6-2024

Alternative Adjacency Matrices and Spatial Analysis

Jaeseong Hwang

Alternative Adjacency Matrices and Spatial Analysis

Jaeseong Hwang, Ph.D.

Western Michigan University, 2024

Spatial analysis is essential for comprehending the spatial distribution of diseases and various phenomena across geographic regions. This study investigates the utilization of alternative adjacency matrices in spatial analysis, with a specific focus on implementing Poisson regression models. This study intricately explores the methodology behind constructing alternative weight matrices, specifying weight matrices, and comparing the performance of Poisson models using five different weight matrices.

The popular Poisson model model is described, and five different definitions of weight matrices are defined, which are the following: binary weight matrix, inverse distance weight matrix using Euclidean distance, Graph distance matrix, Path matrix, and the combination matrix of Graph distance matrix and Path matrix. The first two weight matrices are commonly used in spatial analysis, and the last three weight matrices are introduced in the study.

In particular, we introduce three new weight matrices for spatially correlated random effects; Graph distance matrix, Path matrix, and the combination of Graph distance matrix and Path matrix. We investigate the performance of these new weight matrices via simulation study. Using the generated spatially correlated

random effects, the three different kinds of data sets are generated, each representing a different underlying spatial structure. The models are evaluated using the standard error and the mean square error of estimated parameters. In result, Graph distance matrix, Path matrix, and the combination matrix of Graph distance matrix and Path matrix perform well and the models using these new weight matrices have better performance than binary weight matrix and inverse distance weight matrix using Euclidean distance in the simulated data. We also apply new weight matrices for the real data analysis. The opioid-related drug overdose deaths data collected by the Michigan Death Certificates are used for the real data analysis. In result of the real data analysis, the Poisson model with the combination matrix of Graph distance matrix and Path matrix has the best fit. In conclusion, the specification of the spatial weight matrices significantly influences both model fit and parameter estimation. The simulation results provide some evidence that all weight matrices that generated in the study have a good fit of the simulated data and the combination matrix of Graph distance matrix and Path matrix is the best choice for the real data analysis.

Alternative Adjacency Matrices and Spatial Analysis

by

Jaeseong Hwang

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Statistics
Western Michigan University
June 2024

Doctoral Committee:

Dr. Duy Ngo, Ph.D., Chair
Dr. Joshua Naranjo, Ph.D.
Dr. Kevin Lee, Ph.D.
Dr. Sangwoo Lee, Ph.D.

ACKNOWLEDGMENTS

First and foremost, I extend my heartfelt gratitude to Dr. Duy Ngo, the chair of my doctoral committee. Over the past four years, he has exemplified what it means to be a great scientist and a diligent individual. I am immensely fortunate to have had such an exceptional advisor. Thank you for dedicating your time and effort to guide me.

Acknowledging the indispensable role played by the Department of Statistics, I express my deep appreciation. Dr. Hyun Bin Kang, the graduate student academic advisor, and Jeffrey Terpstra, the chair of the department, provided crucial encouragement during challenging moments.

I am equally thankful to the other members of my dissertation committee: Dr. Joshua Naranjo, Dr. Kevin Lee, and Dr. Sangwoo Lee. Their unwavering support and invaluable advice have been instrumental in my academic journey.

Lastly, my heartfelt love goes out to my parents, Duhwan Hwang and Jeongbok Huh, my wife, Juyoung Hwang, and my children, Yechan Hwang and Yena Hwang. Without their boundless support and love, I would not have been able to pen these acknowledgments.

Jaeseong Hwang

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Spatial statistics is the area of research that focuses on statistical techniques incorporating spatial elements and relationships, such as distance, in their mathematical calculations. Due to the rapid rise in the availability of spatial data and advancements in computing technology, there is a notable increase in interest in spatial analysis. When we analyze spatial data, it is important to account for spatial autocorrelation for producing reliable and interpretable results when analyzing spatial data. It allows for more accurate inference, better understanding of spatial patterns, and informed decision-making in various fields such as public health, environmental science, urban planning, and criminology, among others. The Conditional Auto-Regressive (CAR) model was introduced in 1974 to account for spatial autocorrelation with a weight matrix which controls the behavior and degree of spatial smoothing meaning that data points are averaged with their neighbors. In this study, we used the intrinsic conditional autoregressive (ICAR) model which is a special case of CAR models. While both the CAR and ICAR models are used to model spatial dependence in data, they differ in how they specify the spatial

structure and the associated conditional distributions. The CAR model relies on an explicit definition of neighbor relationships, while the ICAR model captures spatial dependence implicitly through a spatial structure, making it more flexible in certain situations. The weight matrix is involved in the ICAR model. For a set of N sub-regions, the size of the weight matrix is $N \times N$. The weight matrix is utilized to describe the spatial relationships between observations. The binary weight matrix is the most commonly used weight matrix for the ICAR model. In spatial analysis, the creation of a new weight matrix is important as it helps to measure the spatial relationships between various locations or elements in a dataset. The choice of weight matrix impacts the results of spatial analyses, so it's essential to explore different options and assess their impact on your specific research question. The purpose of this study is to review the commonly used specifications of weight matrix, introduce new specifications of weight matrix, and using these specifications, assess the effect on model performance.

1.2 Literature review

The mapping of disease incidence and prevalence has long been a part of public health, epidemiology, and the study of disease in human populations. The most famous example is John Snow's studies of the cholera epidemic in the Golden Square area of London in 1854. Usually the object of the disease mapping is to estimate the true relative risk of a disease of interest across a geographical study area normally using generalized linear mixed models (GLMMs). The term disease mapping derives from Clayton and Kaldor (1987), who defined Bayesian methods building from Poisson regression.

Suppose we observe counts of disease cases y_i for a set of regions $i = 1, \ldots, N$

partitioning the study region. We model the counts as Poisson random variables in generalized linear models, using a logarithm. In some cases we may also have observed values of region-specific covariates $\vec{\boldsymbol{X}}_i$ with associated parameters $\vec{\boldsymbol{\beta}}$. Other data often include the number of cases expected, denoted e_i. We assume the values e_i are fixed and known.

Poisson model in its most basic, fixed effects-only form is

$$y_i \sim Poisson(e_i exp(\vec{\boldsymbol{X}}_i \vec{\boldsymbol{\beta}})), \quad i = 1, \ldots, N, \tag{1.1}$$

This method operates under the assumption that including all known and observable confounding variables will yield a map devoid of artifacts, accurately depicting the true excess risk surface. However, it's important to recognize that unobserved effects may also be present in the data. Incorporating these effects into the analysis is crucial, and they are commonly referred to as random effects In the mapping context, a random effect is an extra variation and it is important to distinguish two basic forms of extra variation in spatial applications. First, as in the non-spatial scenario, there is a type of independent and spatially uncorrelated additional variation that can be considered. This is commonly known as uncorrelated heterogeneity. Another type of random effect comes from a model where it is believed that the spatial unit is correlated with neighboring spatial units. This is often referred to as correlated heterogeneity. Essentially, this indicates that there is spatial autocorrelation among the spatial units. In the model for relative risks, random effects are decomposed into a component that models the effects that vary in an unstructured way between areas (uncorrelated random effect) and a component that takes into account the effects that vary in a structured manner in areas (correlated random effect).

One commonly used spatial model is the Besag-York-Mollié (BYM) model, de-

veloped by Besag, York, and Mollié in 1991. This model considers the spatial correlation of data, recognizing that observations in neighboring areas are likely to be more similar than those in more distant areas. This model includes both the uncorrelated random effect and the correlated random effect formulated as follows:

$$y_i \sim Poisson(\mu_i), \quad i = 1, \ldots, N \tag{1.2}$$

$$log(\theta_i) = \alpha + \vec{\boldsymbol{X}}_i\vec{\boldsymbol{\beta}} + v_i + u_i, \quad i = 1, \ldots, N, \tag{1.3}$$

where $\mu_i = e_i\theta_i$, θ_i is the relative risk in the ith area, v_i is the spatially uncorrelated random effect, and u_i is the spatially correlated random effect.

The distribution model for the uncorrelated random effect is

$$v_i \sim N(0, \sigma_v^2). \quad i = 1, \ldots, N \tag{1.4}$$

For the correlated random effect, the intrinsic conditional autoregressive (ICAR) model is used

$$u_i \mid u_{-i} \sim N(\frac{1}{\sum_{j\in\delta_i} w_{ij}} \sum_{j\in\delta_i} w_{ij}u_j, \frac{1}{\sum_{j\in\delta_i} w_{ij}}\sigma_u^2), \quad i = 1, \ldots, N, \tag{1.5}$$

In 2012 Changping Zhang introduced binary weight and distance decay weight for spatial weight matrix W in his study. The binary weight W is represented as follows:

$$w_{ij} = \begin{cases} 1 & \text{if areas } i \text{ and } j \text{ are neighbours,} \\ 0 & \text{otherwise} \end{cases} \tag{1.6}$$

where neighbours are defined to be areas which share a common boundary or vertex.

The distance decay weight W is usually expressed as follows:

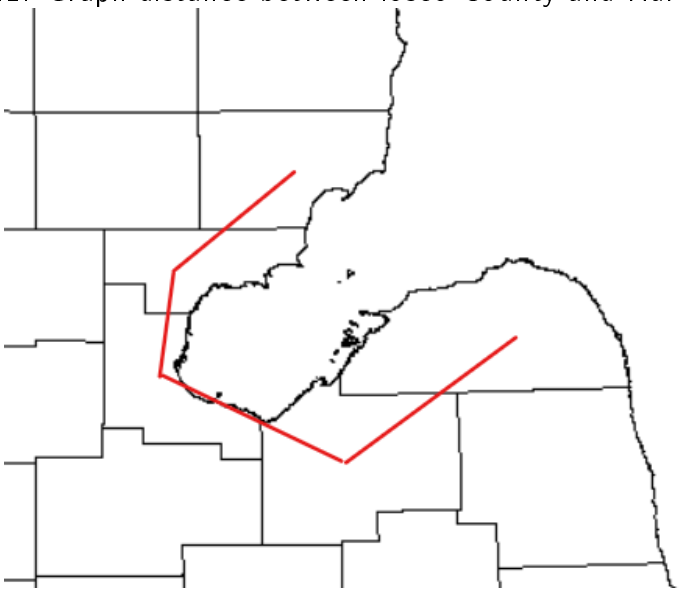
$$w_{ij} = \frac{1}{d_{ij}}, \tag{1.7}$$

where d_{ij} is the Euclidean distance from the center of units i to the neighboring units j. In this papar, we use the binary weight and the distance decay weight as traditional weight matrices.

1.3 Limitation

In this section, we discuss the disadvantage of commonly used weight matrices. The binary weight matrix provides the same weight 1 even if the distances from the centroids of counties are different. For example, in the Michigan Counties, the distance between Charlevoix County and Antrim County is 16.3 miles and the distance between Charlevoix County and Cheboygan County is 30.4 miles. Even if Cheboygan County is farther from Charlevoix County than Antrim County, the weight for Charlevoix County and Antrim County and the weight for Charlevoix County and Cheboygan County are the same. For this disadvantage of the binary weight, Distance decay weight is also commonly used for a weight matrix. The weight needs to decrease with a growing distance for a weight matrix. But sometimes Euclidean distances can be different from the distances we actually use.. For example, from Figure 1.1, the Euclidian distance between Iosco County and Huron County is 47.4 miles, buy there is Saginaw Bay between Iosco County and Huron County, so the trip distance between Iosco County and Huron County is much longer. For these limitations of common weight matrices, we consider the alternative weight matrices.

Figure 1.1: Graph distance between Iosco County and Huron County



1.4 Alternative weight matrices

One of the purpose of this study is to create new specifications of weight matrix. For this purpose, we bring some definitions from Graph Theory. Appendix A contains definitions of graph distance and shortest path, and a theorem for calculating the number of shortest paths. The set of areas of a map can be represented as a (undirected) graph that has a vertex for each area and an edge for every pair of areas that share a boundary. The graph distance between two counties is the number of boundaries in a shortest path connecting them. In other word, the graph distance between two counties is the minimum number of counties connected in one line between the two counties plus one. Figure 1.1 shows Graph distance between Iosco County and Huron County. The Graph distance between Iosco County and Huron County is 4 since there are four boundries between Iosco County and Huron County. We can define the weight matrix K using Graph distance as follows:

$$w_{ij} = \frac{1}{d_{ij}}, \tag{1.8}$$

where d_{ij} is the Graph distance between county i and county j. Generally, there are more paths between more distant areas. We can define another weight matrix using the number of paths. Let s_{ij} be the number of the shortest paths between county i and j. Then we can define a weight matrix as

$$w_{ij} = \frac{1}{s_{ij}}. \tag{1.9}$$

The value s_{ij} tells us that some pairs of areas, have multiple ways of reaching other areas to which they are not directly connected using shortest paths. If there are two areas with the same distance from a certan area and one area has more number of the shortest paths from the certan area than the other area, then the area that has more number of the shortest paths from the certan area might have more weights. Having concept of it, we can fine another weight matrix as

$$w_{ij} = \frac{1}{d_{ij} - (1 - \frac{1}{s_{ij}})}, \tag{1.10}$$

where d_{ij} is the Graph distance between county i and county j. When d_{ij} is fixed in equation, as s_{ij} is increasing, w_{ij} is increasing as well.

