2011

Bayesian 2-Stage Space-Time Mixture Modeling with Spatial Misalignment of the Exposure in Small Area Health Data

Andrew B. Lawson
Medical University of South Carolina

Jungsoon Choi
Medical University of South Carolina

Bo Cai
Medical University of South Carolina

Md. Monir Hossain
Medical University of South Carolina

Russell S. Kirby
Medical University of South Carolina

See next page for additional authors

Authors

Andrew B. Lawson, Jungsoon Choi, Bo Cai, Md. Monir Hossain, Russell S. Kirby, and Jihong Liu

Paper Number/ Resource Identifier: 11-007

Date: 2011

MUSC Author(s): Lawson, Andrew B, Choi, Jungsoon

Paper Title: Bayesian 2-stage space-time mixture modeling with spatial misalignment of the exposure in small area health data

Complete Author List: Lawson, Andrew B.; Choi, Jungsoon; Cai, Bo; Hossain, Md. Monir; Kirby, Russell S.; Liu, Jihong

Bayesian 2-stage space-time mixture modeling with spatial misalignment of the exposure in small area health data

Andrew B. Lawson[1,*], **Jungsoon Choi**[1], **Bo Cai**[2], **Md. Monir Hossain**[3], **Russell S. Kirby**[4], **and Jihong Liu**[2]

[1]Division of Biostatistics and Epidemiology, College of Medicine, Medical University of South Carolina, Charleston, SC 29403, USA

[2]Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC 29208, USA

[3]Biostatistics, Epidemiology and Research Design (BERD) Core, Center for Clinical and Translational Sciences, The University of Texas, the University of Texas Health Science Center at Houston, Houston, TX 77030, USA

[4]Department of Community and Family Health, College of Public Health, University of South Florida, Tampa, FL, USA

*email: lawsonab@musc.edu

The authors thank the National Institutes of Health (5R21HL088654-02) for support of this work.

Bayesian 2-stage space-time mixture modeling with spatial misalignment of the exposure in small area health data

Abstract

We develop a new Bayesian two-stage space-time mixture model to investigate the effects of air pollution on asthma. The two-stage mixture model proposed allows for the identification of temporal latent structure as well as the estimation of the effects of covariates on health outcomes. In the paper, we also consider spatial misalignment of exposure and health data. A simulation study is conducted to assess the performance of the 2-stage mixture model. We apply our statistical framework to a county-level ambulatory care asthma data set in the US state of Georgia for the years 1999-2008.

Key words: Space-time mixture model; air pollution; covariate adjustment; asthma; Bayesian modeling

1 Introduction

Respiratory diseases such as asthma, chronic obstructive pulmonary disease (COPD), and bronchitis are important health problems in the United States. In 2008, it was estimated that more than 23 million Americans have asthma and approximately 13 million adults have COPD (Centers for Disease Control and Prevention, 2008; Pleis *et al.*, 2009). In addition, respiratory diseases have a high cost in medical expenses. For example, the annual cost of asthma associated with medical expenses was estimated at about \$50.1 billion in 2007 (Centers for Disease Control and Prevention, 2011). Thus, finding the risk factors of respiratory diseases is important to policy-makers and program planners wishing to reduce incidence.

Numerous epidemiologic studies have found the risk factors that showed significant association with asthma, which is a common chronic disease in the US, about 1% of all ambulatory visits (Dockery and Pope, 1994; Ponka and Virtanen, 1996; Eisner *et al.*, 2001; Ellison-Loschmann *et al.*, 2007). For example, socioeconomic and ethnic characteristics such as income and African-American race have been linked with greater risks of asthma (Eisner *et al.*, 2001; Ellison-Loschmann *et al.*, 2007). Elevated concentrations of air pollutants (e.g. particulate matter and ozone) have been shown to be associated with increased incidence of asthma (Stieb *et al.*, 1996; Lin *et al.*, 2002; Sheppard, 2003; Lin *et al.*, 2008).

Recently, the study of the association between $\text{PM}_{2.5}$ known as fine PM (ambient particles less than or equal to $2.5\mu m$ in diameter) and asthma has received much attention in public health studies (e.g. Freidman *et al.*, 2001; Sheppard, 2003). However, most researches have been conducted using time-series analysis in specific locations because $\text{PM}_{2.5}$ data are available only in the limited locations. In addition, $\text{PM}_{2.5}$ concentrations and asthma data are collected over space and time so the relative risks of asthma may have space-time dependence structures and the association between $\text{PM}_{2.5}$ exposure and asthma may vary across space and time. Thus, spatiotemporal analysis of the association between $\text{PM}_{2.5}$ exposure and asthma is important and necessary.

In most environmental health effects studies, relative risk within the fixed space and time period is modeled using a linear function of air pollutants and covariates as well as space-time random effects. The coefficients of risk factors are constructed in various ways depending on the modeling approach (e.g. constant, space-varying, or space-time varying coefficients). Along with this coefficient structure, the relative risk model includes a function of space-time random effects (Bernardinelli *et al.*, 1995; Xia *et al.*, 1997; Knorr-Held and Besag, 1998; Knorr-Held, 2000; Mugglin, *et al.* 2002; Richardson *et al.*, 2006; Tzala and Best, 2008). A commonly-used approach (global modeling) has space, time, and space-time interaction random components in risk, and each random component explains the overall risk effect over their space-time domain (Knorr-Held, 2000). However, temporal risk effects, for instance, can vary within the space-time domain, and a subset of spatial areas can have a homogeneous temporal profile in risk. In this situation, global modeling is not appropriate because it has the restrictive assumption of common risk effects across all areas.

Recently, Lawson *et al.* (2010) developed Bayesian space-time latent models using mixture structures in order to capture the heterogeneous temporal profiles of relative risks in space-time health data. They also proposed the use of entry parameters in the space-time mixture (STM) model for the estimation of the number of the underlying temporal risk patterns. Choi *et al.* (2011) evaluated the performance of STM models in terms of a range of measures and also compared space-time Dirichlet process mixture models with the STM models. They found that STM models are better than Dirichlet process mixture models in terms of recovery of spatial clustering of temporal profiles and how well they estimate the true number of latent temporal components.

When incorporating space-time varying risk factors such as air pollution and socioeconomic factors in space-time health modeling where space-time random effects are included, confounding bias problems may arise (Reich *et al.*, 2006; Ma *et al.*, 2007; Hodges and Reich, 2010; Paciorek, 2010). For example, air pollution varying spatiotemporally may be correlated with the space-time random effects so the confounding bias in estimating the effects of air pollution on health outcomes can appear in the model. However, there are a few statistical studies related to this bias problem in spatial models (Clayton *et al.*, 1993; Hodges and Reich, 2010; Paciorek, 2010). In the STM models, space-time varying risk factors on health outcomes may be correlated with the locally varying temporal risk patterns so it can be difficult to exactly estimate both the effects of the risk factors on the outcomes and the underlying temporal components.

In this paper, we introduce a Bayesian 2-stage space-time mixture model to reduce confounding bias, which provides better estimates of the association between exposure to fine PM and health outcomes as well as the underlying temporal components. This method first obtains Poisson residuals from the covariates-only model and then using these residuals as inputs, a space-time mixture structure is fitted to find the locally different temporal components. From the estimation of the mixture structure and covariate information, the effects of covariates on health outcomes are finally estimated. We evaluate this approach using a simulation study in terms of recovering the coefficients of covariates and latent components. We also compare the 2-stage mixture model with the full space-time mixture model where relative risk is expressed as both a function of covariates and a space-time structure, in order to investigate how they bias the estimated health risks and they estimate the number of latent component. We conduct an analysis of the relationship between ambulatory care visits for asthma and exposure to $\mathrm{PM}_{2.5}$ and socioeconomic factors. Since we have different sources of PM and health data, a “change of support”problem needs to be considered (Gotway and Young 2002; Banerjee *et al.*, 2004; Fuentes *et al.*, 2006). Thus, we consider a space-time model for $\mathrm{PM}_{2.5}$ to provide county-level estimates of $\mathrm{PM}_{2.5}$, which allows for the estimation of the effects of fine PM on asthma outcomes. This work presented here is the first attempt to consider confounding bias problems in space-time models and then gain better estimates of the coefficients of space-time varying covariates and the true latent components, by introducing a 2-stage mixture model.

The remainder of this paper is organized as follows. In Section 2, we describe the asthma data, air pollution data, and socioeconomic data used in this study. In Section 3 we present the space-time $PM_{2.5}$ model and the 2-stage space-time mixture model. In Section 4 a simulation study is performed to verify the performance of the 2-stage space-time mixture model in comparison with the full space-time mixture model. In Section 5, the data analysis results from the 2-stage mixture model proposed are provided. Finally, a general discussion of our approach is provided in Section 6.

2 Data Description

In the paper, we use county-level counts of ambulatory case sensitive asthma in the state of Georgia USA, for the year 1999 to 2008, which were obtained from the Georgia health information system OASIS (`http://oasis.state.ga.us/`), Georgia Division of Public Health. There are 159 counties and 10 time periods (years) in the available data. We used standardization to provide expected counts within counties for each time period. The expected count was calculated by using the internal standardization method (Banerjee *et al.*, 2004) based on the statewide population-based rate. Figure 1 displays the standardized incidence ratios for asthma for each year where the standardized incidence ratio is defined as the count of asthma divided by the expected count. Overall, the standardized incidence ratios are high in the south-east areas of Georgia over years.

