Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
21(5) 445–456

DOI: 10.1177/0962280212446318
smm.sagepub.com
SAGE

A Bayesian latent model with spatio-temporally varying coefficients in low birth weight incidence data

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

Abstract

In spatial epidemiology studies, the effects of covariates on adverse health outcomes could vary over space and time so examining the spatio-temporally varying effects is useful. In particular, the association between covariates and health outcomes could have locally different temporal patterns. In this article, we develop a Bayesian spatio-temporal latent model to identify spatial clusters in each of which covariate effects have homogeneous temporal patterns as well as estimate heterogeneous temporal effects of covariates depending on spatial groups. We compare the proposed model to several alternative models to assess the performance of the proposed model in terms of a range of model assessment measures. Low birth weight incidence data in Georgia for the years 1997–2006 are used.

Keywords

Low birth weight, spatial cluster, spatio-temporal mixture model, latent model

1 Introduction

Numerous epidemiology studies have mainly focused on the association between risk factors and health outcomes such as child-related diseases. Health data and covariate data are often collected over space and time and they have space-time variation. So, the association between covariates and health outcomes may vary spatio-temporally. In particular, the relations in spatial health data could be different depending on areas within the spatial domain. They could also have locally different temporal patterns in the spatio-temporal domain. However, the commonly used spatial model assumes that covariate effects on health outcomes are constant within their entire space and time

[1]Division of Biostatistics and Epidemiology, College of Medicine, Medical University of South Carolina, Charleston, SC, USA
[2]Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC, USA
[3]Center for Clinical and Translational Sciences, The University of Texas, Houston, TX, USA
[4]Department of Community and Family Health, College of Public Health, University of South Florida, Tampa, FL, USA

Corresponding author:
Jungsoon Choi, Division of Biostatistics and Epidemiology, College of Medicine, Medical University of South Carolina, Charleston, SC, USA.
Email: choju@musc.edu

domains, although relative risk is constructed with a function of space-time random effects.[1–4] Thus, investigating the spatio-temporal association between covariates and health outcomes is useful and important.

In this article, the focus is on identifying the spatial groups where covariate effects have different temporal patterns depending on groups as well as estimating their temporal profiles varying locally. In analyzing small area health data, covariates may have different temporal effects on disease outcomes, depending on geographical areas (e.g. urban/rural areas). In some cases, researchers often consider dynamic models to capture spatio-temporal variation on the coefficients in linear models[5]. But, these approaches have difficulty to estimate the locally heterogeneous temporal profiles of the covariate effects. Recently, Lawson et al.[6] proposed a Bayesian spatio-temporal mixture model for analyzing the spatio-temporal health data to estimate temporal patterns varying across space in relative risks. They only focused on understanding the heterogeneous temporal patterns of relative risks not covariate effects. So, the development of statistical modeling is needed to find the spatial groups that are determined by distinct temporal patterns of the covariate effects and then estimate the locally varying effects.

To investigate the locally varying temporal behavior of covariate effects, we develop a new Bayesian latent model with space-time varying coefficients. The proposed model allows for the heterogeneous temporal patterns of coefficients depending on areas within the study domain. In each spatial group, a coefficient has a temporal profile which is different from the profiles of the coefficients in other groups. Since the temporal patterns of the covariate effects within neighboring areas may be different and the patterns in disconnected areas may be same, the proposed model does not require the groups to be spatially adjacent, which is a general assumption. Instead, we model the group memberships of coefficients using spatially dependent weight structures to consider spatial variation in grouping patterns. A Bayesian hierarchical approach is also employed to model locally temporal profiles of the coefficients as well as space clustering in the coefficients simultaneously. Thus, the novelty of the proposed model is flexibility and practicality in analyzing space-time health data. To our best knowledge, this work is the first attempt to consider the spatial groups where the covariate effects have distinct temporal profiles within a statistical model and then better understand the locally varying temporal effects of covariates.

In this article, we assume that the number of spatial groups is finite and can be estimated. However, the determination of the number of spatial groups is one difficulty in clustering approaches. To avoid the specification of the number of groups, Dirichlet process mixture modeling or the use of entry parameters can be considered.[7–9] However, these approaches have a computational burden because the maximum number of groups considered in the model should be assumed to be very large and dramatically increased with the size of the spatial domain. In our proposed modeling, we overcome this problem by using several model assessment criteria. We initially consider the range of the number of spatial groups based on the spatio-temporal variation of the data of interest and the size of the spatial domain. The proposed model with different number of groups is fitted and the best number of spatial groups in the proposed model is determined by using various model diagnostics. This approach is especially appropriate for small area health data because the number of spatial groups in small area data would be finite in general.