Chapter 2

Methodology

In the realm of disease mapping, where visualizing the spread of a disease across different geographic regions is crucial, the Bayesian model developed by Besag, York, and Mollié, known as BYM and represented by Equation (1.3), stands out as the predominant method for estimating relative risks in small areas. Relative risk is calculated by dividing the death or disease risk in a specific population area by the risk of people from all other areas. In the following, we assume that our region of interest is divided into N disjoint areas. This chapter primarily delves into the model utilized in this paper. The models used in this chapter require specifying prior distributions for the model's coefficients and random effects. Later in this chapter, we provided five different weight matrix specifications for simulation studies and analysis. To evaluate the effectiveness of our model, we compared the proposed models with the different weight matrices.

2.1 Poisson regression model

Suppose that there are N small areas in the target region. Let θ_i denote the relative risk and e_i denote the expected number of cases or deaths in the ith area,

$i = 1, \ldots, N$. These expected numbers e_i are commonly calculated as

$$e_i = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} Pop_i} Pop_i, \tag{2.1}$$

where Pop_i is the population over the ith area. Let y_i denote the number of observed cases or deaths in the ith area. When the disease is non-contagious and rare, it is usually reasonable to assume that y_i has the following a Poisson distribution, i.e,

$$y_i \sim Poisson(\mu_i), \quad i = 1, \ldots, N, \tag{2.2}$$

where $\mu_i = e_i\theta_i$.

When we assume that there are J covariates and coefficients, the log-relative risk $\log(\theta_i)$ is defined as

$$log(\theta_i) = \alpha + \vec{\boldsymbol{X}}_i\vec{\boldsymbol{\beta}} + v_i + u_i, \quad i = 1, \ldots, N, \tag{2.3}$$

where α is the intercept that represents the overall risk of the log-relative risk, $\vec{\boldsymbol{X}}_i = (x_{i1}, \ldots, x_{iJ})$ is the covariate vector of the size of $1 \times J$ for the ith area, $\vec{\boldsymbol{\beta}} = (\beta_1, \ldots, \beta_J)^T$ is the coefficient vector of the size of $J \times 1$ of $\vec{\boldsymbol{X}}_i$. The v_i is the spatially uncorrelated random effect, and the u_i is the spatially correlated random effect.

We assume that the uncorrelated random effects are independently and identically distributed (iid) and each of spatially uncorrelated random effects of $\vec{v} = (v_1, \ldots, v_N)^T$ has the following distribution,

$$v_i \sim N(0, \sigma_v^2), \quad i = 1, \ldots, N \tag{2.4}$$

Assuming that spatially correlated random effects arise from a Gaussian random field (GRF) is a common choice in spatial statistics for several reasons. The first reason is that Gaussian random fields have well-understood statistical properties, making them convenient for modeling spatial dependence. They are fully characterized by their mean and covariance functions, allowing for flexible modeling of spatial correlation structures. Another reason is that Gaussian random fields provide interpretable measures of spatial correlation, such as the spatial range and strength of dependence. These measures can inform decisions about the appropriate scale of spatial smoothing or the degree of spatial autocorrelation in the data.

The spatially correlated random effects $\vec{u}$ are assumed to have arisen from a Gaussian random field which is consistent with the assumtion that neighbouring areas have similar spatial effects. This spatial dependence is formalized by applying the intrinsic conditional autoregressive (ICAR) model on the spatially correlated random effect. The spatially correlated random variable $\vec{u} = (u_i, \ldots, u_N)^T$ follows the intrinsic conditional autoregressive (ICAR) distribution the the following,

$$u_i \mid u_{-i} \sim N(\frac{1}{\sum_{j \in \delta_i} w_{ij}} \sum_{j \in \delta_i} w_{ij} u_j, \frac{1}{\sum_{j \in \delta_i} w_{ij}} \sigma_u^2), \quad i = 1, \ldots, N, \tag{2.5}$$

where u_{-i} denotes all spatially correlated random effects except u_i, δ_i is the set of neighbors of the ith area, and w_{ij} is the entry of a symmetric weight matrix W corresponding to row i and column j.

2.2 Weights matrix specifications

According to the purpose of this paper, five different weights matrix specifications were selected for the analysis.

The weight matrix $\boldsymbol{W}_B$ is defined as follows:

$$w_{ij} = \begin{cases} 1 & \text{if counties } i \text{ and } j \text{ are neighbours,} \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

The weight matrix $\boldsymbol{W}_E$ is defined as follows:

$$w_{ij} = \frac{1}{d_{ij}}, \tag{2.7}$$

where d_{ij} is the Euclidean distance from the center of county i to the neighboring county j. Suppose that the longitude and the latitude of the county i are ϕ_i and λ_i and the longitude and the latitude of the county j are ϕ_j and λ_j and the latitude and longitude coordinates on maps are expressed in radians. The Euclidean distance d_{ij} formula is given by

$$d_{ij} = R\sqrt{(\Delta\phi)^2 + (cos(\phi_m)\Delta\lambda)^2} \tag{2.8}$$

where R is the radius of the Earth, 3,958.761 miles, ϕ_m is the average of ϕ_i and ϕ_j, $\Delta\phi = \phi_i - \phi_j$, and $\Delta\lambda = \lambda_i - \lambda_j$.

The weight matrix $\boldsymbol{W}_G$ is defined as follows:

$$w_{ij} = \frac{1}{d_{ij}}, \tag{2.9}$$

where d_{ij} is the Graph distance between county i and county j.

The weight matrix $\boldsymbol{W}_P$ is defined as follows:

$$w_{ij} = \frac{1}{s_{ij}}, \tag{2.10}$$

Figure 2.1: Map of Southwest Michigan including Berrien County and Calhoun County

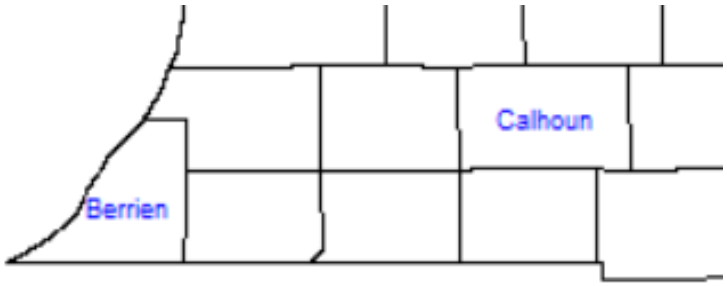


Table 2.1: Weights between Berrien County and Calhoun County

Weight specification	Weight
$\boldsymbol{W}_B$ binary,	0
$\boldsymbol{W}_E$, Euclidean dist.	0.013
$\boldsymbol{W}_G$, Graph dist.	0.33
$\boldsymbol{W}_P$, num. of paths	0.25
$\boldsymbol{W}_C$, Graph dist. with Paths	0.44

where s_{ij} is the number of the shortest paths between county i and county j.

The weight matrix $\boldsymbol{W}_C$ is defined as follows:

$$w_{ij} = \frac{1}{d_{ij} - (1 - \frac{1}{s_{ij}})}, \tag{2.11}$$

where d_{ij} is the Graph distance between county i and county j and s_{ij} is the number of the shortest paths between county i and county j.

According to Equation (2.8), the distance between Berrien County and Calhoun County is around 75 miles when calculated using the Euclidean method. The Figure 2.1 shows that the Graph distance between Berrien County and Calhoun County is 3 and the number of the shortest paths Berrien County and Calhoun County is 4. The denominator $d_{ij} - (1 - \frac{1}{s_{ij}})$ in Equation (2.10) is $3 - (1 - \frac{1}{4}) = \frac{9}{4}$.

Table 2.1 shows the distinct weights between Berrien County and Calhoun County. The binary weight between these counties is 0 since they are not adjacent.

The Euclidean weight, measuring 0.013, is notably smaller than the new weights. Let's take Kalamazoo County and Calhoun County as neighboring examples. The Euclidean distance weight between them is calculated as 1/59, which equals 0.016, given that the distance between them is approximately 59 miles. Consequently, when considering distance decay, the new weights seem to be more advantageous than the Euclidean distancer weight.

2.3 Prior distributions

Since our models are implemented in the Bayesian setting we assign prior distributions for each parameter. For the intercept α and the coefficents β_j,

$$\alpha \sim N(0, \sigma_\alpha^2) \tag{2.12}$$

$$\beta_j \overset{\text{i.i.d.}}{\sim} N(0, \sigma_{\beta_j}^2), \quad j = 1, \ldots, J, \tag{2.13}$$

where the hyperparameters $\sigma_\alpha, \sigma_{\beta_1}, \ldots, \sigma_{\beta_j}$ were assumed to have a Uniform(0,10) distribution. Suitable prior distributions for the spatially correlated random variable effect and associated variance are

$$v_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma_v^2), \quad i = 1, \ldots, N, \tag{2.14}$$

$$\sigma_v \sim Unif(0, 10) \tag{2.15}$$

We assume that u_i follows the intrinsic conditional autoregressive (ICAR) distribution discussed in Equation (2.5) and the hyperparameter σ_u for ICAR distribution is following gamma distribution:

$$\sigma_u \sim Gamma(0.01, 0.01), \tag{2.16}$$

Let $\vec{\psi}$ be a vecter of the prior and the hyperprior parameters, $(\alpha, \beta_1, \ldots, \beta_J, \sigma_\alpha, \sigma_{\beta_1}, \ldots, \sigma_{\beta_J}, \sigma_v, \sigma_u)^T$. Then we propose a join prior for $\vec{\psi}$ to be of the form

$$\begin{aligned} p(\vec{\psi}) &= p(\alpha, \beta_1, \ldots, \beta_J, \sigma_\alpha, \sigma_{\beta_1}, \ldots, \sigma_{\beta_J}, \sigma_v, \sigma_u) \\ &= p(\alpha)p(\beta_1)\cdots p(\beta_J)p(\sigma_\alpha)p(\sigma_{\beta_1})\cdots p(\sigma_{\beta_J})p(\sigma_v)p(\sigma_u) \\ &\propto \frac{1}{\alpha, \beta_1, \cdots, \beta_J} exp(-\frac{\alpha^2}{2\sigma_\alpha} - \frac{1}{2}\sum_{j=1}^{J}\frac{\beta_j^2}{\sigma_{\beta_j}^2}) \times (\sigma_u^2)^{-0.99} exp(-\frac{\sigma_u^2}{0.01}) \\ &= \frac{(\sigma_u^2)^{-0.99}}{\alpha, \beta_1, \cdots, \beta_J} exp(-\frac{\alpha^2}{2\sigma_\alpha} - \frac{\sigma_u^2}{0.01} - \frac{1}{2}\sum_{j=1}^{J}\frac{\beta_j^2}{\sigma_{\beta_j}^2}). \end{aligned} \tag{2.17}$$

Let $\vec{y}$ be a vecter of $(y_1, \ldots, y_N)^T$. Then the likelihood function for $\vec{y}$ given $\vec{\psi}$ based on Equation (2.2) and Equation (2.3) is given by

$$L(\vec{y}|\vec{\psi}) = \prod_{i=1}^{N} \frac{[e_i exp(\alpha + \sum_{j=1}^{J}\beta_j X_j + v_i + u_i]^{y_i} exp[-e_i exp(\alpha + \sum_{j=1}^{J}\beta_j X_j + v_i + u_i]}{(y_i)!}. \tag{2.18}$$

The posterior density of $\vec{\psi}$ given $\vec{y}$ for our study is given by

$$P(\vec{\psi}|\vec{y}|) \propto L(\vec{y}|\vec{\psi}) \times p(\vec{\psi}), \tag{2.19}$$

where $p(\vec{\psi})$ is the prior in Equation (2.17) and $L(\vec{y}|\vec{\psi})$ is the likelihood function in Equation (2.18).

The posterior was simulated via MCMC chains, which converges to the desired posterior distribution. We used a Metropolis-Hastings algorithm to generate MCMC chains. Algorithm consists of the following steps repeated T iterations of the chain. We obtain the chain values for each parameters, which we summarize to get the posterior mean. A model was run for 100000 iterations, the first 50000 being discarded as burn-in.

2.4 Convergence diagnostics

Generally, convergence is assessed first informally using visual examination of the trace plots. Also, we use the potential scale reduction factor, $\hat{R}$. If the chains have converged to the target posterior distribution, then $\hat{R}$ should be close to 1. Brooks and Gelman (1997) have suggested that if $\hat{R} < 1.2$ for all model parameters, we can be fairly certain that convergence has been reached. Even more reassuring is applying the more stringent condition $\hat{R} < 1.1$. We construct two chains from the remaining iterations after discarding the burn-in.