We use a $PM_{2.5}$ data set where $PM_{2.5}$ is the air quality standard set by the U.S. Environmental Protection Agency (EPA). The $PM_{2.5}$ data set from the Federal Reference Method (FRM) monitoring network was used. There are 31 monitoring stations in Georgia for the period 1999-2008. Originally, $PM_{2.5}$ concentrations were measured either every day, every third day, or every sixth day and the yearly averaged $PM_{2.5}$ values at each station were used in this study. Figure 2 (a) presents the map of PM stations and Figure 2 (b) shows the temporal plots of $\mathrm{PM}_{2.5}$ for the selected two stations. Two stations have the decreasing temporal patterns of $\mathrm{PM}_{2.5}$, but the station located in the urban area (A) has high $\mathrm{PM}_{2.5}$ concentrations over time. As covariates in PM modeling, we also consider yearly-averaged weather variables such as temperature (°F), dew point temperature (°F), and wind speed (miles per hour to tenths) from the U.S. National Climate Data Center. The coverage of th monitoring stations reflects the population concentrations and is relatively sparse in the more rural areas. This means that interpolation of effects will lead to less variation in estimated mean level in such areas.

County-level socioeconomic census data for year 2000 and estimated data for the other years are obtained from the Area Resource File (ARF) from the U.S. Department of Health and Human Services (`http://arf.hrsa.gov`). The ARF is a collection of county-level data sets from more than 50 sources such as American Hospital Association, American Medical Association, National Center for Health Statistics, and US Census Bureau. It contains a wide range of information and includes county level geographic information, socioeconomic, and environmental characteristics. Based on previous study and considering the availability of county-level data, the variables we considered as relevant predictors are: the proportion of black people (the black or African American population divided by total population), median household income (unit: $1000), and unemployment rate, as covariates in the health model (Castro *et al.*, 2001; Eisner *et al.*, 2001; Ellison-Loschmann *et al.*, 2007). Unemployment rate data is also available at the US Bureau of Labor Statistics (`http://www.bls.gov`).

3 Models

Our environmental health framework has two main parts due to a "change of support" problem. In the first part, we estimate the county-level $\mathrm{PM}_{2.5}$ concentrations using a space-time $\mathrm{PM}_{2.5}$ model, which are used as the inputs for the health model in the next part. In the second part, we introduce a 2-stage space-time mixture modeling for asthma and air pollution, along with socioeconomic covariates in order to investigate the association between asthma and exposure to $\mathrm{PM}_{2.5}$ as well as the estimation of temporal risk profiles. This approach is the type of 'Directional' Bayesian approach (Gelman, 2004), mainly used for computational reasons. Gelman (2004) presented the computational and practical benefits for this plug-in approach in comparison to a joint model. Unike a joint model, the approach does not allow the health data to influence the air pollution modeling, which might be seen to be a reasonable approach. Therefore, the posterior distributions are obtained separately at each stage. Of course, measurement error in the plug-in estimates can accommodate some of the biases.

3.1 Spatio-temporal model of exposure

We consider a space-time model of $\mathrm{PM}_{2.5}$ introduced by Fuentes *et al.* (2006) and Choi *et al.* (2009). We assume that $\mathrm{PM}_{2.5}(\mathbf{s}_m, t_j)$ is the yearly-averaged $\mathrm{PM}_{2.5}$ concentration at station $\mathbf{s}_m$ $(m = 1, \cdots, M)$ and time t_j $(j = 1, \cdots, J)$ and is not the "true" PM value because of measurement error. Thus, the $\mathrm{PM}_{2.5}$ model is given by

$$\mathrm{PM}_{2.5}(\mathbf{s}_m, t_j) = Z(\mathbf{s}_m, t_j) + \epsilon_1(\mathbf{s}_m, t_j),$$

where $Z(\mathbf{s}_m, t_j)$ is the unobserved "true" PM process at space $\mathbf{s}_m$ and time t_j and $\epsilon_1(\mathbf{s}_m, t_j) \sim N(0, \sigma_1^2)$ is the measurement error. We model the true process $Z(\mathbf{s}_m, t_j)$ as

$$Z(\mathbf{s}_m, t_j) = \mu_z(\mathbf{s}_m, t_j) + \epsilon_2(\mathbf{s}_m, t_j),$$

where $\mu_z(\mathbf{s}_m, t_j)$ is the mean function and $\boldsymbol{\epsilon}_2{}^T = (\epsilon_2^T(\mathbf{s}_1), \cdots, \epsilon_2^T(\mathbf{s}_M))$, where $\epsilon_2^T(\mathbf{s}_m) = (\epsilon_2(\mathbf{s}_m, t_1), \cdots, \epsilon_2(\mathbf{s}_m, t_J))$, has a multivariate normal distribution with mean zero and space-time covariance function $\mathbf{\Sigma}_Z$. The mean function $\mu_z(\mathbf{s}_m, t_j)$ can be modeled with coordinates or meteorological variables. In this study, the mean function is assumed to be $\mu_z(\mathbf{s}_m, t_j) = \mathbf{W}^T(\mathbf{s}_m, t_j)\boldsymbol{\eta}_Z$, where $\mathbf{W}(\mathbf{s}_m, t_j)$ is a vector of coordinates (longitude and latitude) and meteorological variables (temperature, dew point temperature, and wind speed) with corresponding coefficient vector $\boldsymbol{\eta}$. Based on exploratory analysis and previous research (Choi *et al.*, 2009), we use the separable space-time covariance $\boldsymbol{\Sigma}_Z = \sigma_e^2 H_s(\phi) \otimes H_t(\rho)$ where $\otimes$ denotes the Kronecker product. The matrix $H_s(\phi)$ is $M \times M$ with $(H_s(\phi))_{mm'} = \exp\left(-\phi||\mathbf{s}_m - \mathbf{s}_{m'}||\right)$ and $H_t(\rho)$ is $J \times J$ with $(H_t(\rho))_{jj'} = \rho^{|t_j - t_{j'}|}/(1-\rho^2)$, where $\phi \sim \text{Unif}(0.01, 20)$ and $\rho \sim \text{Beta}(1,1)$, that is uniform on (0,1).

The posterior estimate of true $\text{PM}_{2.5}$ at unobserved site $\mathbf{s}_0$ and time t_j is calculated using Markov Chain Monte Carlo (MCMC) algorithms from the posterior predictive distribution of $Z(\mathbf{s}_0, t_j)$ given the observed information,

$$p(Z(\mathbf{s}_0, t_j)) = \int p(Z(\mathbf{s}_0, t_j)|\text{PM}_{2.5}, \mathbf{M}, \Theta_z) p(\Theta_z|\text{PM}_{2.5}, \mathbf{M}) d\Theta_z,$$

where Θ_z is a set of all parameters included in the PM model, and σ_1^2 and σ_e^2 have uniform prior distributions (Gelman, 2006). The "true" $\text{PM}_{2.5}$ of county i at time t_j of interest (Z_{ij}) is defined as

$$Z_{ij} = \frac{1}{|B_i|} \int_{B_i} Z(\mathbf{s}_m, t_j) d\mathbf{s} \tag{1}$$

where B_i is the spatial domain within a county i. The estimate of Z_{ij} (Z_{ij}^*) is the average of estimates of true $\text{PM}_{2.5}$ at several locations randomly selected within a county i at time t_j. We could consider block Kriging or MC integration for estimation of Z_{ij}.We have chosen the latter for convenience as this can be achieved by averaging simulated point level predictions at random locations within each county. We have found this approach to be reasonably accurate compared to block Kriging in preliminary evaluation studies.

3.2 Health model: 2-stage space-time mixture modeling

In space-time epidemiological studies, little is known about the impact of space-time random effects on the health effect of spatiotemporally varying covariates. Commonly-used approach to the space-time association between covariates and human health outcomes is a Poisson regression model where risk is modeled as a linear function of covariates and space-time random effects. However, this full space-time modeling may cause confounding problems, not distinguishing the effects of covariates from unmeasured space-time random effects. In space-time mixture modeling, it is important to estimate the effects of covariates as well as the space-time mixture structure. Thus, we propose a 2-stage space-time mixture model. The value of this model lies in the ability to provide estimates of spatial disaggregation of risk while also providing good overall description of risk variation (Lawson *et al.*, 2010).

Denote the count of disease in the ith area at the t_jth time period as y_{ij}, where $i = 1, \cdots, I$ and $j = 1, \cdots, J$. We make the conventional assumption that y_{ij} follows a Poisson distribution as

$$y_{ij} \sim Pois(e_{ij}\theta_{ij}),$$

where e_{ij} is the observed expected count and θ_{ij} is the relative risk.

In the first-stage, the effects of covariates (PM$_{2.5}$ and socioeconomic factors) are only considered in the log-relative risk model:

$$\log \theta_{ij} = \alpha_0 + Z_{ij}^{*}\gamma_{ij} + \mathbf{X}_{ij}^{T}\boldsymbol{\beta}_{ij}, \tag{2}$$

where α_0 is the intercept parameter. The value Z_{ij}^{*} is the estimate of the "true" unobserved county-level PM$_{2.5}$ from the exposure model presented in Section 3.1 and the corresponding parameter γ_{ij} can be considered in various dependence structures. The vector $\mathbf{X}_{ij}$ includes p socioeconomic covariates of area i at time t_j with the corresponding parameter vector $\boldsymbol{\beta}_{ij} = (\beta_{ij1}, \cdots, \beta_{ijp})^T$. The parameters γ_{ij} and $\boldsymbol{\beta}_{ij}$, for example, can be assumed to be space-time dependent structures as follows:

$$\begin{aligned} \gamma_{ij} &= \gamma_0 + \gamma_i^1 + \gamma_j^2 \\ \beta_{ijp} &= \beta_{0p} + \beta_{ip}^1 + \beta_{jp}^2, \end{aligned} \tag{3}$$

where γ_0 and β_{0p} are the overall mean parameters of the coefficients over space and time, γ_i^1 and β_{ip}^1 are the spatially correlated components, and γ_j^2 and β_{jp}^2 are the temporally correlated components.

This covariates-only model provides the estimated relative risk $\hat{\theta}_{ij}$ from the posterior distribution using a Bayesian approach. The Poisson residuals using these estimates and the data are calculated as

$$\hat{r}_{ij} = \log\left(y_{ij}/e_{ij}\right) - \log\hat{\theta}_{ij}.$$

These residuals are used for the estimation of space-time mixture structures.