To assess the performance of the proposed model, we conduct the comparison of our proposed model to several alternative models that do not identify the heterogeneous behavior of coefficients depending on areas. The competing models include constant coefficients over space and time, temporally varying coefficients over space, spatially varying coefficients over time, and globally space-time varying coefficients. A Bayesian model choice criterion[10] and prediction measures[11] are considered to examine the performance of the models considered. For the comparison, a county-level

low birth weight (LBW) incidence data set from the state of Georgia is used. We also explore the results of the covariate effects on LBW from the proposed model.

The outline of this article is as follows. In Section 2, we describe the LBW incidence data used in this research. Section 3 proposes our spatio-temporal latent model, and Section 4 explores various model comparison methods. In Section 5, we apply our spatio-temporal model and alternative models to the LBW data, find the best model using model assessment tools, and provide the findings. Conclusions and potential future work are provided in the last Section 6.

2 The data

We used county-level LBW (infant birth weight less than 2500 g) incidence data in Georgia for the year 1997 to 2006. The counts of LBW were collected from the state health information system OASIS (Georgia Division of Public Health, http://oasis.state.ga.us/). The number of counties in Georgia is 159 and the number of years is 10. As socioeconomic covariates of LBW,[12] we consider the county-level population density (defined as population divided by total land area in square miles), the proportion of black people, median household income, and unemployment rate, which were obtained from the US census and the US Bureau of Labor Statistics. In addition, we use aggregate data based on birth certificates for the other socio-demographic and behavioral risk factors during pregnancy related with mothers. The proportion of mothers with less than 12th grade education, the proportion of mothers smoking during pregnancy, and the proportion of mothers with 'Inadequate' value from the Kotelchuck Index[13] (IKI value) are considered.

Figure 1 displays the maps of standardized incidence ratios for LBW births where the standardized incidence ratio is defined as the number of LBW births divided by the expected counts computed using the internal standardization method.[14] Here, the spatio-temporal variation of standardized incidence ratios can be seen. For example, the LBW standardized incidence ratios in some central areas of Georgia are higher than other areas over years. In south-east areas, the standardized incidence ratios have temporal variation, showing that the ratios from 2001 to 2003 are overall lower than those for the other years. We also found the spatio-temporal variation of the covariates considered, so the effects of covariates on LBW in Georgia may vary across space and time. In addition, the spatial domain (Georgia state) is a small area so it is reasonable to assume that the number of spatial groups is finite and small. Thus, we apply the proposed model and the competing models to the LBW data and assess the performance of the models.

3 Statistical model

Let $\{Y_{it} : i = 1, \ldots, I, \text{ and } t = 1, \ldots, T\}$ be the LBW count data collected for county i at time point t, where $I = 159$ and $T = 10$ and let E_{it} be the expected count. Conditional on E_{it}, we model Y_{it} as a standard Poisson distribution

$$Y_{it} \sim \text{Pois}(E_{it} \exp(\theta_{it})) \tag{1}$$

where θ_{it} is the logarithm of the relative risk. To account for spatio-temporal dependence in coefficients, θ_{it} is specified as

