medicine*, 21, 359–370.

Lu, H. and Carlin, B. P. (2005). "Bayesian areal wombling for geographical boundary analysis." *Geographical Analysis*, 37, 265–285.

MacNab, Y. C. and Gustafson, P. (2007). "Regression B-spline smoothing in Bayesian disease mapping: with an application to patient safety surveillance." *Statistics in Medicine*, 26, 4455–4474.

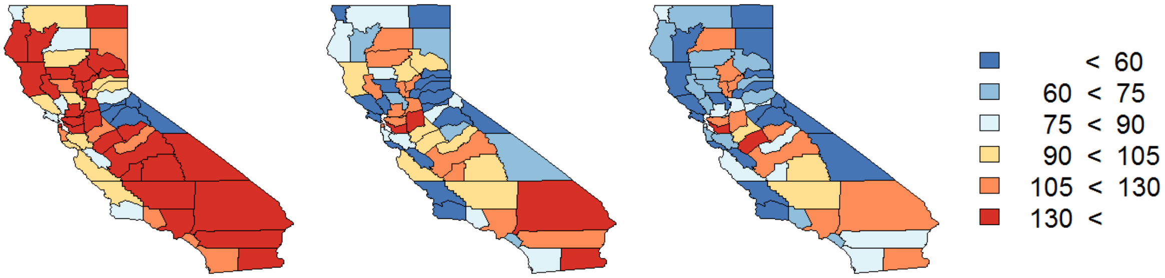
Quick, H., Banerjee, S., and Carlin, B. P. (2013). "Modeling temporal gradients in regionally aggregated California asthma hospitalization data." *Annals of Applied Statistics*, 7, 154–176.

Reich, B. and Hodges, J. (2008). "Modeling longitudinal spatial periodontal data: a spatially adaptive model with tools for specifying priors and checking fit." *Biometrics*, 64, 790–799.

Spiegelhalter, D. J., Best, N., Carlin, B. P., and van der Linde, A. (2002). "Bayesian measures of model complexity and fit (with discussion)." *Journal of the Royal Statistical Society, Series B*, 64, 583–639.

Stern, H. S. and Cressie, N. (2000). "Posterior predictive model checks for disease mapping models." *Statistics in Medicine*, 19, 2377–2397.

Ugarte, M. D., Goicoa, T., and Militino, A. F. (2010). "Spatio-temporal modeling of mortality risks using penalized splines." *Environmetrics*, 21, 270–289.



(a) 1991 Rates (b) 2000 Rates (c) 2008 Rates

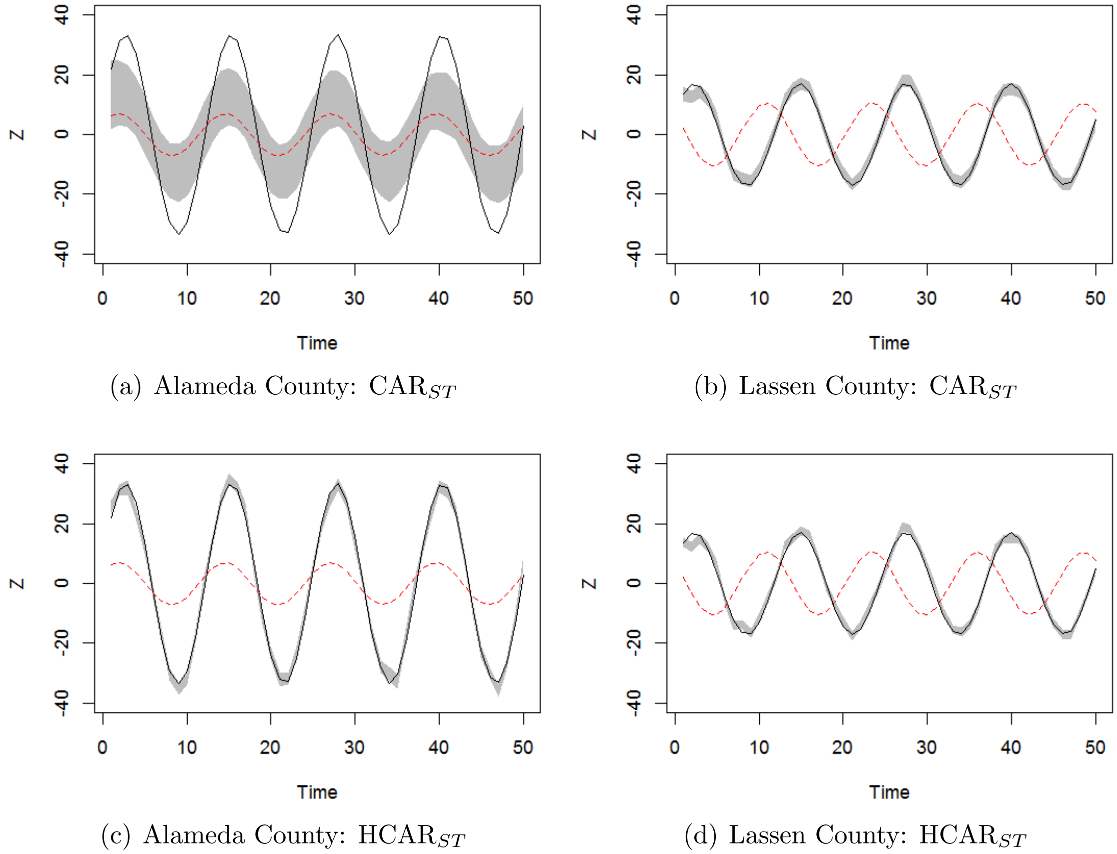
Figure 1: Raw asthma hospitalization rates per 100,000 people for select years.