In the second-stage, we assume that the Poisson residual model is

$$\hat{r}_{ij}|\hat{\theta}_{ij},\ y_{ij},\ e_{ij} \sim \mathrm{N}(\alpha_r + \Lambda_{ij}, \sigma^2_{r_{ij}}), \tag{4}$$

where $\sigma^2_{r_{ij}}$ is the variance and α_r is the intercept to explain the overall difference between the $\log(y_{ij}/e_{ij})$ and the estimated log relative risk. The component Λ_{ij} represents a space-time random effect. Following Lawson *et al.* (2010) and Choi *et al.* (2011), the Λ_{ij} is modeled as a space-time mixture structure:

$$\begin{aligned} \Lambda_{ij} &= \sum_{l=1}^{L} w_{il}\chi_{lj}, \\ w_{il} &= \frac{\psi_l w_{il}^*}{\sum_l \psi_l w_{il}^*}, \qquad w_{il}^* \geq 0, \end{aligned}$$

where L is assumed to be a large value to estimate the “true” number of latent components. The latent component χ_{lj} represents the underlying temporal profile in relative risk by specifying a time-dependent structure, and w_{il} is the corresponding weight at area i. The weight $w_{il} \geq 0$ is the proportion of component l at area i so the sum of all weights for each area should be one and the weight w_{il} is expressed using unstandardized weight $w_{il}^* \geq 0$. We model w^*_{il} as a log-normal distribution with spatially dependent mean ξ_{il} and variance $\sigma^2_{w_l}$

$$\begin{aligned} w^*_{il} &\sim \mathrm{LN}(\xi_{il}, \sigma^2_{w_l}), \\ \xi_{il} &\sim \mathrm{MIAR}(\Sigma_\xi). \end{aligned}$$

The mean ξ_{il} has a multivariate intrinsic autoregressive (MIAR) distribution (Gelfand and Vounatsou, 2002) with cross-covariance function Σ_ξ, which is a relatively smooth spatial process:

$$\xi_{il}|\xi_{i'l,\ i'\neq i} \sim \mathrm{N}\Big(\frac{1}{N_i}\sum_{i'\neq i} G_{ii'}\xi_{i'l}, \frac{1}{N_i}\boldsymbol{\Sigma}_\xi\Big),$$

where $G_{ii'} = 1$ if area i is adjacent to area i', and $G_{ii'} = 0$ otherwise. Also, $N_i = \sum_{i'\neq i} G_{ii'}$ is the number of neighbors of area i. This multivariate spatial process allows for both the spatial dependence structure of the weights and the dependence structure between the different weights given neighboring sites.

The entry parameter ψ_l has a value of 0 or 1 and controls whether the lth temporal component is included in the model or not. If $\psi_l = 1$, then the lth temporal component is involved in the model. Otherwise, the lth temporal component is not involved in the model. In this study, the entry parameter is assumed to have a Bernoulli distribution, $\psi_l \sim \mathrm{Bern}(0.5)$, where 0.5 is a non-informative value.

By fitting the residual model in Equation (4) the estimated temporal components and weights ($\widehat{\chi}_{lj}$ and $\widehat{w}_{il}$) are obtained. We adjust the temporal components using the intercept α_r, $\widehat{\chi}^*_{lj} = \widehat{\alpha}_r + \widehat{\chi}_{lj}$, to improve the estimation performance. These estimates along with covariate information are used as the inputs in the final model expressed as

$$\log(\theta_{ij}) = \alpha_0 + Z^*_{ij}\gamma_{ij} + \mathbf{x}^T_{ij}\boldsymbol{\beta}_{ij} + \sum_{l=1}^{L} \widehat{w}_{il}\widehat{\chi}^*_{lj} + \eta_{ij}, \tag{5}$$

where α_0, γ_{ij}, and $\boldsymbol{\beta}_{ij}$ are parameters for estimation and have the same structures as those of the covariates-only model in Equation (2) . The random component $\eta_{ij} \sim \mathrm{N}(0, \sigma^2_\eta)$ is the uncorrelated space-time interaction term. This restricted Poisson regression model provides the final estimates for α_0, γ_{ij}, and $\boldsymbol{\beta}_{ij}$, which are our main focus.

To conduct the spatial allocation of the temporal components in the 2-stage space-time mixture model, a post-processing method based on the posterior distributions of the weights is considered. The spatial cluster indicator C_i $(= 1, \cdots, L)$ is defined as

$$C_i = \arg\max_{l}\{w_{il}\}. \tag{6}$$

This indicator C_i has the label index of the temporal component with the largest weight value in area i. Thus, a subset of areas within the space-time domain is assigned to one of the temporal components included in the model, which represents the principal temporal profile of the area in relative risk.

In the covariates-only model in Equation (2) and the restricted regression model in Equation (5), the prior distributions of the intercept and the overall mean parameters in the coefficients are specified as $\alpha_0 \sim \mathrm{N}(0, \sigma^2_{\alpha_0})$, $\gamma_0 \sim \mathrm{N}(0, \sigma^2_{\gamma_0})$, and $\beta_{0p} \sim \mathrm{N}(0, \sigma^2_{\beta_{0p}})$. We use an intrinsic autoregressive (IAR) distribution for the spatial components γ^1_i and β^1_{ip} (Besag *et al.*, 1991), that corresponds to a univariate spatial process ($L = 1$) in the MIAR distribution. The temporal components γ^2_j and β^2_{jp} are assigned to be random walk Gaussian distributions. All the standard deviation parameters in the models have uniform prior distributions (Gelman, 2006). For both models, the likelihoods are defined as

$$\begin{aligned} p(\mathbf{y}|\Theta_1) &= \prod_{i=1}^{I}\prod_{j=1}^{J} \mathrm{Pois}(y_{ij}|e_{ij}, \alpha_0, \gamma_{ij}, \boldsymbol{\beta}_{ij}), \\ p(\mathbf{y}|\Theta_3) &= \prod_{i=1}^{I}\prod_{j=1}^{J} \mathrm{Pois}(y_{ij}|e_{ij}, \alpha_0, \gamma_{ij}, \boldsymbol{\beta}_{ij}, \eta_{ij}, \widehat{w}_{il}, \widehat{\chi}^*_{lj}), \end{aligned}$$

where Θ_1 and Θ_3 are the sets of the parameters in the covariates-only model and the restricted regression model, respectively. Based on the likelihood and the prior distributions, the posterior distributions of the parameters Θ_1 and Θ_3 are obtained.

For the Poisson residual model in Equation (4), the likelihood is derived as

$$p(\hat{r}_{ij}|\Theta_2) = \prod_{i=1}^{I}\prod_{j=1}^{J} \mathrm{N}(\hat{r}_{ij}|\alpha_r, w_{il}, \chi_{lj}, \sigma^2_{r_{ij}})$$

where $\alpha_r \sim \mathrm{N}(0, \sigma^2_{\alpha_r})$ and $\sigma_{r_{ij}}$ is assigned to be a uniform distribution. The covariance matrix of the MIAR distribution ($\mathbf{\Sigma}_\xi$) is specified as an inverse Wishart prior distribution, Inv-Wishart$((0.01 I_L)^{-1}, L)$ and I_L is the $L \times L$ identity matrix. In this study, the temporal component χ_{lj} has an AR(1) model and each temporal parameter has a beta prior distribution, Beta(1,1), that is uniform on (0,1). Similarly, the posterior distribution of all the parameters Θ_2 is obtained based on the likelihood and the prior distributions. For the estimation of all the parameters in these models, the Gibbs sampling algorithm and the Metropolis adaptive rejection sampling algorithm are implemented. The posterior means are used for the estimation of all the parameters except the cluster indicator C_i while the posterior mode is used for the estimation of C_i because C_i is the nominal value.

An identifiability problem of components in Bayesian space-time mixture modeling can appear because of the invariance of the likelihood under the permutation of the component labels (Stephens, 2000; Jasra *et al.*, 2005). We assume that the latent components in the proposed model follow temporally correlated structures while the corresponding weights follow spatially correlated structures. Moreover, during MCMC simulation, it could be possible that the components switch labels if multiple chains are used (Choi *et al.*, 2011). In this study, a single chain is used to avoid the label switching problem.

4 Simulation Study

In this section we perform a simulation study to compare the 2-stage space-time mixture model proposed in the previous section with the full mixture model where risk is modeled as a linear function of covariates and a space-time mixture structure. We examine the performance of the 2-stage mixture model by investigating the capability of recovering the true coefficients and the true space-time mixture structure.

We simulate data under three designs. In all the designs, the 159 counties of the state of Georgia are used as a space domain because there are many counties with similar spatial shapes in Georgia and we conduct real data analysis within this spatial domain in Section 5. As the time domain, $J = 10$ time points are used. All the designs have $L = 3$ temporal components and the spatial design of the cluster indicator (C_i) created in Georgia (Figure 3 (a)). Each spatial cluster is assigned to one temporal component χ_{lj} that has an AR(1) structure with the temporal parameter ρ_l and the standard deviance 0.1. To make the components different, the temporal parameters are specified as $\boldsymbol{\rho} = (0.9, 0.7, 0.5)$ and the true temporal profiles are presented in Figure 3 (b).