$$\theta_{it} = \mathbf{X}'_{it}\beta_{it}$$

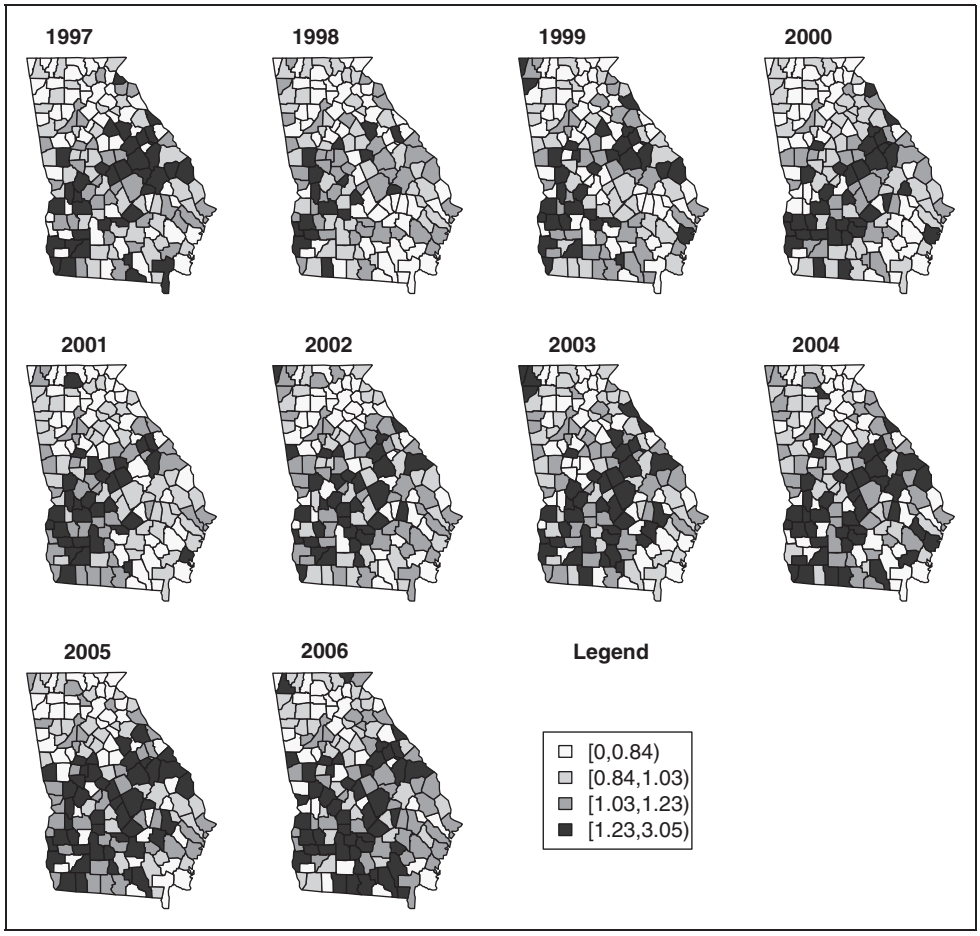

Figure 1. The standardized incidence maps of county-level LBW data in Georgia for each year.
LBW: low birth weight.

where $\mathbf{X}_{it} = (1, X_{it1}, \ldots, X_{itP})'$ denotes the vector of both intercept and P covariates of county i at time t and $\boldsymbol{\beta}_{it} = (\beta_{it0}, \beta_{it1}, \ldots, \beta_{itP})'$ denotes the corresponding coefficient vector.

In this study, we assume that there are spatial clusters in which each has a set of homogeneous time-varying coefficients. Thus, covariate effects on LBW change over spatial clusters and they within a spatial cluster have own temporal pattern. To satisfy this assumption, we consider $\boldsymbol{\beta}_{it}$ as

$$\boldsymbol{\beta}_{it} = \beta_{Z(i),t} \tag{2}$$

where $Z(i)(=1, \ldots, G)$ is the spatial cluster indicator and G is the number of spatial clusters.

To construct spatial clusters, we model $Z(i)$ as

$$Z(i) \sim \text{Categorical}(q_{i1}, \ldots, q_{iG}) \tag{3}$$

where $\sum_{g=1}^{G} q_{ig} = 1$ and $q_{ig} \geq 0$. We again model the probabilities using the unnormalized structure

$$q_{ig} = \frac{\delta_{ig}}{\sum_{g=1}^{G} \delta_{ig}}$$

where $\delta_{ig} \geq 0$. To consider spatial-dependent structures in the unnormalized weights δ_{ig}, we introduce a spatial dependence structure to construct the parameter δ_{ig}

$$\log(\delta_{ig}) \sim N(\eta_{ig}, \sigma_g^2)$$

where η_{ig} is the spatially correlated mean and σ_g^2 is the variance. To account for spatial dependence in the mean, the prior distribution of η_{ig} is specified as

$$\eta_{ig} | \eta_{i'g}, i' \neq i \sim N\left(\frac{1}{n_i}\sum_{i' \sim i} \eta_{i'g}, \frac{\sigma_{\eta_g}^2}{n_i}\right)$$

independently for each g, where $i' \sim i$ denotes that county i' is a neighbor of county i, n_i is the total number of neighbors of county i, and $\sigma_{\eta_g}^2$ controls the magnitude of spatial variation. This is denoted as a conditional autoregressive (CAR) distribution proposed in Besag et al.[15]