2.5 Goodness of fit

We choose the models from BYM models with different weight matrices. We use the deviance information criterion (DIC) proposed by Spiegelhalter et al (2002) as the model selection criteria. DIC is widely used in Bayesian modeling and DIC is a hierarchical modeling generalization of the Akaike information criterion (AIC). The parameters degrees of freedom (pD) reflect the model complexity, which is considered in the DIC calculations, higher pD indicating a more complex model. Models with smaller DIC should be preferred to models with larger DIC.

Chapter 3

Simulation Study

In this chapter, we will introduce the Poisson model for our simulation study, incorporating specifications regarding the weights matrix. Additionally, we will explore techniques for generating spatially correlated variables and datasets for the simulation study, assessing the efficacy of the model. Finally, the chapter concludes with an analysis of the simulation outcomes.

3.1 Simulation setting

We chose the county map of the lower peninsula of Michigan as our target regions, USA which has 68 areas (counties). Hence, we consider the 68 counties in the lower peninsula as the spatial domain and let $i = (1, \ldots, 68)$ be the county set.

We assume that the outcomes y_i, $i = 1, \ldots, 68$ follow Poisson distributions:

$$y_i \sim Poisson(\mu_i), \quad i = 1, \ldots, 68, \tag{3.1}$$

where $\mu_i = e_i\theta_i$. We fix $e_i = 1$ for all weights matrix specifications. For the log

relative risk, it is calculated by

$$log(\mu_i) = \alpha + \vec{\boldsymbol{X}}_i\vec{\boldsymbol{\beta}} + v_i + u_i,$$

where α is the overall fixed effect, $\vec{\boldsymbol{\beta}}$ is the effect of the covariate $\vec{\boldsymbol{X}}_i$, and v_i and u_i are spatially uncorrelated and correlated random effects respectively.

We considered two simulation scenarios, in the first one we included three covariates, $\vec{\boldsymbol{X}}_i = (x_{i1}, x_{i2}, x_{i3})^T$, the truth intercept $\alpha = 1$ and the truth coefficients $\vec{\boldsymbol{\beta}} = (\beta_1, \beta_2, \beta_3)^T = (1, 1, 1)^T$, while in the second scenario we included three covariates $\vec{\boldsymbol{X}}_i = (x_{i1}, x_{i2}, x_{i3})^T$, the true intercept $\alpha = 0.7$ and the true coefficients $\vec{\boldsymbol{\beta}} = (\beta_1, \beta_2, \beta_3)^T = (0.7, 0.7, 0.7)^T$.

3.2 Spatially correlated random effect

We assumed that the initial $u = (u_1, \ldots, u_{68})$ follows Normal distribution. Specifically,

$$u_i \sim N(0, 1), \quad i = 1, \ldots, N, \tag{3.2}$$

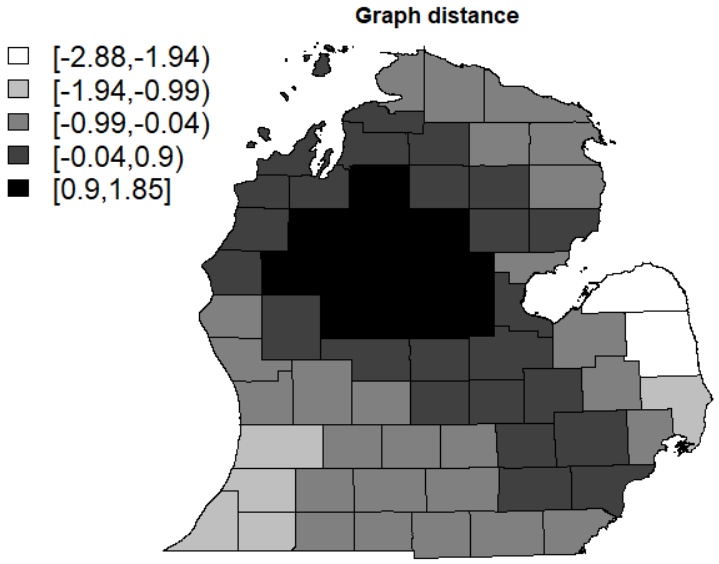
Let r_i be a row vector of a weight matrix W. Then $\langle r_i, \vec{1}\rangle$ is the sum of all weights between ith county and its neighbors, where $\langle \cdot, \cdot \rangle$ is the inner product of two vectors. We assume that σ_u is the following gamma distribution:

$$\sigma_u \sim Gamma(0.01, 0.01), \tag{3.3}$$

We choose u_i as the following Normal distribution:

$$u_i \mid u_{-i} \sim N(\frac{1}{\langle r_i, \vec{1}\rangle}\langle r_i, u\rangle, \frac{1}{\langle r_i, \vec{1}\rangle}\sigma_u^2), \quad i = 1, \ldots, N, \tag{3.4}$$

Figure 3.1: Map of simulated spatially correlated random effects using Weight matrix $\boldsymbol{W}_G$ over the lower peninsula of Michigan



We replaced the new u_i with the ith element of u. We repeat the process until u convergences. The R package, mclcar, offers various tools for the simulation of different kinds of the spatial random variable. We used the R function, $CAR.simWmat$ from R package, mclcar, to generate the spatially correlated random effects u.

Figure 3.1 is the map of the spatially correlated random effects generated by the R function, $CAR.simWmat$ with the Weight matrix specification $\boldsymbol{W}_G$. Based on the simulated map, by checking the deepness of color of each region, we can find there is a variation of the spatially structured covariates in the lower peninsula Michigan and the regions in deep dark indicate spatial clusters. Figure 3.2 is the map of the spatially correlated random effects generated by the R function, $CAR.simWmat$ with the Weight matrix specification $\boldsymbol{W}_P$ and it shows that the spatial clusters are located in eastern and northern regions of the state. Figure 3.3 is the map of the spatially correlated random effects generated by the R function, $CAR.simWmat$ with the Weight matrix specification $\boldsymbol{W}_C$ and it shows that the

Figure 3.2: Map of simulated spatially correlated random effects using Weight matrix W_P over the lower peninsula of Michigan

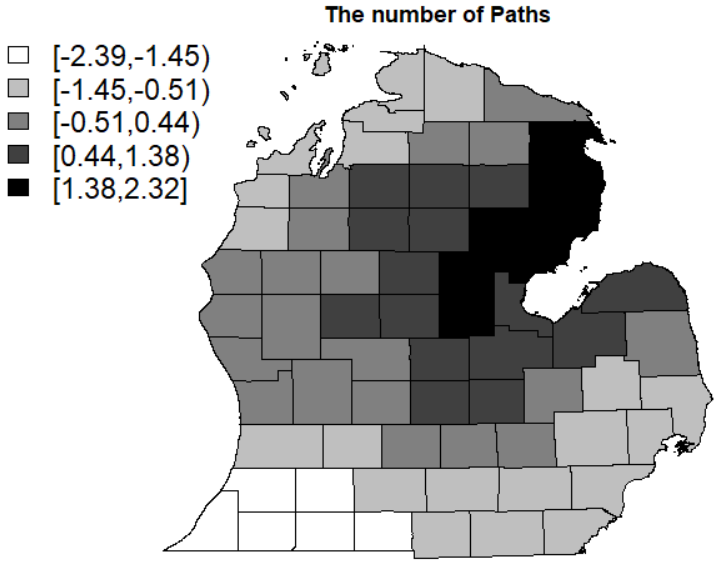
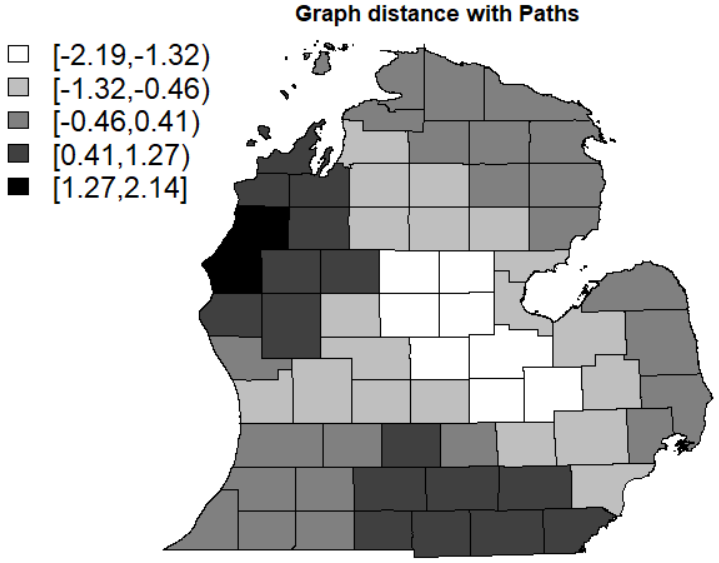


Figure 3.3: Map of simulated spatially correlated random effects using Weight matrix W_C over the lower peninsula of Michigan



spatial clusters are located in western region and southen region of the state.

3.3 Simulated datasets

A shapefile is a simple format for storing the geometric location and attribute information of geographic features. We have a Michigan shapefile from DATA.GOV website to create datasets. We generated 200 datasets over the raster of the lower peninsula of Michigan that was created by reading shapefile. We simulated 68 standard Normal random numbers with mean 0 and standard deviation 1 for each covariate x_{i1}, x_{i2}, and x_{i3}, $i = 1, \ldots, 68$. Also, we simulated 68 standard Normal random numbers with mean 0 and standard deviation 1 for spatially uncorrelated random effect v_i, $i = 1, \ldots, 68$. We simulated 68 observed counts y_i, $i = 1, \ldots, 68$, from Poisson distribution with mean $exp(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i + v_i)$, where $(\alpha, \beta_1, \beta_2, \beta_3)^T = (1, 1, 1, 1)^T$ for scenario 1 and $(\alpha, \beta_1, \beta_2, \beta_3)^T = (0.7, 0.7, 0.7, 0.7)^T$ scenario 2 and u_i is the simulated spatially correlated random effect.

3.4 Model evaluation

After fitting the simulated data with the created variables, we evaluated the performance o fthe models by checking the accuracy of the estimated coefficients though mean square error (MSE). The MSE is calculated as

$$MSE_j = \frac{\sum_{k=1}^{n} (\hat{\beta}_{jk} - \beta_{jk})^2}{n}, \tag{3.5}$$

where n is the number of simulated data.

3.5 Simulation results

In this section, we will examine how well the models perform on datasets created by specific weight matrices $\boldsymbol{W}_G$, $\boldsymbol{W}_P$, and $\boldsymbol{W}_C$. We will compare the model performances by analyzing the mean square error and standard deviation of estimated intercepts and coefficients. It is anticipated that the model utilizing a particular weight matrix will exhibit the highest performance on the dataset generated by that weight matrix.

3.5.1 Weight matrix specification $\boldsymbol{W}_G$

We compared the performance of the models with five distinct weight matrices and the model without a spatially correlated random effect for the datasets generated by the spatially correlated random variables and the weight matrix $\boldsymbol{W}_G$. Table 3.1 shows the results from the first simulation scenario, which the truth value of the intercept and the coefficients were set to 1. We displayed the mean square error (MSE) for estimated intercept and each coefficient to fit the models in Table 3.1.

The lower MSE value indicates a better estimate. We found that the model with the weight matrix specification $\boldsymbol{W}_G$ has better estimates based on the value of the MSE of the intercept and each coefficient. Table 3.1 presents the standard deviation of estimated intercept and each coefficient respectively from scenario 1. We found that the standard deviation of the estimated intercepts and coefficients for $\boldsymbol{W}_G$ is smaller than the standard deviations of the intercepts and coefficients for the other weight matrices.