In Design 1, two covariates (X_{1ijk} and X_{2ijk}) for county i and time j of the kth simulation data ($k = 1, \cdots, K$) are considered, where X_{1ijk} is generated as $X_{1ijk} \sim \mathrm{N}(0, 1)$, independent over space, time, and simulation, and X_{2ijk} is generated from the IAR prior distribution with the overall variance 1, independent over time and simulation. Thus, X_{2ijk} has a spatial dependence structure while X_{1ijk} has no spatial dependence structure. We generate simulated count y_{ijk} as follows:

$$\begin{aligned} y_{ijk} &\sim \mathrm{Pois}(e_{ijk}\theta_{ijk}), \qquad k = 1, \cdots, K, \\ \log(\theta_{ijk}) &= \beta_0 + \beta_1 X_{1ijk} + \beta_2 X_{2ijk} + \chi_{C_i,j} + \eta_{ijk}, \end{aligned}$$

where $\beta_0 = 1$, $\beta_1 = 0.05$, and $\beta_2 = 0.1$. The expected count e_{ijk} is generated independently from the uniform distribution, Unif(15, 20), and the random effect η_{ijk} is generated as $\eta_{ijk} \sim N(0, 0.01^2)$.

For Designs 2 and 3, the true relative risks are assumed to be constant over simulations but the simulated counts are different

$$\begin{aligned} y_{ijk} &\sim \text{Pois}(e_{ij}\theta_{ij}), \qquad k = 1, \cdots, K, \\ \log(\theta_{ij}) &= \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \chi_{C_i,j} + \eta_{ij}, \end{aligned}$$

where β's have the same values with the Design 1, and X_{1ij} and η_{ij} are generated from the same scheme as the Design 1. Here, we assume the second covariate X_{2ij} varies over space and time, and X_{2ij} is generated from the normal distribution with mean 0 and space-time covariance $\Sigma_{X_2} = 0.1\Sigma_S \otimes \Sigma_T$, where Σ_S and Σ_T are the covariance matrices of the IAR prior distribution and the AR(1) distribution. Designs 2 and 3 have values of 0.8 and 0.2 for the temporal parameter in Σ_T, respectively.

For each design we generate $K = 200$ data sets. For each simulated data, we fit two models: the full space-time mixture model and the 2-stage mixture model. We use $L = 6$ entry parameters in fitting the models because the true number of components is 3 and $L = 6$ is enough to estimate the true number of components. After fitting the models, we determine whether a temporal component is included in the model using the estimated corresponding entry parameter. If the estimated entry parameter is larger than 0.5, then the component is included in the model. Thus, the estimated number of components included in the model is computed. In addition, the identification of the estimated temporal components with the true temporal components is required when the estimated number of components is three, because the label switching problem can arise (Stephens, 2000; Jasra *et al.*, 2005). For the allocation of the estimated components and their labels, the mean square error method is used

$$\widehat{\mathcal{C}} = \arg\min_{l'} \sum_{l=1}^{L} \sum_{j=1}^{J} (\hat{\chi}_{l'j} - \chi_{lj})^2,$$

where $\widehat{\mathcal{C}}$ includes the labels of the estimated components corresponding to the true components.

For each simulated data set and each model we compute posterior means as point estimates and 95% intervals for β_0, β_1, and β_2. For the comparison, we use mean squared error (MSE) and mean absolute error (MAE). Mean squared error and mean absolute error for β_p (p=0, 1, 2) are calculated as

$$\begin{aligned}\text{MSE} &= \frac{1}{KIJ}\sum_{k=1}^{k}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\hat{\beta_p}^{(k)} - \beta_p\right)^2 \\ \text{MAE} &= \frac{1}{KIJ}\sum_{k=1}^{k}\sum_{i=1}^{I}\sum_{j=1}^{J}\left|\hat{\beta_p}^{(k)} - \beta_p\right|\end{aligned}$$

where $\hat{\beta_p}^{(k)}$ is the estimate of the true β_p for the kth simulation.

Table 1 presents the results for the coefficients: the average of the estimates over simulation (mean), the 2.5th and 97.5th percentiles of the estimates, the averaged widths of 95% intervals over simulation, MSE and MAE. Both models have similar results of β_1 associated with no space-time varying covariate. Overall, the averages of the estimates for β_0 and β_2 in the 2-stage mixture model are closer to the true values in comparison with the full mixture model. Also, the 2-stage models have the smallest MSE and MAE, which justifies that the 2-stage mixture models estimate the true coefficients very well. In some cases, the 2.5th and 97.5th percentiles of the estimates in the 2-stage model do not include the true coefficient values, but the averaged widths of 95% intervals for the coefficients in the 2-stage model are much smaller than those in the full mixture model. Thus, these results suggest that the 2-stage mixture models are better than the full mixture models in terms of recovering the true coefficients. Especially, the 2-stage mixture models dramatically improve the performance of the estimates of the intercept and coefficients associated with space (or space-time) varying covariates.

Table 2 summarizes the estimated number of components included in the models by using a percentage table. Clearly, the 2-stage mixture models estimate the true number of components very well while the full mixture models estimate the small number of components. It is shown that 93%, 87%, and 94.5% of the simulations in the 2-stage models estimate the exact true number of components in Designs 1-3, respectively. However, the full mixture models estimate the true number of components with less than 25% of the simulations (10.5% of the simulations in Designs 1 and 3, 23.0% in Design 2). In estimating the true number of temporal components, the 2-stage mixture models are much better than the full mixture models.

In Figure 4, the plots of the true temporal components and their estimates with 95% credible intervals in Designs 1 and 2 are displayed using only the output when the models estimate the exact true number of components. As you can see the plots, all the intervals of the estimated temporal components from the 2-stage models contain the true profiles while the intervals for Component 2 from the full models do not include the true ones. Overall, the widths of the intervals in the 2-stage models are smaller than those in the full mixture models. Design 3 also has similar results. This suggest that the 2-stage mixture models fit the true temporal components well.

Finally, we examine the performance of spatial clustering in both models with the outputs when the estimated number of components is equal to the true number of components. To check the ability of the models in detecting the spatial clusters, we use the accuracy cluster rate, $A = \sum_{i=1}^{I} A_i/I$ and $A_i = \sum_{k=1}^{K} \mathbf{I}(C_i^T = \hat{C}_{ik})/K$, where $\mathbf{I}(\cdot)$ is the indicator function, C_i^T is the true spatial cluster indicator for county i and $\hat{C}_{ik}$ is the estimated cluster indicator for the ith county at the kth simulation. This accuracy measure explains how well the model recovers the true spatial clusters over space and simulation. In Design 1, the 2-stage mixture models (0.59) provide higher A value than the full mixture model (0.55). In Designs 2 and 3, the full mixture models have a little bit higher A values than the 2-stage models, but a quite small output from the full mixture models is only used to compute the accuracy rate in comparison with the 2-stage mixture models. Thus, there is no big difference between both models in terms of recovering the true spatial clusters.

5 Real data analysis

We apply our statistical framework to data in Georgia for the years 1999-2008 described in Section 2. We first analyze monitored $\mathrm{PM}_{2.5}$ data using the space-time PM model proposed in Section 3.1 to produce the estimated county-level $\mathrm{PM}_{2.5}$ concentrations. Using these $\mathrm{PM}_{2.5}$ estimates and asthma data, the 2-stage mixture model is fitted to investigate the effects of air pollution on asthma and examine the space-time mixture structure.

For the PM model and the health model we use a single chain with a total of 70,000 iterations to satisfy convergence criteria. The number of iterations for the burnin period is 20,000, and the thinning rate is 10 so the number of samples used for the estimation of the parameters is 5000. MCMC convergence diagnostics using the Geweke convergence diagnostic (Geweke, 1992), autocorrelation functions, and trace plots are conducted. The deviance and several representative parameters meet acceptable MCMC convergence.

Figure 5 presents the maps of the estimated county-level $\mathrm{PM}_{2.5}$ concentrations for the years 1999-2008. The estimated $\mathrm{PM}_{2.5}$ concentrations for the first two years (1999 and 2000) are the highest values over the state of Georgia. For almost areas, $\mathrm{PM}_{2.5}$ concentrations tend to decrease from 1999 to 2008 (on average, the $\mathrm{PM}_{2.5}$ concentration was $18.33\mu g/m^3$ for 1999 and $13.08\mu g/m^3$ for 2008). Also, the estimated $\mathrm{PM}_{2.5}$ concentrations in the Atlanta areas were higher than the other areas for the years 2001-2006.

To evaluate the prediction performance of the proposed $\mathrm{PM}_{2.5}$ model, we compare the observed $\mathrm{PM}_{2.5}$ values with the estimated $\mathrm{PM}_{2.5}$ at the monitoring locations. The percentage of the observations that are outside the 95% prediction intervals is 1.09% This suggests that the $\mathrm{PM}_{2.5}$ model considered here performs well in terms of the prediction.

To examine the performance of the 2-stage mixture model as the health model, we fit four different models:

1) Model 1: simple linear Poisson model in Equation (2)

2) Model 2: space-time random effect model proposed by Knorr-Held (2000)

$$\log \theta_{ij} = \alpha_0 + Z^*_{ij}\gamma_{ij} + \mathbf{X}'_{ij}\boldsymbol{\beta}_{ij} + u_i + v_i + \xi_j + \delta_j + \eta_{ij},$$

where u_i has an IAR distribution with the variance σ^2_u, ξ_j has an AR(1) with the temporal paramter ρ_ξ ∼Beta(1,1), $v_i \sim \mathrm{N}(0, \sigma^2_v)$, $\delta_j \sim \mathrm{N}(0, \sigma^2_\delta)$, and $\eta_{ij} \sim \mathrm{N}(0, \sigma^2_\eta)$. All the standard deviances have uniform prior distributions.

3) Model 3: full space-time mixture model

$$\log \theta_{ij} = \alpha_0 + Z^*_{ij}\gamma_{ij} + \mathbf{X}'_{ij}\boldsymbol{\beta}_{ij} + \sum_{l=1}^{L} w_{il}\chi_{lj} + \eta_{ij},$$

where w_{il}, χ_{lj},and η_{ij} have the same structures as in Section 3.2.

4) Model 4: 2-stage space-time mixture model proposed in Section 3.2.