	$\mathbf{Z}(t) \sim \text{CAR}(\alpha, \sigma^2)$	$\mathbf{Z}(t) \sim \text{SACAR}(\alpha, \boldsymbol{\sigma})$	$\mathbf{Z}(t) \sim \text{HCAR}(\alpha, \boldsymbol{\sigma})$
Precision	$(\sigma^2)^{-1}(D - \alpha W)$	$S - \alpha {\Sigma_S}^{-1} W {\Sigma_S}^{-1}$	${\Sigma_S}^{-1}(D - \alpha W){\Sigma_S}^{-1}$
$E\left[Z_i(t) \,\vert\, Z_{(i)}(t), \alpha, \boldsymbol{\sigma}\right]$	$\alpha \frac{1}{n_i} \sum_{k \sim i} Z_k(t)$	$\alpha \frac{1}{\sum_{k \sim i} \sigma_k} \sum_{k \sim i} \frac{Z_k(t)}{\sigma_k}$	$\alpha \frac{\sigma_i}{n_i} \sum_{k \sim i} \frac{Z_k(t)}{\sigma_k}$
$V\left[Z_i(t) \,\vert\, Z_{(i)}(t), \alpha, \boldsymbol{\sigma}\right]$	σ^2 / n_i	$\sigma_i / \left[\sum_{k \sim i} \frac{1}{\sigma_k}\right]$	σ_i^2 / n_i

Table 1: Comparison of the CAR, SACAR, and HCAR models for $\mathbf{Z}(t) = (Z_1(t), \ldots, Z_{N_s}(t))$. Here, S is a diagonal matrix with $S_{ii} = \frac{1}{\sigma_i} \sum_{j \sim i} \frac{1}{\sigma_j}$ and Σ_S is a diagonal matrix with (i, i)-element σ_i. Note that the CAR model is a special case of both the SACAR and HCAR models where $\sigma_i = \sigma$ for all i.