Table 3.2 displays the results from the second simulation scenario, which the truth value of the coefficients were set to 0.7. We displayed the mean square error

Table 3.1: MSE and Standard deviation of estimated intercept and coefficients from scenario 1 using W_G for simulated data

Weight specification		$\hat{\alpha}(=1)$	$\hat{\beta}_1(=1)$	$\hat{\beta}_2(=1)$	$\hat{\beta}_3(=1)$
W_B	MSE	0.1223	0.0453	0.0654	0.0567
	SD	0.1796	0.1915	0.2068	0.1983
W_E	MSE	0.1788	0.0678	0.0642	0.0963
	SD	0.2588	0.2467	0.2683	0.2178
W_G	MSE	0.0569	0.0214	0.0155	0.0190
	SD	0.1567	0.1357	0.1166	0.1574
W_P	MSE	0.1455	0.0523	0.0542	0.0490
	SD	0.2178	0.2678	0.1896	0.2246
W_C	MSE	0.1364	0.0422	0.0490	0.0436
	SD	0.1811	0.1890	0.1996	0.1899
No random effect	MSE	0.1890	0.0473	0.0795	0.0679
	SD	0.2102	0.1815	0.2684	0.2190

Table 3.2: MSE, Mean, and Standard deviation of estimated intercept and coefficients from scenario 2 using W_G for simulated data

Weight specification		$\hat{\alpha}(=0.7)$	$\hat{\beta}_1(=0.7)$	$\hat{\beta}_2(=0.7)$	$\hat{\beta}_3(=0.7)$
W_B	MSE	0.0456	0.0413	0.0398	0.0456
	SD	0.1246	0.1374	0.1567	0.1865
W_E	MSE	0.0518	0.0467	0.0511	0.0413
	SD	0.1256	0.1663	0.1845	0.1796
W_G	MSE	0.0199	0.0231	0.0298	0.0311
	SD	0.1132	0.1411	0.1521	0.1134
W_P	MSE	0.0459	0.0413	0.0524	0.0313
	SD	0.1555	0.1735	0.1588	0.1519
W_C	MSE	0.0211	0.0467	0.0516	0.0422
	SD	0.1579	0.1979	0.1674	0.1590
No random effect	MSE	0.0689	0.0535	0.0669	0.0314
	SD	0.1956	0.1607	0.2001	0.1745

Table 3.3: Potential scale reduction factors, $\hat{R}$, of estimated intercept and coefficients from scenario 1, 2 using W_G for simulated data

Weight specification		$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
W_B	scenario 1	1.0981	1.0203	1.0267	1.0279
	scenario 2	1.0223	1.0411	1.0238	1.0224
W_E	scenario 1	1.0230	1.0752	1.0780	1.0223
	scenario 2	1.0212	1.0211	1.0239	1.0279
W_G	scenario 1	1.0376	1.0725	1.0716	1.0779
	scenario 2	1.0228	1.040	1.0204	1.0233
W_P	scenario 1	1.0746	1.0748	1.0112	1.0920
	scenario 2	1.0238	1.0291	1.0133	1.0401
W_C	scenario 1	1.0756	1.0722	1.0761	1.0811
	scenario 2	1.0232	1.0221	1.0227	1.0231
No random effect	scenario 1	1.0326	1.0222	1.0521	1.0711
	scenario 2	1.0331	1.0341	1.0215	1.0267

(MSE) for estimated intercept and each coefficient to fit the models in Table 3.2. Also, Table 3.2 presents the standard deviation of estimated intercept and each coefficient respectively from scenario 2.

Table 3.3 shows the maximums of the estimated potential scale reduction factors ($\hat{R}$) for all model parameters from scenario 1 and 2 respectively. As suggested by Brooks and Gelman (1997), we can be reasonably confident that convergence is reached if $\hat{R} < 1.2$ for all model parameters. According to Table 3.3, we can assume that every estimated intercept or coefficiant from scenario 1 and 2 is convergent.

3.5.2 Weight matrix specification W_P

We evaluated how well the poisson model with weight matrix specification W_P performed compared to poisson models with four other unique weight matrices and a model without a spatially correlated random effect. The datasets were created using spatially correlated random variables and the weight matrix W_P. In Table 3.4, the results from the initial simulation scenario are presented, where the intercept

Table 3.4: MSE and Standard deviation of estimated intercept and coefficients from scenario 1 using W_P for simulated data

Weight specification		$\hat{\alpha}(=1)$	$\hat{\beta}_1(=1)$	$\hat{\beta}_2(=1)$	$\hat{\beta}_3(=1)$
W_B	MSE	0.0966	0.0577	0.1569	0.0486
	SD	0.1998	0.2167	0.2126	0.1812
W_E	MSE	0.0956	0.0578	0.1356	0.0411
	SD	0.2176	0.2678	0.1891	0.1973
W_G	MSE	0.0745	0.0498	0.1130	0.0451
	SD	0.1988	0.2604	0.2862	0.1972
W_P	MSE	0.0402	0.0333	0.0441	0.0132
	SD	0.1346	0.1613	0.1811	0.1212
W_C	MSE	0.0811	0.0375	0.1431	0.0513
	SD	0.2145	0.1864	0.1945	0.2303
No random effect	MSE	0.0786	0.0394	0.1963	0.0456
	SD	0.1980	0.1932	0.2267	0.1896

and coefficients were set to a truth value of 1. The table displays the mean square error (MSE) for the estimated intercept and each coefficient to assess the model fits.

We discovered that the model utilizing the weight matrix specification W_P provides more accurate estimates, as indicated by the MSE values of the intercept and coefficients. Additionally, Table 3.4 displays the standard deviations of the estimated intercept and coefficients from scenario 1. It was observed that the standard deviation of the estimated intercepts and coefficients for W_P is lower compared to the standard deviations of the intercepts and coefficients for the other weight matrices.

Table 3.5 displays the results from the second simulation scenario, which the truth value of the coefficients were set to 0.7. We displayed the mean square error (MSE) for estimated intercept and each coefficient to fit the models in Table 3.5. Also, Table 3.5 presents the standard deviation of estimated intercept and each coefficient respectively from scenario 2.

Table 3.5: MSE and Standard deviation of estimated intercept and coefficients from scenario 2 using W_P for simulated data

Weight specification		$\hat{\alpha}(=0.7)$	$\hat{\beta}_1(=0.7)$	$\hat{\beta}_2(=0.7)$	$\hat{\beta}_3(=0.7)$
W_B	MSE	0.0353	0.0432	0.0563	0.0245
	SD	0.1467	0.1679	0.1786	0.1321
W_E	MSE	0.0345	0.0196	0.0478	0.0511
	SD	0.1578	0.1467	0.1597	0.1532
W_G	MSE	0.0214	0.0256	0.0389	0.0261
	SD	0.1386	0.1545	0.1533	0.1667
W_P	MSE	0.0143	0.0133	0.0243	0.0145
	SD	0.1208	0.1255	0.1455	0.1232
W_C	MSE	0.0452	0.0321	0.0311	0.0352
	SD	0.1374	0.1488	0.1581	0.1678
No random effect	MSE	0.0489	0.0435	0.0648	0.0356
	SD	0.1503	0.1752	0.1660	0.1678

Table 3.6: Potential scale reduction factors, $\hat{R}$, of estimated intercept and coefficients from scenario 1, 2 using W_P for simulated data

Weight specification		$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
W_B	scenario 1	1.0671	1.0578	1.0807	1.0669
	scenario 2	1.0233	1.0140	1.0238	1.0824
W_E	scenario 1	1.0760	1.0702	1.0380	1.0623
	scenario 2	1.0288	1.0414	1.0407	1.0299
W_G	scenario 1	1.0776	1.0485	1.0506	1.0879
	scenario 2	1.0218	1.0808	1.0804	1.0273
W_P	scenario 1	1.0759	1.0548	1.0512	1.0820
	scenario 2	1.0858	1.0551	1.0253	1.0246
W_C	scenario 1	1.0536	1.0822	1.0661	1.0116
	scenario 2	1.0201	1.0251	1.0757	1.0238
No random effect	scenario 1	1.1755	1.0728	1.1721	1.0821
	scenario 2	1.0358	1.0351	1.0257	1.0267

Table 3.6 shows the maximums of the estimated potential scale reduction factors ($\hat{R}$) for all model parameters from scenario 1 and 2 respectively. According to Table 3.6, we can assume that every estimated intercept or coefficiant from scenario 1 and 2 is convergent since $\hat{R} < 1.2$ for all model parameters.

3.5.3 Weight matrix specification W_C

We conducted a comparative assessment of the Poisson model's performance using the weight matrix specification Wc against four other distinct weight matrices and a model without a spatially correlated random effect, utilizing datasets generated by spatially correlated random variables and the weight matrix Wc. Table 3.7 presents the findings from the first simulation scenario, where the intercept and coefficients were set to a truth value of 1. Mean square error (MSE) values for the estimated intercept and each coefficient were provided in Table 3.7, with lower MSE values indicating superior estimates. Our analysis suggests that the model incorporating the weight matrix specification Wc produced better estimates, as indicated by the MSE values for the intercept and coefficients. Additionally, Table 3.7 presents the standard deviation of estimated intercept and each coefficient from scenario 1. We discovered that the standard deviation of the estimated intercepts and coefficients for weight matrix $\boldsymbol{W}_C$ is less than the standard deviations of the intercepts and coefficients for the remaining weight matrices.

In Table 3.8, we present the results from the second simulation scenario, where the truth value of the coefficients was set to 0.7. Mean square error (MSE) values for the estimated intercept and each coefficient are displayed in Table 3.8, along with the standard deviation of estimated intercept and each coefficient from scenario 2.

Table 3.9 shows the maximum estimated potential scale reduction factors ($\hat{R}$)

Table 3.7: MSE and Standard deviation of estimated intercept and coefficients from scenario 1 using W_C for simulated data

Weight specification		$\hat{\alpha}(=1)$	$\hat{\beta}_1(=1)$	$\hat{\beta}_2(=1)$	$\hat{\beta}_3(=1)$
W_B	MSE	0.1341	0.0786	0.0438	0.0839
	SD	0.1743	0.1563	0.1632	0.1527
W_E	MSE	0.1783	0.0731	0.0561	0.0676
	SD	0.1656	0.1463	0.1585	0.1741
W_G	MSE	0.1420	0.0413	0.0316	0.0411
	SD	0.1563	0.1689	0.1833	0.1836
W_P	MSE	0.1467	0.0313	0.0465	0.0466
	SD	0.1642	0.1887	0.1909	0.2362
W_C	MSE	0.0212	0.0221	0.0145	0.0315
	SD	0.1247	0.1156	0.1264	0.1342
No random effect	MSE	0.1673	0.0563	0.0485	0.0466
	SD	0.1875	0.1861	0.1977	0.1679

Table 3.8: MSE and Standard deviation of estimated intercept and coefficients from scenario 2 using W_C for simulated data

Weight specification		$\hat{\alpha}(=0.7)$	$\hat{\beta}_1(=0.7)$	$\hat{\beta}_2(=0.7)$	$\hat{\beta}_3(=0.7)$
W_B	MSE	0.0347	0.0415	0.5872	0.0768
	SD	0.1365	0.1955	0.3475	0.3531
W_E	MSE	0.0622	0.0788	0.0999	0.0780
	SD	0.1466	0.1898	0.2567	0.3145
W_G	MSE	0.0291	0.0412	0.0556	0.0578
	SD	0.1377	0.1890	0.2996	0.2784
W_P	MSE	0.0285	0.0787	0.0556	0.0784
	SD	0.1378	0.1883	0.2453	0.2731
W_C	MSE	0.0142	0.0216	0.0241	0.0311
	SD	0.1298	0.1733	0.2114	0.2327
No random effect	MSE	0.0356	0.0463	0.0818	0.0759
	SD	0.1787	0.1985	0.2678	0.2931

Table 3.9: Potential scale reduction factors, $\hat{R}$, of estimated intercept and coefficients from scenario 1, 2 using W_C for simulated data

Weight specification		$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
W_B	scenario 1	1.0681	1.0703	1.0747	1.0149
	scenario 2	1.0543	1.0664	1.0748	1.0154
W_E	scenario 1	1.0570	1.0782	1.0850	1.0653
	scenario 2	1.0224	1.0241	1.0247	1.0235
W_G	scenario 1	1.0746	1.0585	1.0746	1.0759
	scenario 2	1.0224	1.0258	1.0235	1.0843
W_P	scenario 1	1.0779	1.0478	1.0412	1.0720
	scenario 2	1.0288	1.0791	1.0241	1.0214
W_C	scenario 1	1.0736	1.0752	1.0861	1.0611
	scenario 2	1.0235	1.0241	1.0257	1.0801
No random effect	scenario 1	1.1726	1.082	1.884	1.0421
	scenario 2	1.0531	1.0481	1.0247	1.0267

for all model parameters from scenarios 1 and 2. As suggested by Brooks and Gelman, convergence can be reasonably assumed if $\hat{R} < 1.2$ for all model parameters. Analysis of Table 3.6 suggests that every estimated intercept or coefficient from scenarios 1 and 2 has converged.

Figure 3.4, 3.5, and 3.6 include include the spatially structured random variable maps for all models discussed with the weight specification W_G, W_P, and W_C. We find that the estimated spatially correlated random effect is consistent in all models discussed and that the major cluster is located in the northwestern region of the state. By comparing the estimated spatially structured random variable maps for all models discussed with the generated spatially correlated structure variable, we found that estimations are close to the simulated values that the maps look similar.