For Models 3 and 4, we use $L = 10$ entry parameters because it seems to be large enough to find the true number of latent components. For all the models, we also consider three different structures for the coefficients (γ_{ij} and $\boldsymbol{\beta}_{ij}$) in Equation (3):

(i) Constant: The coefficients are constant over space and time ($\gamma_{ij} = \gamma_0$ and $\beta_{ijp} = \beta_{0p}$).

(ii) Space-varying: The coefficients are constant over time but vary over space ($\gamma_{ij} = \gamma_0 + \gamma^1_i$ and $\beta_{ijp} = \beta_{0p} + \beta^1_{ip}$).

(iii) Space-time varying: The coefficients vary over space and time, presented in Equation (3).

To assess how well the models considered fit the data and predicts, we use the DIC$_3$ measure proposed by Celeux *et al.* (2006) that uses a posterior estimate of likelihood in computing the effective number of parameters, pD. This measure is defined as $\mathrm{DIC}_3 = \overline{D(\Theta)} + \mathrm{pD}_3 = \overline{D(\Theta)} + [\overline{D(\Theta)} + 2\log \hat{p}(\mathbf{y}|\Theta)]$, where $\overline{D(\Theta)}$ is the posterior mean of the deviance. We use this DIC_3 measure instead of the standard DIC measure (Spiegelhalter *et al.*, 2002) because DIC_3 is easily calculated by MCMC and it performs well in mixture models. It also provides stable and reliable evaluations. For the prediction performance, we consider the Marginal Predictive-likelihood (MPL) and the mean square prediction error (MSPE). The MPL computed using the Conditional Predictive Ordinate (CPO) (Dey *et al.*, 1997) is specified as $\text{MPL} = \sum_{i,j} \log(\text{CPO}_{ij})$, where CPO_{ij} is the marginal posterior predictive density of y_{ij} given the data omitting y_{ij}. Thus, the CPO represents a cross-validation measure, and the MPL explains a predictive measure for a future replication of the given data. The model with a larger value of MPL is better (Ibrahim *et al.*, 2001; Congdon, 2005). The MSPE is given by $\text{MSPE} = \frac{1}{IJ}\sum_{i,j}(y_{ij} - \hat{y}_{ij})^2$, where $\hat{y}_{ij}$ is the predicted value of the observed value y_{ij} from the posterior predictive distribution.

Table 3 reports these measures for the models considered and the estimated number of latent temporal components for Models 3 and 4. For each coefficient structure, the simple linear Poisson model (Model 1) has much larger DIC_3 and MSPE values and lower MPL values than the other models. Therefore, the simple linear Poisson model is not appropriate for this data set, and this implies that space, time, or space-time random effect needs to be considered in the model. In terms of DIC_3, MPL and MSPE, the constant coefficient structure over space and time in the space-time random effect model (Model 2) is much better than the other coefficient structures for that Model. Similarly, the 2-stage space-time mixture model (Model 4) with constant coefficients over space and time is better than the model with the other coefficient structures in terms of DIC_3 and MPL. The 2-stage mixture models with different coefficient structures provide similar MSPE values. In contrast, the full space-time mixture model (Model 3) with spatiotemporally varying coefficients has smaller DIC_3 and MPL than those with constant (or space-varying) coefficients. From these results, we can see that the 2-stage mixture model (Model 4) with constant coefficients over space and time has the smallest DIC_3 and MPL overall. Thus, this model is the best fit model among these models. Also, it appears that the 2-stage mixture models estimate 4 components included in the models while the full mixture models estimate the small number of components (1 or 2), which is consistent with the results obtained from the simulation study.

In Table 4 the posterior means and 95% credible intervals for the model parameters in the 2-stage mixture model with constant coefficients over space and time are presented. The proportion of black population and the unemployment rate are significant positive risk factors of the asthma while the $\mathrm{PM}_{2.5}$ and the household median income are significant negative risk factors. For example, a higher proportion of black people or the unemployment rate is associated with increased risk of the asthma. The lower income is associated with increased risk of the asthma. For $\mathrm{PM}_{2.5}$ a slightly surprising result was found. The $\mathrm{PM}_{2.5}$ parameter posterior mean is negative (-0.028) with a small 95% credible interval (-0.034,-0.022). Our results for $\mathrm{PM}_{2.5}$ are inconsistent with some air pollution-related time series studies (Tolbert *et al.*, 2000; Sheppard *et al.*, 2003). However, all the other models (Models 1-3) also provide negative estimates for the $\mathrm{PM}_{2.5}$ coefficient, adjusting for the socioeconomic covariates. This seems to imply that the estimates of $\mathrm{PM}_{2.5}$ are smoother, since $\mathrm{PM}_{2.5}$ data in some areas are sparsely sampled. This may lead to less spatial variation in areas with high disease risk and may tend to produce the negative effects of $\mathrm{PM}_{2.5}$ while controlling for the non-PM covariates and space-time mixture structures.

Figure 6 shows the plots for the temporal components included in the 2-stage mixture model after adjusting for the covariates. Component 1 has a stable increasing pattern and component 4 increases dramatically over year. On the other way, component 3 has a decreasing pattern. Component 2 tends to increase until 2002 and then decrease. In addition, component 2 has the largest relative risks over time while component 3 has the smallest relative risks. The maps of the estimated weights corresponding to the temporal components are displayed in Figure 7 (a). Based on the allocation approach presented in Equation (6), the map of the spatial cluster indicator (C_i) is also displayed in Figure 7 (b). Overall, the Atlanta areas and some of south areas are assigned to component 2 (increasing and then decreasing from the year 2002) and some of center areas and south-east areas are assigned to component 1 or 4. North areas and a few east or south areas are assigned to component 3.

As mentioned previously, it is possible to consider $\mathrm{PM}_{2.5}$ estimated in counties with added measurement error. This could hope to partially address the issue of bias induced by using plug-in estimates. To explore the impact of adding $\mathrm{PM}_{2.5}$ measurement error in the health model, we re-fit the full mixture model and the 2-stage mixture model with constant coefficients over space and time, with measure error added to the $\mathrm{PM}_{2.5}$. We assume Berkson measurement error (Berkson, 1950) in the $\mathrm{PM}_{2.5}$ with $(Z^*_{ij} + \epsilon_{ij})$ and $(Z^*_{ij} + \mathrm{se}(Z^*_{ij}) * \epsilon_{ij})$ replacing Z^*_{ij}, where $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2_\epsilon)$. The value of $\mathrm{se}(Z^*_{ij})$ was computed from the set of county prediction values used to estimate Z^*_{ij}. Table 5 displays the comparison of results from the 2-stage model with measurement error. As compared to the results from the 2-stage model without measurement error in Table 3, including the measurement error has an effect of reducing DIC_3 values and increasing MPL values. The models both including measurement error or not, have similar MSPE values. The DIC_3 and MPL measures favor the 2-stage mixture model with $(Z^*_{ij} + \mathrm{se}(Z^*_{ij}) * \epsilon_{ij})$ among these measurement error models. This model has DIC_3 =10352 and MPL = -5432 although the MSPE is similar to that for the non measurement error version of this model. These results suggest that the measurement error models do provide a better fit overall to these data. However, the posterior mean estimate of $\mathrm{PM}_{2.5}$ is -0.031 with 95% interval (-0.038, -0.023) and the estimated number of temporal components is 4, which is close to that for the 2-stage mixture model without measurement error. Thus, the models with measurement errors have little effect on the estimate of PM and the estimated number of components.

6 Discussion

We have presented a novel approach to the incorporation of covariates within a space-time modeling framework. In particular, we have examined the use of space-time mixture models where covariates are to be introduced. A 2-stage procedure was proposed and applied both in simulations and real data. In simulated comparisons, the 2-stage model yielded lower error in the estimation of predictor parameters and also yielded much greater accuracy in the estimation of latent component numbers than other models. There was little difference in the ability to detect spatial grouping or clustering of risk. In application to the ambulatory asthma data for 1999-2008 we found that we could model the data well with the 2 stage model approach and found 4 components to be optimal. In this case, we also found a small but negative posterior mean for the $\mathrm{PM}_{2.5}$ parameter which is different from previously reported results based on time series studies. The result holds across different space-time modeling scenarios and so we have concluded that it is substantive, but that the negative association could be partly contributed to by the smoothness of the interpolation in sparse areas.

In future analysis we would aim to consider the development of models that could combine predictor information with temporal components so that we could directly relate temporal effects to predictor temporal variation. We would want to consider directly modeling areas with higher densities of monitoring stations so that a more direct link with disease outcome could be examined

References

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004), *Hierarchical Modeling and Analysis for Spatial Data,* Boca Raton, FL: Chapman & Hall/CRC.

Berkson, J. (1950), "Are there two regressions?," *Journal of the American Statistical Association*, 45, 164–180.

Bernardinelli, L., Clayton, D. G., Pascutto, C., Montomoli, C., Ghislandi, M., and Songini, M. (1995), "Bayesian analysis of space-time variation in disease risk," *Statistics in Medicine*, 14, 2433–2443.

Besag, J., York, J., and Mollie, A. (1991), "Bayesian image restoration, with two applications in spatial statistics (with discussion),"*Annals of the Institute of Statistical Mathematics*, 43, 1–59.

Castro, M., Schechtman, B. K., Halstead J., and Bloomberg, G. (2001), "Risk Factors for Asthma Morbidity and Mortality in a Large Metropolitan City,"*Journal of Asthma*, 38, 625–635.

Celeux, G., Forbes, F., Robert, C., and Titterington, M. (2006), "Deviance information criteria for missing data models,"*Bayesian Analysis*, 1, 651–674.

Centers for Disease Control and Prevention. (2008), National Health Interview Survey Raw Data, 2008. Atlanta, GA: U.S. Department of Health and Human Services, CDC.

Centers for Disease Control and Prevention. (2011), "Vital Signs,May 2011: Asthma in the US,"Atlanta, GA: CDC. Available at `http://www.cdc.gov/vitalsigns/pdf/2011-05-vitalsigns.pdf`.