	Median (95% CI)			Median (95% CI)	
Parameter	$CAR(\sigma^2)$	$HCAR_{ST}(\boldsymbol{\sigma})$	Parameter	$CAR(\sigma^2)$	$HCAR_{ST}(\boldsymbol{\sigma})$
β_0 (Intercept)	9.89 (8.83, 11.05)	9.49 (8.53, 10.43)	$\beta_{15} - \beta_{26}$ (Ozone)		
β_1 (Pop Den)	0.59 (0.50, 0.70)	0.64 (0.55, 0.71)	— January	0.88 (-0.65, 2.57)	0.51 (-0.92, 1.94)
β_2 (% Black)	1.23 (1.13, 1.33)	1.78 (1.37, 1.89)	— February	0.52 (-0.91, 1.81)	0.39 (-0.62, 1.47)
β_3 (% $<$ 18)	1.13 (1.02, 1.24)	1.24 (1.13, 1.34)	— March	0.44 (-0.03, 0.86)	0.41 (-0.06, 0.89)
β_4 (Feb)	-0.16 (-0.52, 0.20)	-0.10 (-0.42, 0.23)	— April	0.25 (-0.03, 0.53)	0.20 (-0.05, 0.49)
β_5 (Mar)	-0.63 (-1.86, 0.52)	-0.25 (-1.34, 0.82)	— May	-0.17 (-0.34, 0.00)	-0.17 (-0.34, 0.00)
β_6 (Apr)	-2.15 (-3.44, -0.99)	-1.60 (-2.66, -0.52)	— June	-0.35 (-0.50, -0.18)	-0.36 (-0.52, -0.20)
β_7 (May)	-1.95 (-3.25, -0.74)	-1.40 (-2.46, -0.32)	— July	-0.24 (-0.38, -0.1)	-0.22 (-0.35, -0.09)
β_8 (June)	-3.45 (-4.76, -2.24)	-2.45 (-3.60, -1.39)	— August	-0.28 (-0.42, -0.13)	-0.20 (-0.33, -0.07)
β_9 (July)	-4.47 (-5.77, -3.26)	-3.29 (-4.49, -2.21)	— September	-0.39 (-0.56, -0.22)	-0.28 (-0.42, -0.12)
β_{10} (Aug)	-4.21 (-5.49, -2.98)	-3.16 (-4.35, -2.09)	— October	-0.05 (-0.23, 0.14)	0.05 (-0.14, 0.25)
β_{11} (Sep)	-2.55 (-3.84, -1.37)	-1.94 (-3.05, -0.89)	— November	0.48 (-0.02, 1.02)	0.51 (0.04, 1.05)
β_{12} (Oct)	-2.13 (-3.41, -1.02)	-1.79 (-2.83, -0.72)	— December	3.31 (1.48, 5.21)	3.14 (1.44, 5.01)
β_{13} (Nov)	-1.08 (-2.36, 0.05)	-0.87 (-1.93, 0.22)	α	0.78 (0.74, 0.82)	0.88 (0.85, 0.90)
β_{14} (Dec)	2.34 (1.06, 3.55)	2.42 (1.13, 3.65)	ϕ	0.93 (0.87, 0.99)	1.24 (1.18, 1.29)

Table 2: Posterior medians and 95% credible intervals (CI) for $\boldsymbol{\beta}$ and ϕ from our asthma hospitalization rate data.



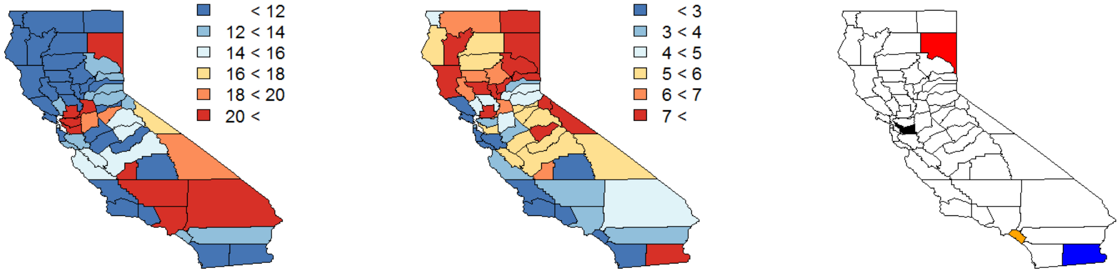
(a) Alameda County: CAR$_{ST}$

(b) Lassen County: CAR$_{ST}$

(c) Alameda County: HCAR$_{ST}$

(d) Lassen County: HCAR$_{ST}$

Figure 2: 95% CI (gray bands) of $\mathbf{Z}$ for Alameda (left) and Lassen (right) Counties from the simulation study. Results from the CAR$_{ST}$ are displayed in the top row, and those from the HCAR$_{ST}$ are displayed in the bottom row. Solid black lines denote the true underlying curves and dashed red lines denote the mean of each county's neighboring regions.