Graphical methods are also used for convergence diagnosis. Figures 3.7, 3.8, 3.9, 3,10, and 3,11 show the trace plots which show the realizations of the Markov chain at each iteration against the iteration numbers for the estimated coefficients and deviances for the all models discussed. We observe a stable, dense cloud of

Figure 3.4: Spatially structured random variable map for the weight specification W_G

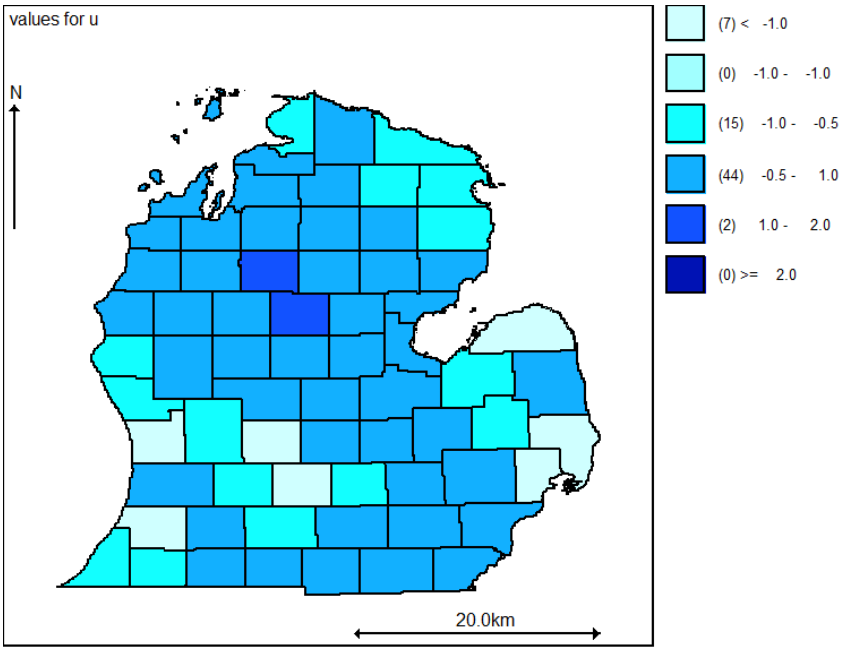


Figure 3.5: Spatially structured random variable map for the weight specification W_P

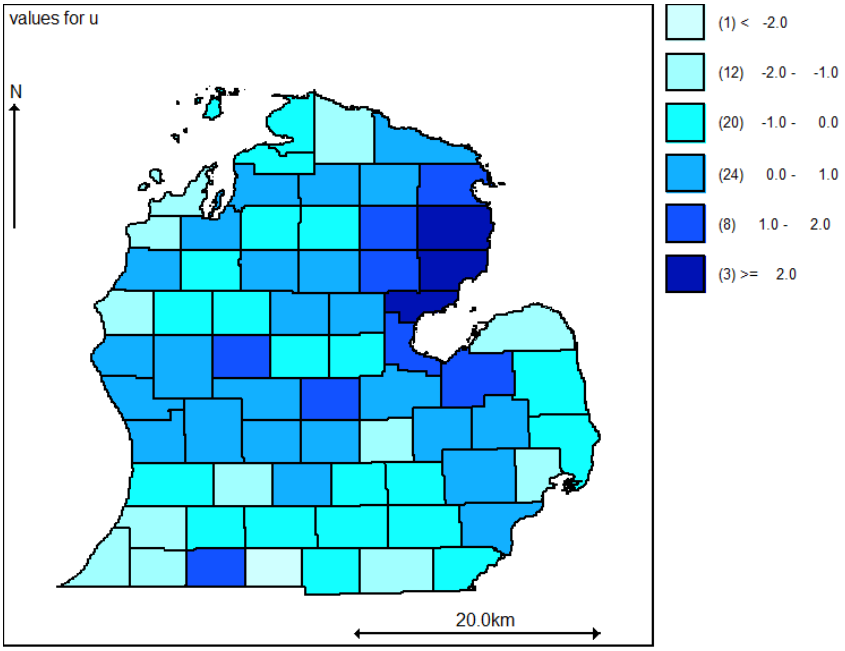
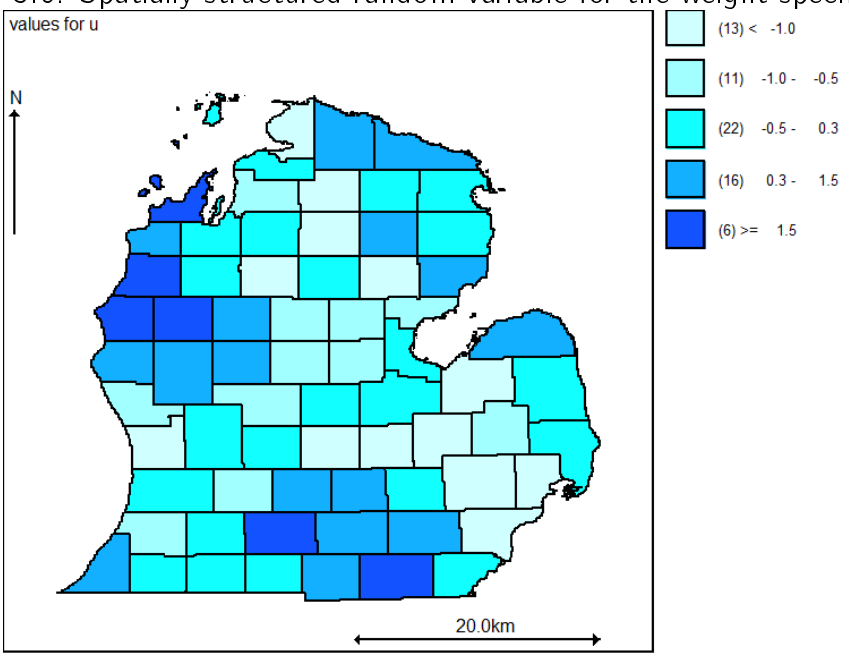


Figure 3.6: Spatially structured random variable for the weight specification W_C



points with no systematic trends. The two chains in each trace plot are well mixed which indicates a good convergence.

We anticipated that the model utilizing a specific weight matrix would exhibit superior performance on the dataset produced by that same weight matrix. For the datasets created by weight matrices W_G, W_P, and W_C, we found that models incorporating spatially random effects generated by the corresponding weight matrices would perform best.