Choi, J., Fuentes, M., and Reich, B. J. (2009), "Spatial-temporal association between fine particulate matter and daily mortality,"*Computational Statistics and Data Analysis*, 53, 2989–3000.

Choi, J., Lawson, A. B., Cai, B., and Hossain, M. M. (2011), "Evaluation of Bayesian spatial-temporal latent models in small area health data,"*Environmetrics*, *DOI: 10.1002/env.1127*, In press.

Clayton, D., Bernardinelli, L., and Montomoli, C. (1993), "Spatial correlation in ecological analysis,"*International Journal of Epidemiology*, 22, 1193–1202.

Congdon, P. (2005), *Bayesian Models for Categorical Data*, New York, NY: John Wiley and Sons.

Dey, D., Chen, M. H., and Chang, H. (1997), “Bayesian approach for nonlinear random effects models,” *Biometrics*, 53, 1239–1252.

Dockery, D. W., and Pope, C. A. III. (1994), “Acute respiratory effects of particulate air pollution,” *Annual Review of Public Health*, 15, 107–132.

Eisner, D. M., Katz, P. P., Yelin, H. E., Shiboski, C. S., and Blanc, D. P. (2001), “Risk factors for hospitalization among adults with asthma: the influence of sociodemographic factors and asthma severity,” *Respiratory Research*, 2, 53–60.

Ellison-Loschmann, L., Sunyer, J., Plana, E., Pearce, N., Zock, J. P., Jarvis, D., Janson, C., Anto, J. M., and Kogevinas, M. (2007), “Socioeconomic status, asthma and chronic bronchitis in a large community-based study,” *European Respiratory Journal*, 29, 897–905.

Friedman, M. S., Powell, K. E., Hutwagner, L., Graham, L., M., and Teague, W. G. (2001), “Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma,” *The journal of the American Medical Association*, 285, 897–905.

Fuentes, M., Song, H., Ghosh, S. K., Holland, D. M., and Davis, J. M. (2006), “Spatial association between speciated fine particles and mortality,” *Biometrics*, 62, 855–863.

Gelfand, A. E., and Vounatsou, P. (2002), “Proper multivariate conditional autoregressive models for spatial data analysis,” *Biostatistics*, 4, 11–25.

Gelman, A. (2004), “Parameterization and Bayesian modelling,” *Journal of the American Statistical Association*, 99, 537–545.

Gelman, A. (2006), “Prior distributions for variance parameters in hierarchical models,” *Bayesian Analysis*, 1, 515–533.

Geweke, J. (1992), *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments,* In Bayesian Statistics 4 (ed JM Bernado, JO Berger, AP Dawid and AFM Smith), Oxford, UK: Oxford University Press.

Gotway, C. A., and Young, L. J. (2002), "Combining incompatible spatial data," *Journal of the American Statistical Association*, 97, 632–648.

Hodges, J., and Reich, B. (2010), "Adding spatially-correlated errors can mess up the fixed effect you love," *The American Statistician*, 64, 325–334.

Ibrahim, J., Chen, M. H., and Sinha, D. (2001), *Bayesian Survival Analysis*, New York, NY: Springer.

Jasra, A., Holmes, C. C., and Stephens, D. A. (2005), "Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling," *Statistical Science*, 20, 50–67.

Knorr-Held, L. (2000), "Bayesian modelling of inseparable space-time variation in disease risk," *Statistics in Medicine*, 19, 2555–2567.

Knorr-Held, L., and Besag, J. (1998), "Modelling risk from a disease in time and space," *Statistics in Medicine*, 17, 2045–2060.

Lawson, A. B., Song, H. R., Cai, B., Hossain, M. M., and Huang, K. (2010), "Space-time latent component modeling of geo-referenced health data," *Statistics in Medicine*, 29, 2012–2027.

Lin, M., Chen, Y., Burnett, R. T., Villeneuve, P. J., and Krewski, D. (2002), "The influence of ambient coarse particulate matter on asthma hospitalization in children: Case-crossover and time-series analyses," *Environmental Health Perspectives*, 110, 575–581.

Lin, S., Liu, X., Le, L. H., and Hwang, S. A. (2008), "Chronic exposure to ambient ozone and asthma hospital admissions among children," *Environmental Health Perspectives*, 116, 1725–1730.

Ma, B., Lawson, A. B., and Liu, Y. (2007), "Evaluation of Bayesian models for focused clustering in health data," *Environmetrics*, 18, 871–887.

Mugglin, A. S., Cressie, N., and Gemmell, I. (2002), "Hierarchical statistical modelling of influenza epidemic dynamics in space and time," *Statistics in Medicine*, 21, 2703–2721.

Paciorek, C. (2010), "The importance of scale for spatial-confounding bias and precision of spatial regression estimators," *Statistical Science*, 25, 107–125.

Pleis, J. R., Lucas, J. W., and Ward, B. W. (2009), Summary health statistics for US adults: National Health Interview Survey, 2008, *Vital Health statistics*, 10(242), National Center for Health Statistics.

Ponka, A., and Virtanen, M. (1996), "Asthma and ambient air pollution in Helsinki," *Journal of Epidemiology and Community Health*, 50 Suppl 1: p. s59–62.

Reich, B., Hodges, J., and Zadnik, V. (2006), "Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models," *Biometrics*, 62, 1197–1206.

Richardson, S., Abellan, J., and Best, N. (2006), "Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (U.K.)," *Statistical Methods in Medical Research*, 15, 97–118.

Sheppard, L. (2003), "Ambient air pollution and nonelderly asthma hospital admissions in Seattle, Washington, 1987-1994," In: Revised analyses of time-series studies of air pollution and health. Boston, MA: Health Effects Institute, 227–230.

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 B*, 64, 583–639.

Stephens, M. (2000), "Dealing with label switching in mixture models," *Journal of the Royal Statistical Society B*, 62, 795–809.

Stieb, D. M., Burnett, R. T., Beveridge, R. C., and Brook, J. R. (1996), "Association between ozone and asthma emergency department visits in Saint John, New Brunswick, Canada," *Environmental Health Perspectives*, 104, 1354–1360.

Tolbert, P. E., Mulholland, J. A., Maclntosh, D. L., Xu, F., Daniels, D., Devine, O. J., Carlin, B. P., Klein, M., Dorley, J., Butler, A. J., Nordenberg, D. F., Frumkin, H., Ryan, P. B., and White, M. C. (2000), "Air quality and pediatric emergency room visits for asthma in Atlanta, Georgia," *American Journal of Epidemiology*, 151, 798–810.

Tzala, T., and Best, N. (2008), "Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality," *Statistical Methods in Medical Research*, 17, 97–118.

Xia, H., Carlin, B. P., and Waller, L. A. (1997), "Hierarchical models for mapping Ohio lung cancer rates," *Environmetrics*, 8, 107–120.

Table 1: Comparison of the estimation results from the full space-time mixture model (M1) and the 2-stage space-time mixture model (M2). True values: $\beta_0 = 1.00$, $\beta_1 = 0.05$, and $\beta_2 = 0.10$.

Design	Model	Parameter	mean	2.5%	97.5%	average width 95% interval	MSE	MAE
1	M1	β_0	0.929	0.600	1.116	0.340	0.01998	0.097
		β_1	0.050	0.042	0.056	0.016	0.00002	0.003
		β_2	0.101	0.091	0.113	0.014	0.00003	0.004
	M2	β_0	0.965	0.948	0.985	0.018	0.00129	0.035
		β_1	0.049	0.041	0.057	0.014	0.00002	0.003
		β_2	0.100	0.096	0.104	0.002	<0.00001	0.002
2	M1	β_0	0.940	0.704	1.126	0.339	0.01594	0.095
		β_1	0.050	0.042	0.056	0.015	0.00001	0.003
		β_2	0.104	0.087	0.120	0.037	0.00009	0.008
	M2	β_0	0.960	0.947	0.975	0.016	0.00167	0.040
		β_1	0.050	0.043	0.056	0.014	0.00001	0.003
		β_2	0.106	0.102	0.109	0.004	0.00003	0.006
3	M1	β_0	0.964	0.744	1.122	0.322	0.01126	0.069
		β_1	0.050	0.043	0.057	0.016	0.00001	0.003
		β_2	0.101	0.084	0.124	0.048	0.00009	0.007
	M2	β_0	0.964	0.951	0.977	0.015	0.00134	0.036
		β_1	0.050	0.043	0.057	0.014	0.00001	0.003
		β_2	0.100	0.095	0.106	0.007	0.00001	0.002

Table 2: Percentage table of the estimation of the number of components included in the model over simulation (%). The true number of components is 3 and the number of simulations is 200. (M1: the full space-time mixture model; M2: the 2-stage space-time mixture model).

		L							
Design	Model	0	1	2	**3**	4	5	6	Total
1	M1	0.5	51.5	36.5	**10.5**	1.0	0	0	100
	M2	0	0	2.0	**93.0**	3.5	1.5	0	100
2	M1	0.5	19.0	55.5	**23.0**	2.0	0	0	100
	M2	1.0	0.5	4.5	**87.0**	5.5	1.5	0	100
3	M1	0	33.5	56.0	**10.5**	0	0	0	100
	M2	0	0	0.5	**94.5**	3.0	2.0	0	100

Table 3: Comparison results from four models and three different coefficient structures for Asthma data in Georgia.