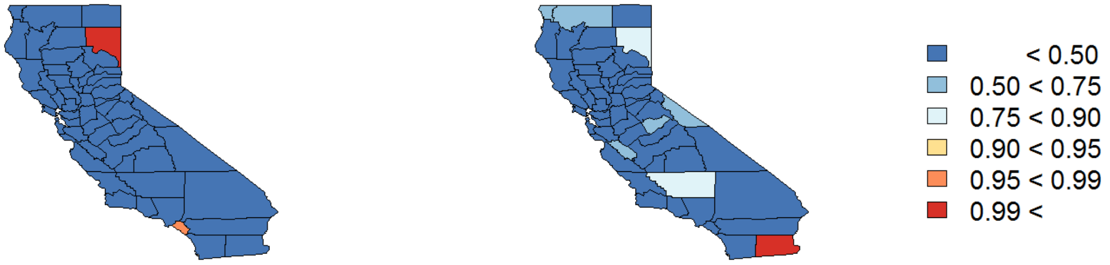


(a) Simulation: σ_i

(b) Asthma Data: σ_i

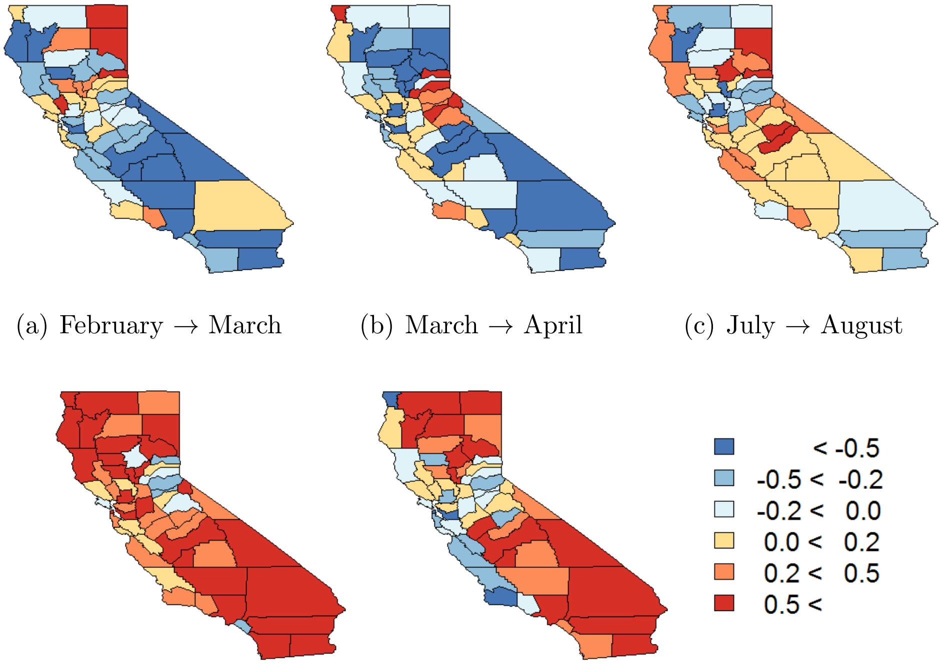
(c) Reference Map

Figure 3: Posterior medians of the σ_i from the simulation study in Section 5.1 and the analysis of the asthma hospitalization data in Section 6. The third panel is a reference for identifying key counties mentioned in the text. Here, Alameda County is shaded black, Imperial County is shaded blue, Lassen County is shaded red, and Orange County is shaded orange.



(a) Simulation: $P\left(Q_i(\mathbf{Z}_i) > \chi^2_{50}(0.95)\right)$ (b) Asthma Data: $P\left(Q_i(\mathbf{Z}_i) > \chi^2_{216}(0.95)\right)$

Figure 4: Posterior probability of $Q_i(\mathbf{Z}_i)$ being greater than $\chi^2_{N_t}(0.95)$, the suggested cutoff to determine outlying regions. Higher values correspond to a stronger indication that a particular county is a spatial outlier.



(a) February → March (b) March → April (c) July → August

(d) August → September (e) September → October

Figure 5: Selected month-to-month temporal gradients from the California asthma hospitalization data.