Figure 3.7: Trace plots for the weight specification W_B

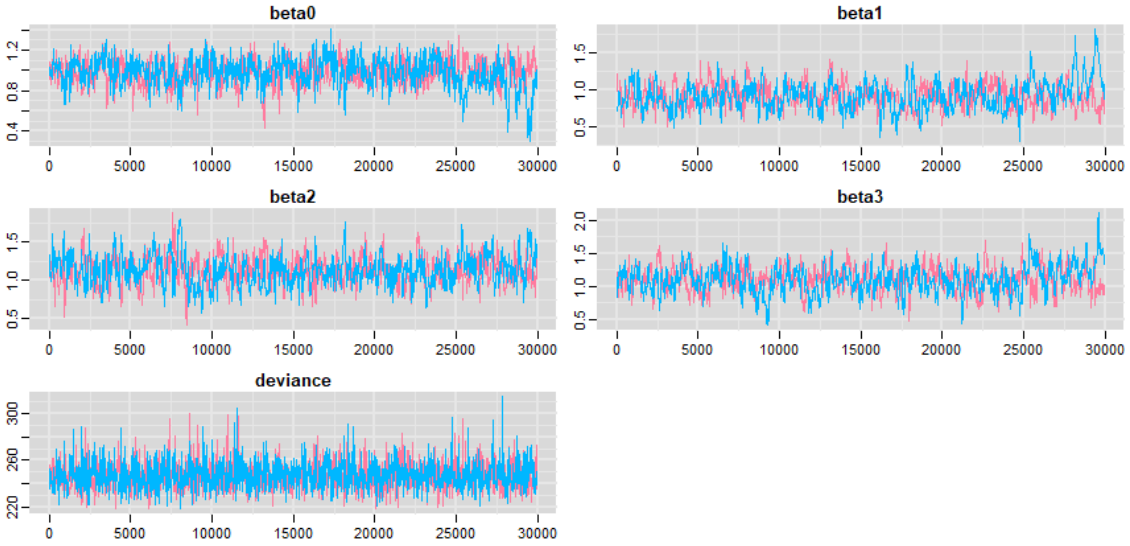


Figure 3.8: Trace plots for the weight specification W_E

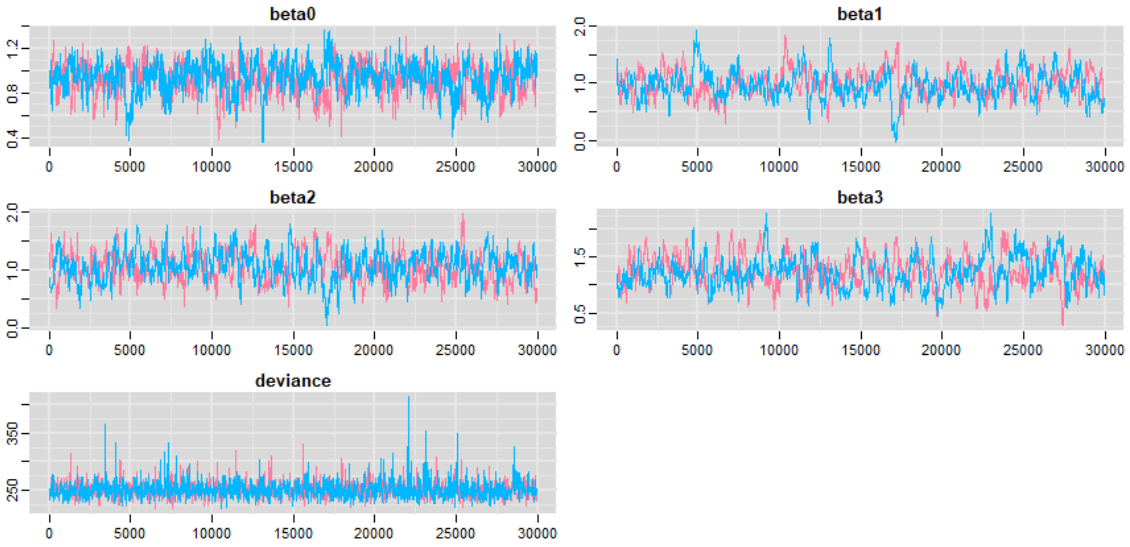


Figure 3.9: Trace plots for the weight specification W_G

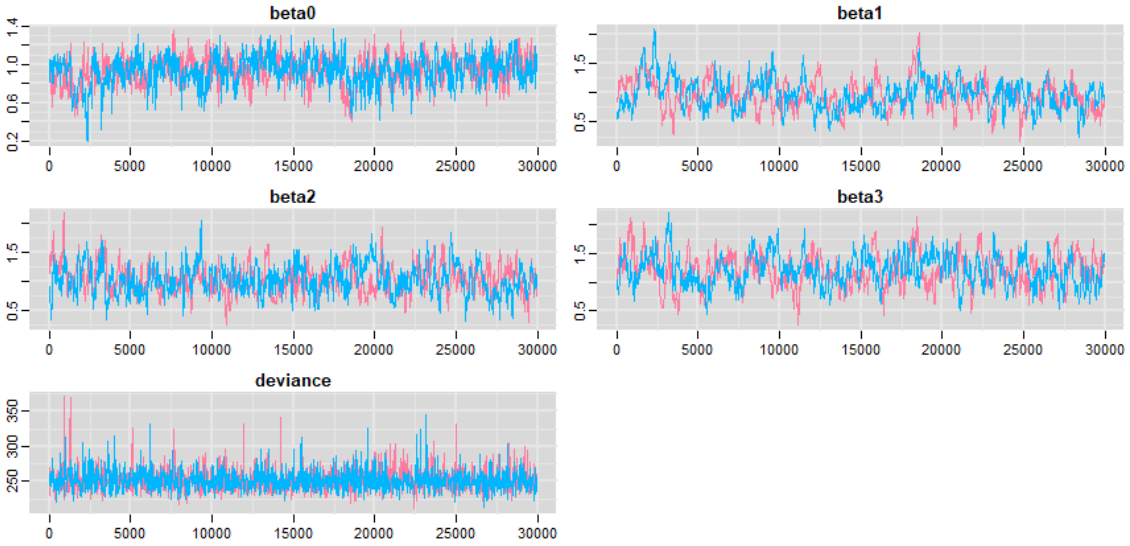


Figure 3.10: Trace plots for the weight specification W_P

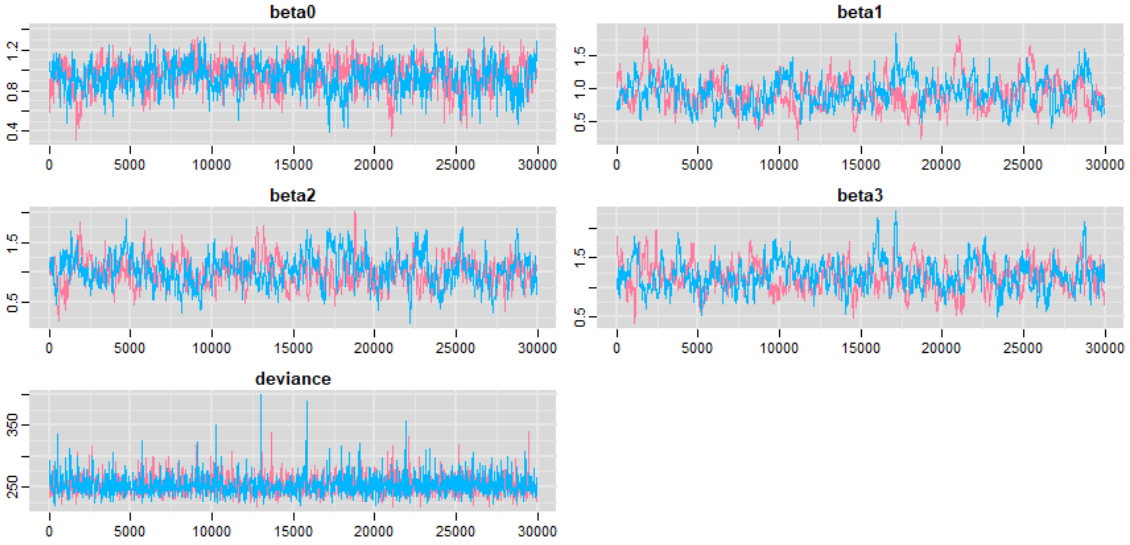
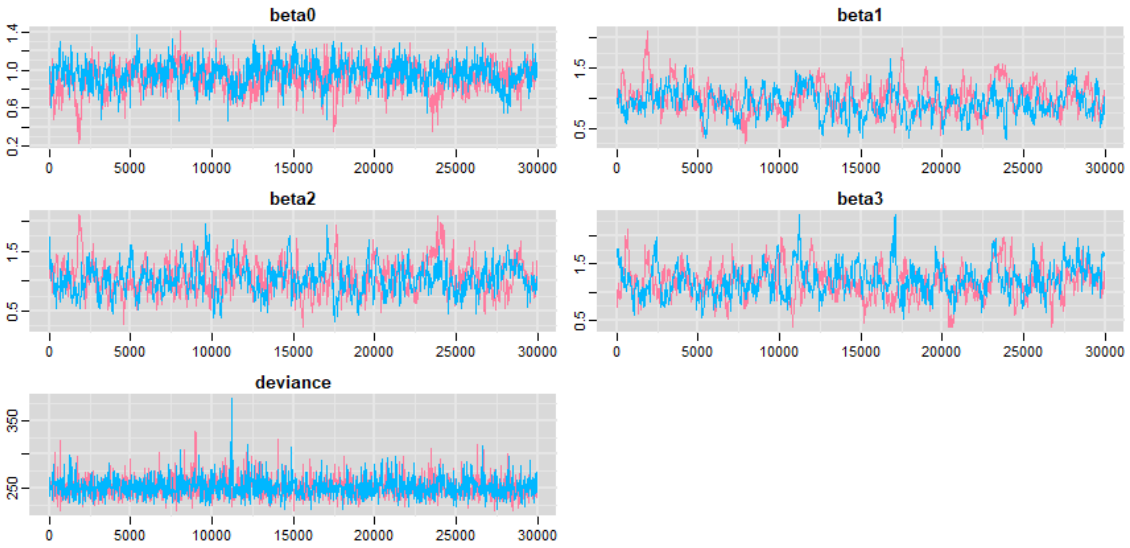


Figure 3.11: Trace plots for the weight specification $\boldsymbol{W}_C$



Chapter 4

Applied project-The spatial distribution of opioid overdose death in Michigan

4.1 Introduction

Opioids refer to a group of compounds that act on opioid receptors (Trescot, Datta, Lee and Hansen, 2008) in nerve cells within the body and brain (National Institute on Drug Abuse, 2019). These compounds can be categorized into several chemical classes: Phenanthrenes, Benzomorphans, Phenylpiperidines, Diphenylheptanes, and Phenylpropanolamines. Opioids encompass a powerful class of medications used for short-term management of acute pain following injuries or surgeries. These drugs include Morphine, Oxycodone, Hydrocodone, Hydromorphone, Fentanyl, and the illegal drug heroin. Opioid prescriptions are only useful for treating some forms of chronic pain temporarily. Opioids include side effects that include addiction, sleepiness, sedation, and overdose when used for extended periods of time. Since the brain adjusts to repeated exposure to opioids, regular treatment

under a doctor's prescription can result in dependence and addiction. This can affect anyone. The National Survey on Drug Use and Health report's findings indicate that over 2.1 million people had either abused or were dependent on opioids. Drug overdoses are the leading cause of death for individuals under 50. Over the past two decades, there has been a steady increase in opioid overdose deaths, peaking in 2013. In comparison to 1999, the incidence of drug overdose deaths related to prescription opioids has multiplied by five in 2017. Notably, 80 percent of heroin users had previously misused prescription opioids before turning to the illicit narcotic. Michigan has been significantly impacted by the opioid epidemic, with overdose deaths increasing 17-fold between 1999 and 2016. In 2015, Michigan ranked 10th in the number of opioid prescriptions written, while the state's opioid overdose death rate stood as the 15th highest in the United States. Astonishingly, Michigan's annual opioid prescriptions reached 11 million in both 2015 and 2016, surpassing an average of one prescription per person. Furthermore, each year, an average of 84 opioid pills are consumed per person. Simultaneously, Michigan's drug overdose deaths have surged since 2012. In 2015, the Michigan Department of Health and Human Services reported that drug overdose deaths had exceeded fatalities from gun-related incidents and traffic accidents. The objectives of this study were to identify the spatial distribution of opioid overdose death in the lower peninsula of Michigan and also the focus of the present study was the comparison of alternate weight matrices for the analysis.

4.2 Data

The study region, which has an area of 40,162 sq.miles, is the the lower peninsula of Michigan. The opioid-related drug overdose deaths data was collected by

the Michigan Death Certificates, specifically the Division for Vital Records and Health Statistics/MDHHS. The opioid-related deaths include overdose deaths related to Hydrocodone, Hydromorphone, Morphine, Codeine, Fentanyl, and heroin. These drug overdose deaths encompass fatalities resulting from unintentional or intentional overdose of a drug, administration of the wrong drug, drug consumption errors, or inadvertent drug intake. Only cases meeting all overdose definitions and involving at least one type of opioid from the list above were counted. To calculate the expected value of opioid overdose deaths in each county, the population of Michigan was necessary. The data available at the Annual Estimates of the Resident Population from the U.S. Census Bureau, Population Division was used. For the opioid overdose and population data, the information from 2017 was utilized. All data were aggregated at the county level, with consideration given to the counties located in the lower peninsula of Michigan due to missing values in the upper peninsula.

Covariates for regression analysis to identify potential risk factors of the opioid overdose deaths are restricted by those available from the 2017 Census. Totally, three covariates related to socioeconomics were tested: 1) poverty rate; 2) education rate; 3) unemployment rate. The values are expressed as percentages.

4.3 Modeling methods

We used Bayesian spatial approach including prior distributions and several alternate specifications of the spatial weights matrix. We briefly describe the two Poisson regression models used in this study: 1) a standard Poission model with uncorrelated random effects(nonspatial model) and 2) Poisson Regression Model with uncorrelated random effects and correlated random effects with diffenrent

weight matrix for ICAR distribution(BYM model). For all approaches it was assumed that, the number of opioid overdose deaths in each of the 68 counties of the lower peninsula of Michigan, $y_1, \ldots, y_{68}$, follows Poisson distribution with mean μ_i, where $\mu_i = e_i\theta_i$, θ_i is the relative risk in ith county, and e_i is the expected count of opioid overdose deaths in ith country calculated by the equation 2.1. Let x_{1i} be the Poverty rate (Persons below 150% of poverty) in ith county, x_{2i} be the Education rate (percent high school graduate or higher) in ith county and x_{3i} be the Unemployment rate in ith county.

4.3.1 Nonspatial Poisson model

A nonspatial Poisson model for the number of opioid overdose deaths can be specified as

$$y_i \sim Poisson(\mu_i), \quad i = 1, \ldots, 68, \tag{4.1}$$

$$log(\mu_i) = log(e_i) + \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + v_i, \quad i = 1, \ldots, 68, \tag{4.2}$$

where α is an intercept term, $\beta_1, \beta_2, \beta_3$, are regression coefficients, x_{1i}, x_{2i}, x_{3i} are a set of covariates, and v_i is uncorrelated random effects specified by $v_i \sim N(0, \sigma_v^2)$.

4.3.2 BYM model

The BYM model take the following form:

$$y_i \sim Poisson(\mu_i), \quad i = 1, \ldots, 68, \tag{4.3}$$

$$log(\mu_i) = log(e_i) + \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + v_i + u_i, \quad i = 1, \ldots, 68, \tag{4.4}$$

where u_i is correlated random effects which follows the intrinsic conditional autoregressive (ICAR) distribution specified by

$$u_i \mid u_{-i} \sim N(\frac{1}{\sum_{j\in\delta_i} w_{ij}} \sum_{j\in\delta_i} w_{ij}u_j, \frac{1}{\sum_{j\in\delta_i} w_{ij}}\sigma_u^2), \quad i = 1, \ldots, N. \tag{4.5}$$

The u_{-i} denotes all spatially correlated random effects except u_i, δ_i is the set of neighbors of the ith area, and w_{ij} is the entry of a symmetric weight matrix W corresponding to row i and column j.

We used five diffenrent weight matrices for ICAR distribution: 1) binery weight matrix $\boldsymbol{W}_B$ specified by Equition 1.6; 2) inverse distance weight matrix with Euclidean distance $\boldsymbol{W}_D$ specified by Equition 1.7; 3) inverse distance weight matrix with Graph distance$\boldsymbol{W}_G$ specified by Equition 1.8; 4) weight matrix of number of Paths between two counties $\boldsymbol{W}_P$ specified by Equition 1.9; 5) inverse distance weight matrix with Graph distance and the number of Paths between two counties $\boldsymbol{W}_C$ specified by Equition 1.10.

For each of α, β_1, β_2, and β_3, we use Gaussian prior distributions $N(0, \sigma_\alpha^2)$, $N(0, \sigma_{\beta_1}^2)$, $N(0, \sigma_{\beta_3}^2)$, and $N(0, \sigma_{\beta_j}^2)$, respectively, where the hyperparameters σ_α, σ_{β_1}, σ_{β_2}, and σ_{β_3} were assumed to have a Uniform(0,10) distribution. The hyperparameters σ_v and σ_u follow $Unif(0, 10)$ and $Gamma(0.01, 0.01)$, respectively.

4.4 Model evaluation

The models with the alternative weight matrices were assessed for goodness-of-fit with the observed data and the criterion employed for assessment was the Deviance Information Criterion (DIC) developed by Spiegelhalter et al., (2002). DIC provides the measure of complexity and fit of the models which are developed from same sample size. A smaller DIC indicates a better model fit. DIC is computed

as the sum of the posterior mean deviance and estimated effective number of parameters:

$$DIC = \bar{D} + P_D, \tag{4.6}$$

where $\bar{D}$ is the sum of the posterior mean deviance which measures how well the model fits the data and P_D represents the effective number of parameters utilized for model building. The data was processed in R version 4.0.3 and the Bayesian analyses were performed in WinBUGS 1.4.3.

4.5 Results

In our research, we examined two distinct models Nonspatial Poisson model included covariates and uncorrelated random effects. BYM model included covariates, uncorrelated random effects, and spatially correlated random effects which follows ICAR distribution with the weight matrix specification $\boldsymbol{W}_B$, $\boldsymbol{W}_D$, $\boldsymbol{W}_G$, $\boldsymbol{W}_P$, and $\boldsymbol{W}_C$ respectively.

Table 4.1 shows that the BYM model with $\boldsymbol{W}_C$ has the lowest DIC (442.375) among the models that we examined in this research.

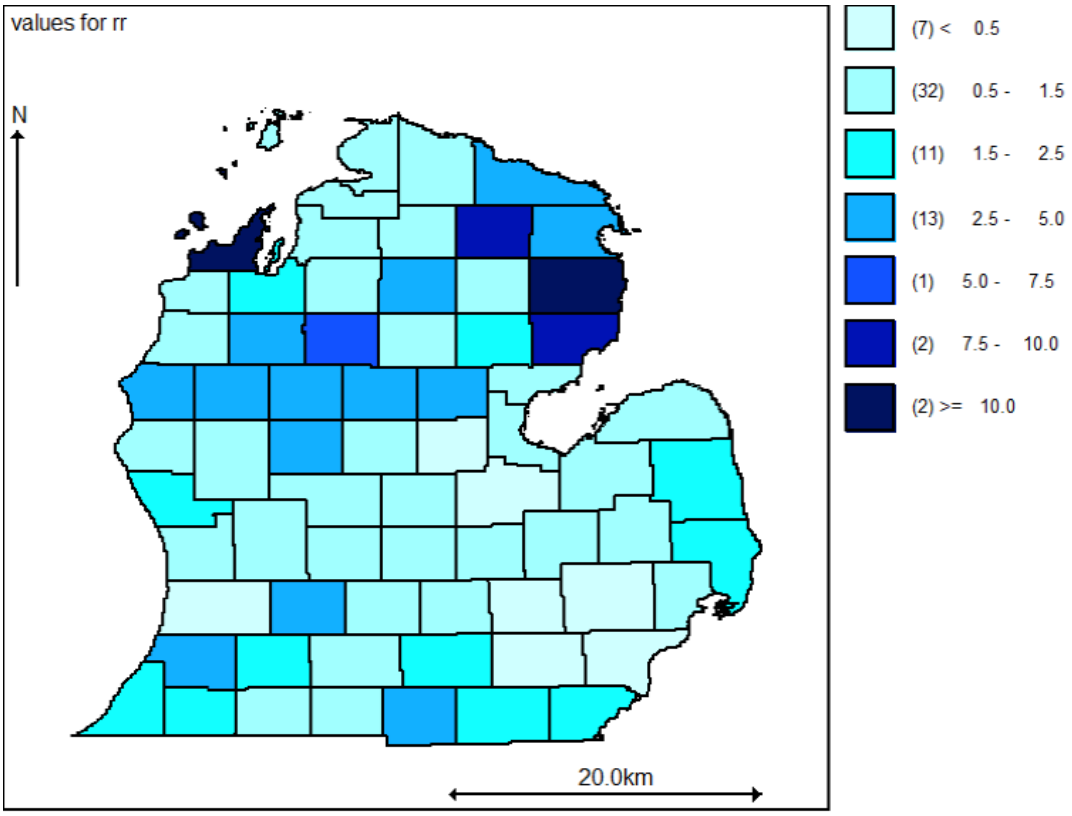
Convergence was monitored by visual examination of the trace plots of the samples for each chain and the Gelman-Rubin convergence statistics. Convergence occurred by 100,000 iterations.