Coefficient structure	Model	DIC_3	pD_3	MPL	MSPE	$\hat{L}$
Constant	Model 1	18977	45	-9490	708.5	
	Model 2	10469	427	-5624	127.9	
	Model 3	10551	487	-5846	127.9	1
	Model 4	10451	365	-5477	128.5	4
Space-varying	Model 1	11765	489	-6073	218.2	
	Model 2	11221	521	-5788	171.4	
	Model 3	10516	436	-5655	128.6	1
	Model 4	10500	432	-5650	128.2	4
Space-time varying	Model 1	11583	499	-6022	209.7	
	Model 2	11150	423	-5754	169.7	
	Model 3	10490	410	-5576	128.1	2
	Model 4	10453	422	-5621	127.6	4

Table 4: Parameter estimation in the best-fitted model (the 2-stage mixture model with constant coefficients over space and time).

covariates	mean	sd	2.50%	97.5%
intercept	0.197	0.0058	0.185	0.208
$PM_{2.5}$	-0.028	0.0031	-0.034	-0.022
black proportion	0.004	0.0004	0.003	0.005
income	-0.019	0.0005	-0.020	-0.018
unemployment rate	0.024	0.0039	0.017	0.032

Table 5: Comparison results from the 2-stage mixture model with constant coefficients over space and time and two different measurement error structures in the $PM_{2.5}$ for Asthma data in Georgia.

Measurement error structure	DIC_3	pD_3	MPL	MSPE	$\hat{L}$
$Z^*_{ij} + \epsilon_{ij}$	10366	399	-5463	128.4	4
$Z^*_{ij} + \text{se}(Z^*_{ij}) * \epsilon_{ij}$	10352	378	-5432	128.7	4

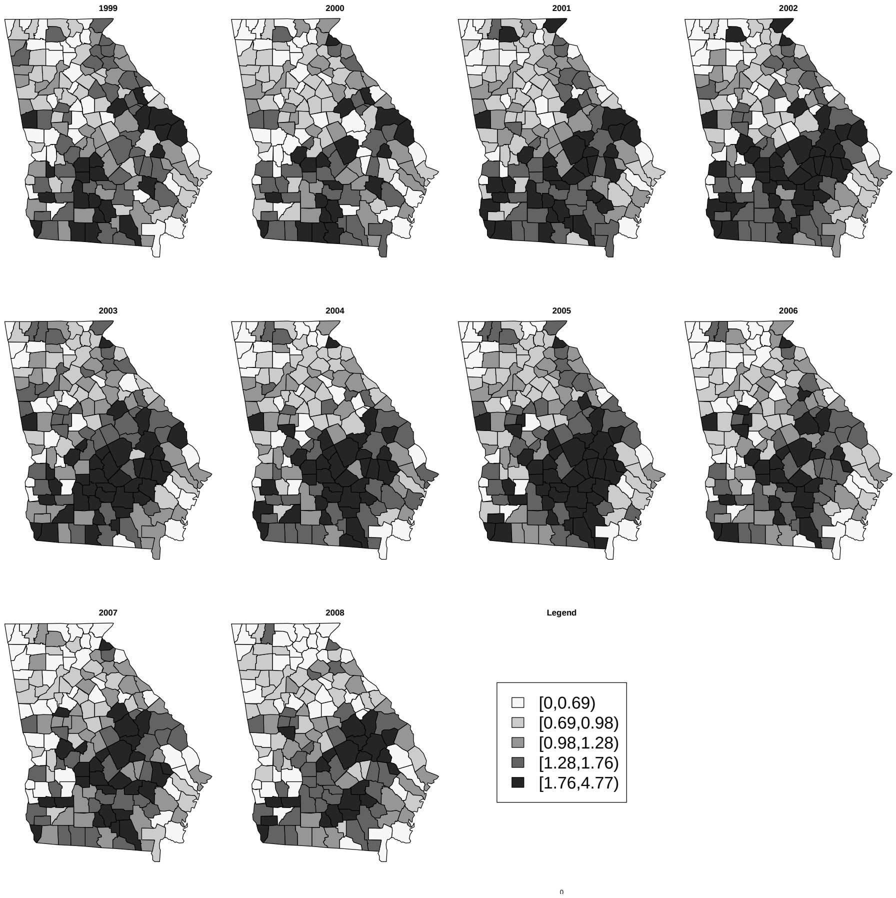
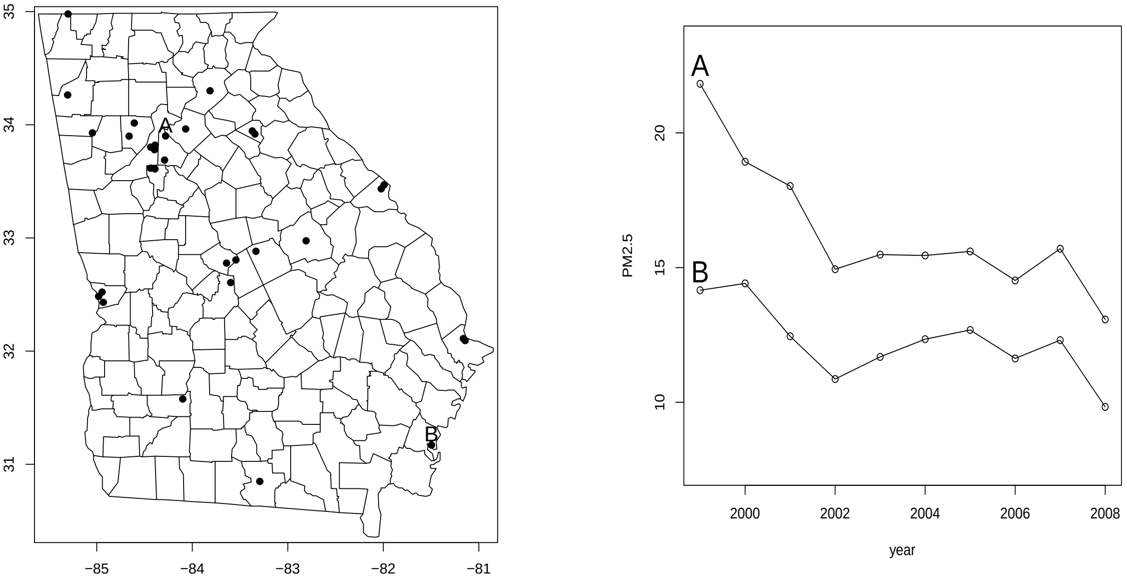


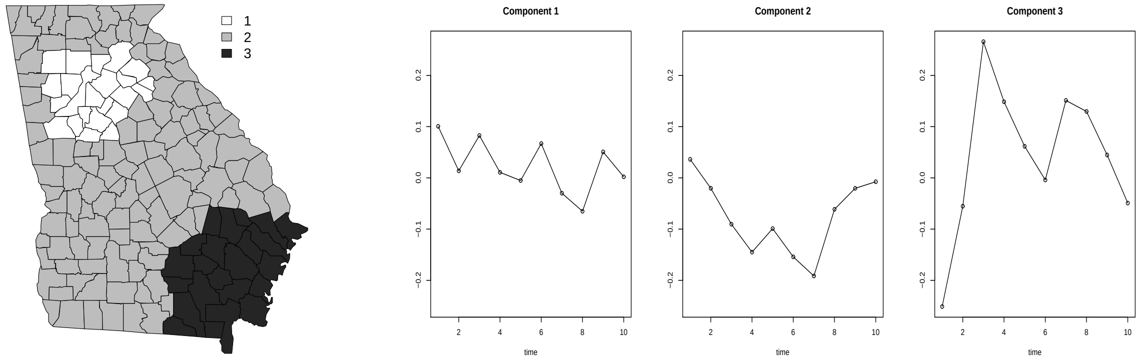
Figure 1: Standardized incidence maps for county-level ambulatory sensitive asthma in Georgia for individual year.



(a)

(b)

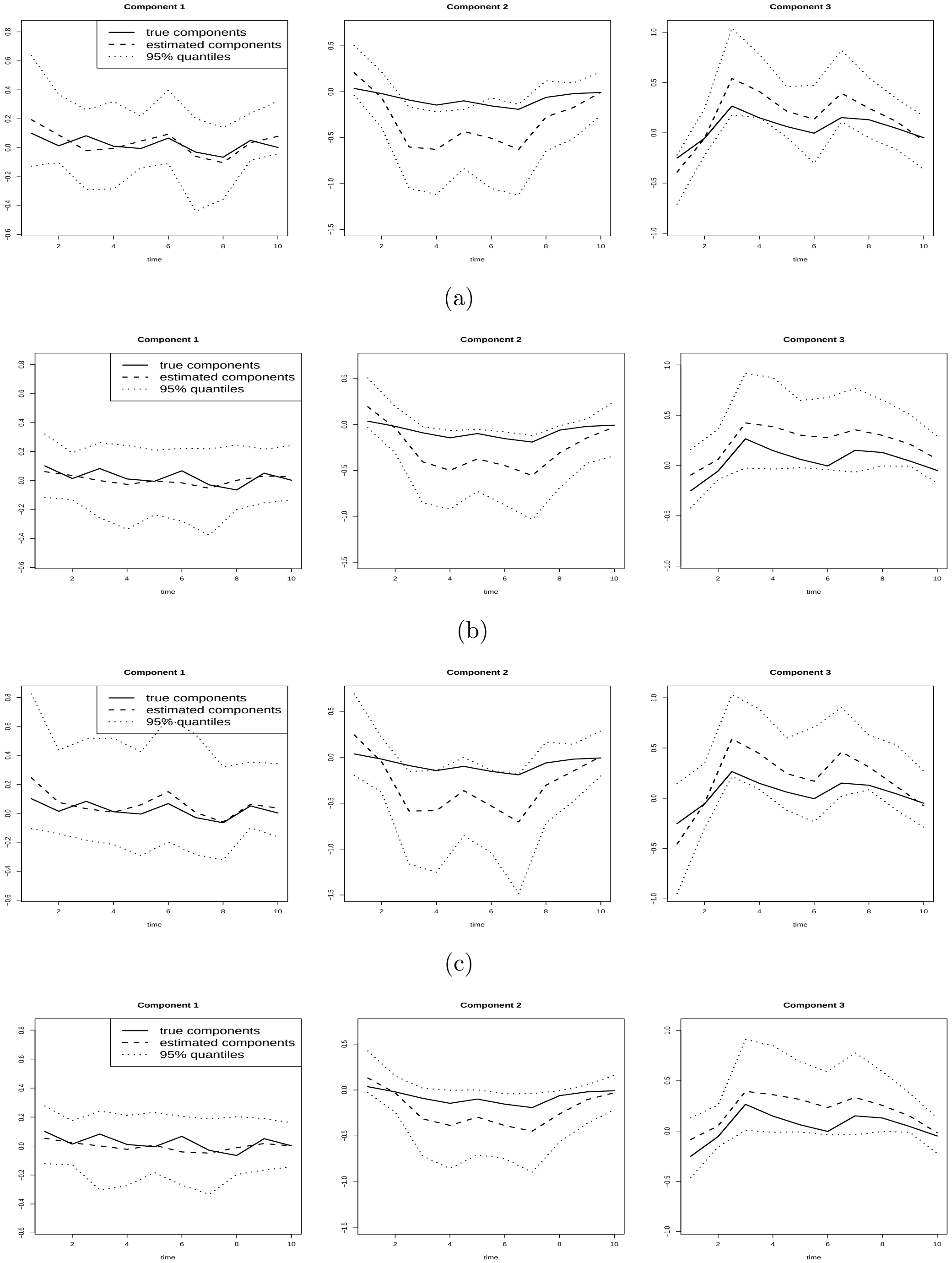
Figure 2: (a) Map of $\mathrm{PM}_{2.5}$ monitoring stations. (b) Temporal trends of $\mathrm{PM}_{2.5}$ for the selected locations (A and B).