Figure 4.1 shows the estimated relative risk map based on the result of the model under inverse distance weight matrix with Graph distance and the number of Paths between two counties, $\boldsymbol{W}_C$. In the map, we found that there is spatial variation of opioid overdoes death in the lower peninsula of Michigan. Deeper color explains higher variation. The spatial clusters of high overdose deathes are located

Table 4.1: Camparision DIC

Model	Var	$Mean$	SD	95%CI	DIC
Nonspatial	int.	1.161	0.386	(0.681, 1.066)	447.72
	Pov.	0.014	0.013	(-0.011, 0.039)	
	Ed.	-0.021	0.008	(-0.039, -0.005)	
	Unemp.	0.046	0.025	(-0.054, 0.142)	
W_B	int.	1.457	0.619	(0.472, 1.768)	446.319
	Pov.	0.013	0.018	(-0.022, 0.049)	
	Ed.	-0.022	0.010	(-0.043, -0.000)	
	Unemp.	0.006	0.069	(-0.135, 0.143)	
W_D	int.	-0.039	0.707	(-1.595, 1.309)	447.584
	Pov.	0.021	0.021	(-0.018, 0.070)	
	Ed.	-0.018	0.012	(-0.042, 0.008)	
	Unemp.	0.051	0.067	(-0.082, 0.179)	
W_G	int.	1.227	0.602	(0.256, 1.543)	445.920
	Pov.	0.014	0.020	(-0.024, 0.053)	
	Ed.	-0.022	0.011	(-0.047, -0.002)	
	Unemp.	0.045	0.062	(-0.067, 0.178)	
W_P	int.	1.329	0.712	(0.285, 1.955)	446.710
	Pov.	0.011	0.021	(-0.033, 0.053)	
	Ed.	-0.025	0.013	(-0.055, -0.002)	
	Unemp.	0.0512	0.069	(-0.084, 0.189)	
W_C	int.	1.123	0.553	(0.267, 1.400)	442.375
	Pov.	0.016	0.017	(-0.0189, 0.049)	
	Ed.	-0.021	0.011	(-0.044, -0.000)	
	Unemp.	0.051	0.061	(-0.065, 0.173)	

Figure 4.1: Mapping based on the model under inverse distance weight matrix with graph distance and the number of paths between two counties, W_C.



in eastern and northern regions of the state.

The trace plots in Figure 4.2 generated for the monitored intercept, coefficient and deviance across the two chains. From the trace plots, we observe that the chains appear to mix rapidly, which is an indication of convergence.

Figure 4.3 shows the Gelman-Rubin Convergence Statistic Diagnostic Plots, which indicates a good convergence. We found that $\hat{R}$, which is the red line in Figure 4.3 is close to 1. A sample of 100,000 iterations with two chains was obtained after the burn-in period of 10,000 samples.

Figure 4.2: Trace Plots for the monitored intercept and deviance for the model 6

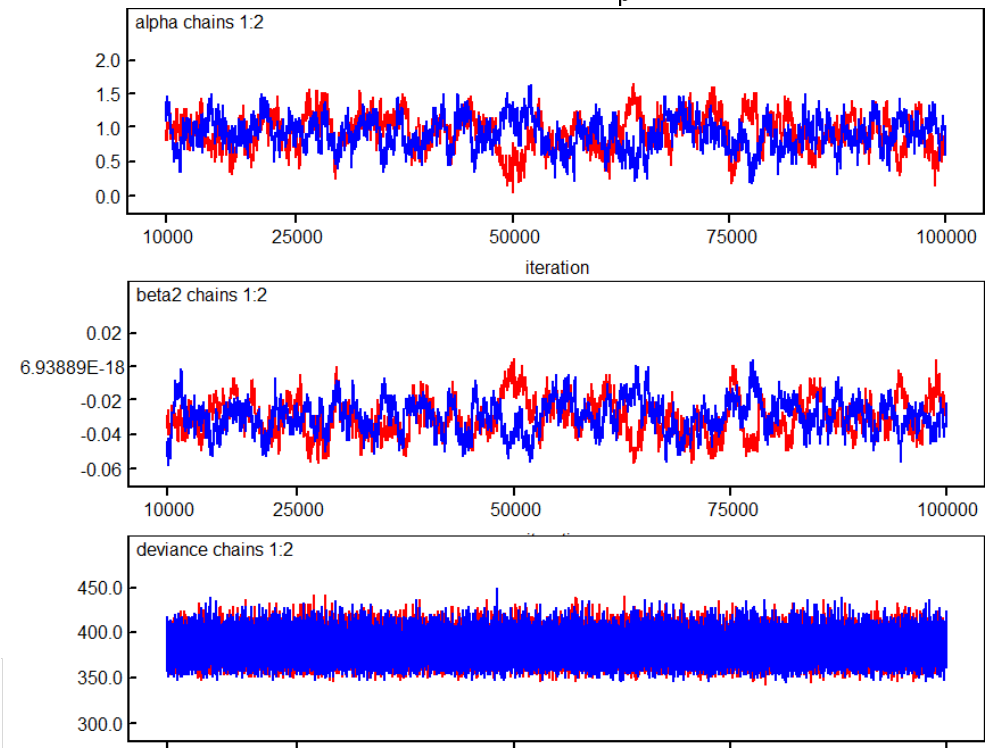
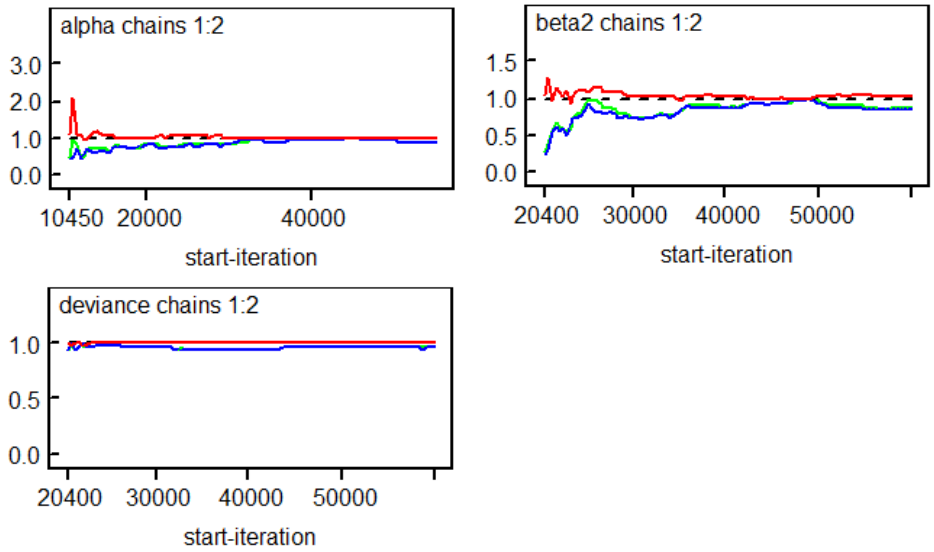


Figure 4.3: Gelman-Rubin Convergence Statisti Diagnostic Plots for the monitored intercept and deviance for the model 6



4.6 Discussion

For the spatial distribution of opioid overdose deaths, we performed a standard Poission model with uncorrelated random effects (nonspatial model) and a Poisson Regression Model with uncorrelated random effect and correlated random effect which follows ICAR distribution with five different weight matix specifications (BYM model). The selected model contains the spatially correlated random effect with the weight matrix created by Graph distance and the number of paths between counties. We found that there is spatial variation of opioid overdose death in the lower peninsula of Michigan. Specifically, counties in eastern and northern Michigan experience higher levels of such fatalities. The counties located in those regions are recommended to be targeted to create opioid prevention strategies, such as recovery services, and increasing the supply of overdose-reversing drugs.

Chapter 5

Conclusions

In this paper, we performed Hierarchical Poisson models for spatial count data with different specification of weight matrices that control the behaviour and degree of spatial smoothing meaning that data points are averaged with their neighbours. The simulation results indicated that all models that we performed in this paper have a good fit of the simulated data with spatially correlated random effect. Especially the Poisson models with a spatially correlated random errect which follows CAR distribution with the weight matrix specificatrions $\boldsymbol{W}_G$, $\boldsymbol{W}_P$, and $\boldsymbol{W}_C$ had better performance than the Poisson models with a spatially correlated random errect which follows CAR distribution with the binary adjacency wight matrix or the inverse Euclidean distance matrix in the simulated data based on the competing methods by checking the accuracy of the estimated coefficients though mean and mean square error (MSE). Hence, the weight matrix specificatrions $\boldsymbol{W}_G$, $\boldsymbol{W}_P$, and $\boldsymbol{W}_C$ appear to a good choice for implementing spatial smoothing and generally perform quite well and offer similar parameter interpretations.. For future work, we can perform a different version of the model in Equation (2.3) using the Laplace distribution instead of Normal distribution in Equation (2.5). For the CAR Laplace prior, we

have

$$u_i \mid u_{-i} \sim L(\frac{1}{\sum_{j\in\delta_i} w_{ij}} \sum_{j\in\delta_i} w_{ij}u_j, \frac{1}{\sum_{j\in\delta_i} w_{ij}}\sigma_u^2), \quad i = 1, \ldots, N \tag{5.1}$$

, where u_{-i} denotes all spatially correlated random effects except u_i, δ_i is the set of neighbors of the ith area, and w_{ij} is the entry of a symmetric weight matrix W corresponding to row i and column j using the following probability density function,

$$L(x|\mu, b) = \frac{1}{2b} exp(-\frac{|x-\mu|}{b}). \tag{5.2}$$

Here μ is a location parameter and $b > 0$ is a scale parameter. Also, we can create a distance based weight matrix of estimated driving distance between two counties using GPS or Google Maps Platform APIs.

Appendix A

Background for Graph Theory

One of the purpose of this study is create new specifications of weight matrix. For this purpose, we bring some definitions from Graph Theory. The set of areas of a map can be represented as a (undirected) graph that has a vertex for each area and an edge for every pair of areas that share a boundary. Figure A.1 shows the county map of the lower peninsula of Michigan and the corresponding graph in red.

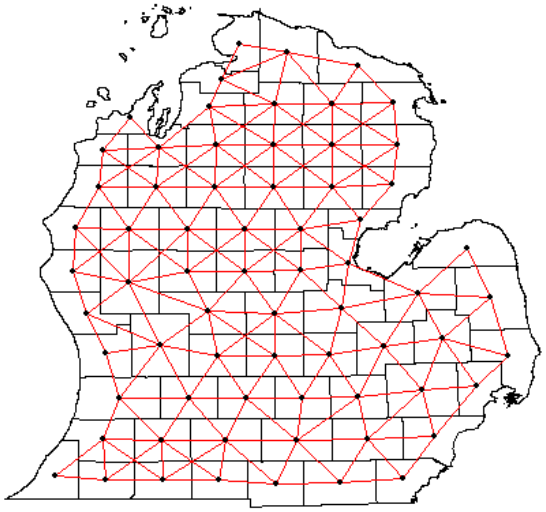
Definition 1. *A walk is defined as a sequence of edges which joins a sequence of vertices. A path is a walk in which all edges and vertices are distinct. The total number of edges covered in a walk is called as length of the walk. A shortest path is a path which has least number of edges. The distance between two vertices in a graph is the length of a shortest path between them.*

By Definition 1, every path is a walk but every walk is not a path.

Theorem 1. *Every walk between two vertices contains a path between them (West, D. B 2000).*

Definition 2. *For a graph with vertex set* $U = \{u_1, \ldots, u_n\}$*, the adjacency matrix is a square* $n \times n$ *matrix A such that its* $(i, j)^{th}$ *entry* a_{ij} *is one when there is an*

Figure A.1: County map of the lower peninsula of Michigan and the corresponding graph in red



edge from vertex u_i to vertex u_j, and zero when there is no edge.