(a)

(b)

Figure 3: (a) Map of the spatial cluster indicator for simulation study. (b) Temporal plots of the true components for simulation study.



(a)

(b)

(c)

(d)

Figure 4: (a) Temporal plots from the full mixture model in Design 1. (b) Temporal plots from the 2-stage mixture model in Design 1. (c) Temporal plots from the full mixture model in Design 2. (d) Temporal plots from the 2-stage mixture model in Design 2.

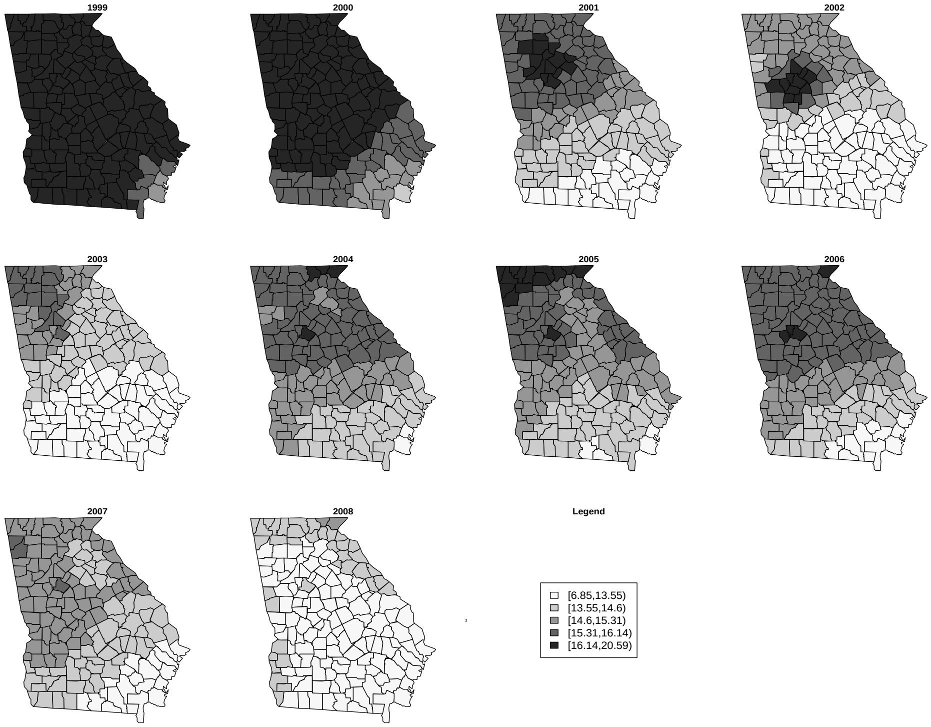


Figure 5: Maps of the estimated $PM_{2.5}$ concentrations for the years 1999-2008 in Georgia.

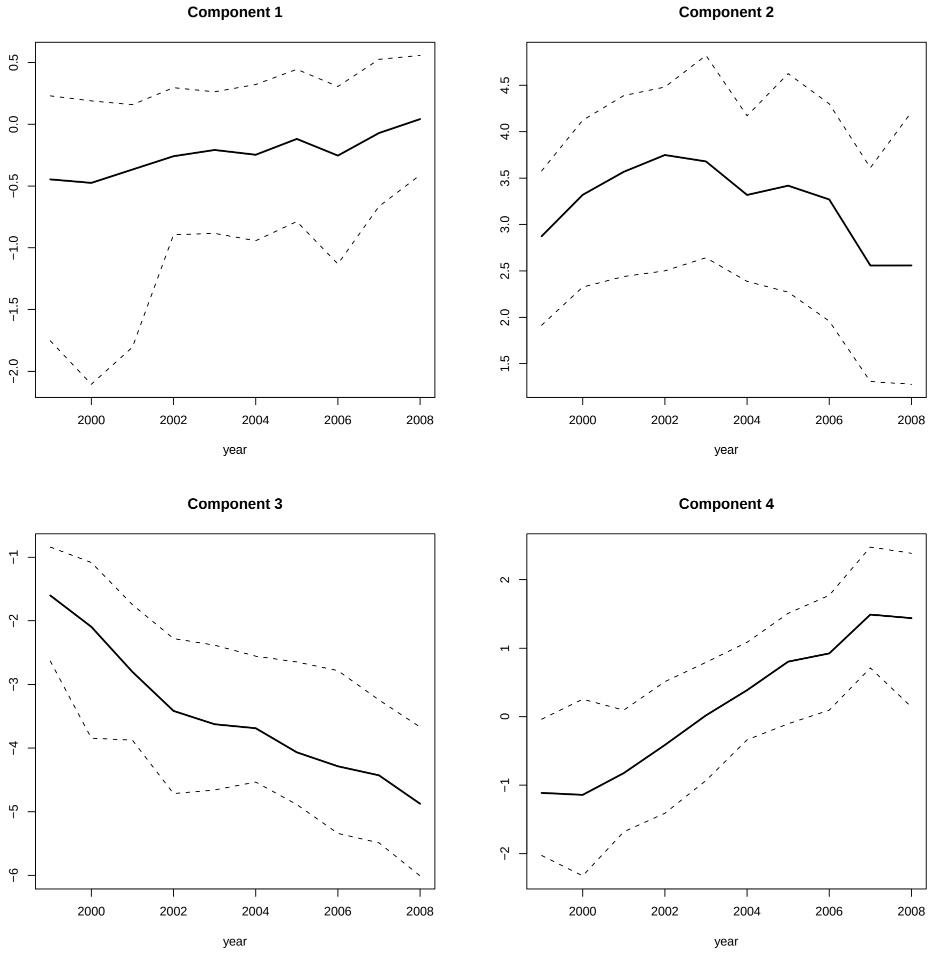
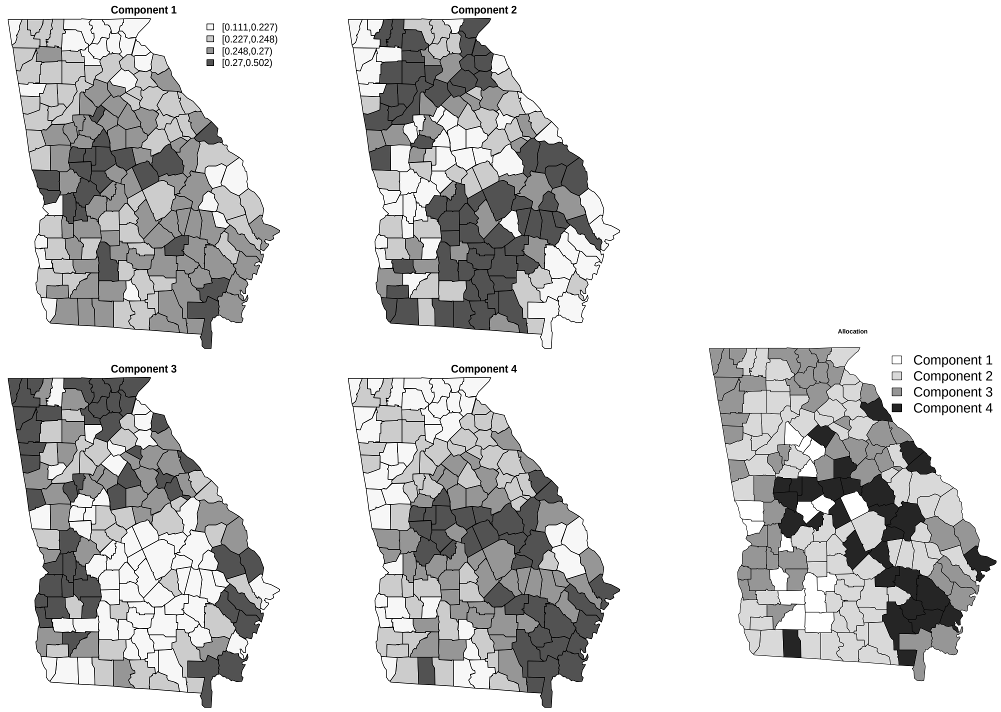


Figure 6: Temporal plots for four estimated components from the 2-stage mixture model.



(a) (b)

Figure 7: (a) Maps of the estimated weights corresponding with the temporal components. (b) Map of the allocation using the weights.