Theorem 2. *The $(i,j)^{th}$ entry a_{ij}^k of A^k, where A the adjacency matrix, counts the number of walks of length k betwean vertex u_i and vertex u_j (Andrew Duncan 2004).*

Theorem 3. *If k is the distance betwean vertex u_i and vertex u_j, then the $(i,j)^{th}$ entry a_{ij}^k of A^k, where A the adjacency matrix, counts the number of shortest paths betwean vertex u_i and vertex u_j.*

Proof. By Theorem 2, there are k walks between vertex u_i and vertex u_j. We want to prove that the k walks are paths. To prove it by contradiction assume that there is a walk but not a path among the k walks between vertex u_i and vertex u_j. Then the walk contains a path between vertex u_i and vertex u_j and its length is shorter than k. Then the distance between vertex u_i and vertex u_j is shorter than k. Then we arrive to a contradiction. □

Appendix B

WinBUGS Codes

B.1 Poisson model without random effect

```
model
{
   for( i in 1 : m ) {
      y[i] ~ dpois(mu[i])
      mu[i] <-  rr[i]
      log(rr[i]) <- beta0+beta1*x1[i]+beta2*x2[i]+beta3*x3[i]
r[i]<-(y[i]-mu[i])
   }

sdbeta0~dunif(0,10)
varbeta0<-sdbeta0*sdbeta0
taubeta0<-1/varbeta0
beta0~dnorm(0,taubeta0)

sdbeta1~dunif(0,10)
varbeta1<-sdbeta1*sdbeta1
```

```
taubeta1<-1/varbeta1
beta1~dnorm(0,taubeta1)

sdbeta2~dunif(0,10)
varbeta2<-sdbeta2*sdbeta2
taubeta2<-1/varbeta2
beta2~dnorm(0,taubeta2)

sdbeta3~dunif(0,10)
varbeta3<-sdbeta3*sdbeta3
taubeta3<-1/varbeta3
beta3~dnorm(0,taubeta3)
```

B.2 Poisson model with random effects

```
model
{
   for( i in 1 : m ) {
      y[i] ~ dpois(mu[i])
      mu[i] <-  rr[i]
      log(rr[i]) <- beta0+beta1*x1[i]+beta2*x2[i]+beta3*x3[i]+u[i]
r[i]<-(y[i]-mu[i])
   }

u[1:m] ~  car.normal(adj[], w[], num[],tau.u)

sdbeta0~dunif(0,10)
varbeta0<-sdbeta0*sdbeta0
taubeta0<-1/varbeta0
```

```
beta0~dnorm(0,taubeta0)


sdbeta1~dunif(0,10)
varbeta1<-sdbeta1*sdbeta1
taubeta1<-1/varbeta1
beta1~dnorm(0,taubeta1)


sdbeta2~dunif(0,10)
varbeta2<-sdbeta2*sdbeta2
taubeta2<-1/varbeta2
beta2~dnorm(0,taubeta2)


sdbeta3~dunif(0,10)
varbeta3<-sdbeta3*sdbeta3
taubeta3<-1/varbeta3
beta3~dnorm(0,taubeta3)


tau.u ~ dgamma(0.01, 0.01)
```

Appendix C

R Codes for graph distance and the number of paths

We created the matrix of Graph distance between counties in Michigan and the matrix of the number of paths between counties in Michigan using the adjacent matrix for counties in Michigan. Therefore, the input of the R functions below is the adjacent matrix for Michigan.

C.1 Graph distance

```
interm <- function(x){
  x1 <- x
  for (j in 1:nrow(x)){
    for(i in 1:nrow(x)){
      a <- which(x1[i,]>j-1)
      b <- as.numeric(sapply(apply(cbind(x[,a],NULL),
      1, sum), as.logical))
      b[i]<-0
      b[which(x1[i,]>0)]<-0
```

```
      b[a] <- 0
      x1[i,] <- x1[i,]+(j+1)*b
    }
    if(all(x1[!diag(nrow(x1))] != 0)){break}
  }
  x1
}
```

C.2 The number of paths

```
npath <- function(y){
  interm <- function(x){
    x1 <- x
    for (j in 1:nrow(x)){
      for(i in 1:nrow(x)){
        a <- which(x1[i,]>j-1)
        b <- as.numeric(sapply(apply(cbind(x[,a],NULL),
        1, sum), as.logical))
        b[i]<-0
        b[which(x1[i,]>0)]<-0
        b[a] <- 0
        x1[i,] <- x1[i,]+(j+1)*b
      }
      if(all(x1[!diag(nrow(x1))] != 0)){break}
    }
    x1
  }
  c <- interm(y)
  j = 1
```

```
  t = rep(0,length(c[upper.tri(c, diag = FALSE)]))
  for (i in c[upper.tri(c, diag = FALSE)]){
    a <- y %^% i
    t[j] <- a[upper.tri(a, diag = FALSE)][j]
    j <- j+1
  }
  c[upper.tri(c, diag = FALSE)] = t
  c[lower.tri(c)]  <- t(c)[lower.tri(c)]
  print(c)
}
```

References

Besag, J. (1974), Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36: 192-225. https://doi.org/10.1111/j.2517-6161.1974.tb00999.x

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43, 1-20.

Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions, *Biometrika*, Volume 82, Issue 4, Pages 733 746, https://doi.org/10.1093/biomet/82.4.733

Bivand, Roger. (2014). SPDEP: Spatial Dependence: Weighting Schemes, Statistics and Models.

Bloom, J. (2017). Comparison of frequentist and Bayesian inference.

Brownstein, J. S., Green, T. C., Cassidy, T. A., & Butler, S. F. (2010). Geographic information systems and pharmacoepidemiology: using spatial cluster detection to monitor local patterns of prescription opioid abuse. *Pharmacoepidemiology and drug safety*, 19(6), 627–637.

Cai, B., Lawson, A. B., Hossain, M., Choi, J., Kirby, R. S., & Liu, J. (2013). Bayesian semiparametric model with spatially-temporally varying coefficients selection. *Statistics in medicine*, 32(21), 3670–3685.

Carroll, R., Lawson, A. B., Faes, C., Kirby, R. S., Aregay, M., & Watjou, K. (2015). Comparing INLA and OpenBUGS for hierarchical Poisson modeling in disease mapping. *Spatial and spatio-temporal epidemiology*, 14-15, 45–54.

Carroll, R., Lawson, A. B., Faes, C., Kirby, R. S., Aregay, M., & Watjou, K. (2016). Bayesian model selection methods in modeling small area colon cancer incidence. *Annals of epidemiology*, 26(1), 43–49.

Carroll, R., Lawson, A. B., Faes, C., Kirby, R. S., Aregay, M., & Watjou, K. (2018). Spatially-dependent Bayesian model selection for disease mapping. *Statistical methods in medical research*, 27(1), 250–268.

Choi, J., & Lawson, A. B. (2018). Bayesian spatially dependent variable selection for small area health modeling. *Statistical methods in medical research*, 27(1), 234–249. https://doi.org/10.1177/0962280215627184

Choi, J., & Lawson, A. B. (2019). A Bayesian two-stage spatially dependent variable selection model for space-time health data. *Statistical methods in medical research*, 28(9), 2570–2582.

Clayton, D., & Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43(3), 671–681.

Dellaportas, P., Forster, J. J., & Ntzoufras, I. (2002). On Bayesian model and variable selection using MCMC. *Statistics and computing*, 12(1), 27-36.

De Oliveira, Victor. (2012). Bayesian analysis of conditional autoregressive

models. *Annals of the Institute of Statistical Mathematics*. 64. 107-133. 10.1007/s10463-010-0298-1.

Duncan, A.J. (2004). Powers of the adjacency matrix and the walk matrix.

Duncan, E. W., White, N. M., & Mengersen, K. (2017). Spatial smoothing in Bayesian models: a comparison of weights matrix specifications and their impact on inference. *International journal of health geographics*, 16(1), 47. https://doi.org/10.1186/s12942-017-0120-x

Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical methods for digital computers*, 191-203.

Getis, A. (2009). Spatial weights matrices. *Geographical Analysis*, 41(4), 404+.

Getis, A. and Aldstadt, J. (2004), Constructing the Spatial Weights Matrix Using a Local Statistic. *Geographical Analysis*, 36: 90-104. https://doi.org/10.1111/j.1538-4632.2004.tb01127.x

Gill, Gurdiljot & Sakrani, T. & Cheng, W. & Zhou, J.. (2017). COMPARISON OF ADJACENCY AND DISTANCE-BASED APPROACHES FOR SPATIAL ANALYSIS OF MULTIMODAL TRAFFIC CRASH DATA. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-2/W7. 1157-1161.

Ginestet, Cedric. (2011). Bayesian Decision-theoretic Methods for Parameter Ensembles with Application to Epidemiology.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4), 711-732.

Law J. (2016). Exploring the Specifications of Spatial Adjacencies and Weights in Bayesian Spatial Modeling with Intrinsic Conditional Autoregressive Priors in a Small-area Study of Fall Injuries. *AIMS public health*, 3(1), 65–82. https://doi.org/10.3934/publichealth.2016.1.65

Lawson, A. B. (2013). Statistical methods in spatial epidemiology. *John Wiley & Sons*.

Lawson, A. B., Browne, WJ., & Vidal Rodeiro, CL. (2003). Disease mapping with WinBUGS and MLwiN. *Wiley-Blackwell*.

Lawson, A. B., Song, H. R., Cai, B., Hossain, M. M., & Huang, K. (2010). Space-time latent component modeling of geo-referenced health data. *Statistics in medicine*, 29(19), 2012–2027. https://doi.org/10.1002/sim.3917

Markus J. Fülle & Philipp Otto. (2023) Spatial GARCH models for unknown spatial locations – an application to financial stock returns. *Spatial Economic Analysis* 0:0, pages 1-14.

Modarai, F., Mack, K., Hicks, P., Benoit, S., Park, S., Jones, C., Proescholdbell, S., Ising, A., & Paulozzi, L. (2013). Relationship of opioid prescription sales and overdoses, North Carolina. *Drug and alcohol dependence*, 132(1-2), 81–86. https://doi.org/10.1016/j.drugalcdep.2013.01.006.

NIDA. 2021, June 1. Prescription Opioids DrugFacts. Retrieved from https://nida.nih.gov/publications/drugfacts/prescription-opioids on 2024, February 13

Onicescu, G., & Lawson, A. B. (2018). Bayesian cure-rate survival model with spatially structured censoring. *Spatial statistics*, 28, 352–364.

Onicescu, G., Lawson, A., Zhang, J., Gebregziabher, M., Wallace, K., & Eberth, J. M. (2018). Spatially explicit survival modeling for small area cancer data. *Journal of applied statistics*, 45(3), 568–585.

Onicescu, G., Lawson, A. B., Zhang, J., Gebregziabher, M., Wallace, K., & Eberth, J. M. (2019). Spatially-explicit survival modeling with discrete grouping of cancer predictors. *Spatial and spatio-temporal epidemiology*, 29, 139–148. https://doi.org/10.1016/j.sste.2018.06.001

Otto, P., & Steinert, R. (2023). Estimation of the Spatial Weighting Matrix for Spatiotemporal Data under the Presence of Structural Breaks. *Journal of Computational and Graphical Statistics*, 32(2), 696–711. https://doi.org/10.1080/10618600.2022.2107530

Shen, Y. (2019). A Two-Stage Bayesian Variable Selection Method with the Extension of Lasso for Geo-Referenced Count Data. *Dissertations. 3471.*

Sillanpää, M. J., & Bhattacharjee, M. (2005). Bayesian association-based fine mapping in small chromosomal segments. *Genetics*, 169(1), 427–439.

Sillanpää, M. J., & Corander, J. (2002). Model choice in gene mapping: what and why. *Trends in genetics* : TIG, 18(6), 301–307.

Sommer, S., & Wade, T. (2006). A to Z GIS: an illustrated dictionary of geographic information systems. Esri Press.

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, 12(3), 1–16.

Trescot, A. M., Datta, S., Lee, M., & Hansen, H. (2008). Opioid pharmacology. *Pain physician*, 11(2 Suppl), S133–S153.

Waller, L. A. (2005). Bayesian thinking in spatial statistics. *Handbook of Statistics*, 25, 589-622.

West, Douglas. (2000). Introduction to Graph Theory (2nd Edition).

Xia, H., Carlin, B. P., & Waller, L. A. (1997). Hierarchical models for mapping Ohio lung cancer rates. *Environmetrics: The official journal of the International Environmetrics Society*, 8(2), 107-120.

Yokoi, Takahisa & Ando, Asao. (2008). Spatial structures in a spatial autoregressive model and one-directional adjacency matrices. *The Japanese Journal of Real Estate Sciences*. 21. 115-125.

Zhang, C. (2015). Spatial Weights Matrix and its Application.

Zhang, Xinyu & Yu, Jihai. (2017). Spatial weights matrix selection and model averaging for spatial autoregressive models. *Journal of Econometrics*. 203. 10.1016/j.jeconom.2017.05.021.

Zou, H. (2005). Some perspectives of sparse statistical modeling. *Stanford University*.