The coefficient vector $\boldsymbol{\beta}_{g,t} = (\beta_{gt0}, \beta_{gt1}, \ldots, \beta_{gtP})'$ within the spatial group $Z(i) = g$ is modeled as

$$\beta_{gtp} \sim N(\beta_{g,t-1,p}, \sigma_{gp}^2), \quad p = 0, 1, \ldots, P.$$

This suggests that the coefficient within a spatial cluster has a temporal trend. In this paper, the coefficients are assumed to follow an independent random walk process.

From the previous equations, we can easily express the likelihood of the observed counts $\mathbf{Y}$ as

$$f(\mathbf{Y}|\Theta) = \prod_{i=1}^{I}\prod_{t=1}^{T} \text{Pois}(Y_{it}|E_{it}, \mathbf{X}_{it}, \beta_{Z(i),t})$$

where Θ is a set of all the parameters included in the model. As suggested by Gelman,[16] we consider the prior distributions of the variance parameters in the model as

$$\sigma_g, \ \sigma_{n_g}, \ \text{and} \ \sigma_{gp} \sim \text{Uniform}(0, c)$$

where c is a constant. Thus, inference on the all the parameters Θ is based on its posterior distribution as follows

$$f(\Theta|\mathbf{Y}) \propto f(\mathbf{Y}|\Theta) \times p(\Theta)$$

The estimations of the parameters given the data are obtained by Markov Chain Monte Carlo (MCMC) algorithms: Gibbs sampling or Metropolis adaptive rejection sampling algorithm. Since the spatial cluster indicator $Z(i)$ is the nominal value, we use the posterior mode for its estimation. For other parameters, the posterior means are used as their estimations.

In Bayesian mixture modeling, component identifiability problems can emerge since the likelihood is not variant with respect to the permutation of the components labels, which was found in Stephens.[17] Recently, Choi et al.[9] examined when label switching problems appeared in spatio-temporal mixture modeling and found out that components could switch labels during

MCMC simulation if multiple chains were used. We thus use a single chain run in this analysis to avoid this label switching problem.

4 Model comparison

As a Bayesian model selection method, the deviance information criterion (DIC) of Spiegelhalter et al.[18] is widely used. DIC forms the goodness of fit as well as the complexity of the Bayesian model. It is defined as

$$\mathrm{DIC} = \overline{D(\Theta)} + \mathrm{pD}$$

where $\overline{D(\Theta)} = E_{\Theta}[D(\Theta)]$ is the posterior mean of the deviance, $D(\Theta) = -2\log f(\mathbf{Y}|\Theta)$ is the deviance, and pD is the effective number of parameters in the model. Spiegelhalter et al.[18] proposed that pD is approximated as $\overline{D(\Theta)} - D(\hat{\Theta})$, where $D(\hat{\Theta})$ is the deviance of the estimate of the parameter Θ. Although this DIC is commonly used in comparing models, Celeux et al.[10] suggested the DIC_3 as an alternative DIC, which is better than the standard DIC in comparing mixture models. The DIC_3 is specified as $\mathrm{DIC}_3 = \overline{D(\Theta)} + \mathrm{pD}_3$, where

$$\mathrm{pD}_3 = \overline{D(\Theta)} + 2\log\hat{f}(\mathbf{Y}|\Theta)$$

Since the posterior estimate of likelihood is used in computing pD_3, the DIC_3 is easily obtained by MCMC. In Celeux et al.,[10] it is showed that the DIC_3 is a stable model comparison tool in mixture models and it provides reliable evaluations.

In this article, we also consider a comparison method through the conditional predictive ordinate (CPO).[19,20] This CPO is a cross-validation measure and obtained using the marginal posterior predictive density so the CPO for county i at time t is defined as

$$\begin{aligned}\mathrm{CPO}_{it} &= f(Y_{it}|\mathbf{Y}_{(-it)}) \\ &= \int f(Y_{it}|\Theta, \mathbf{X}_{it}) f(\Theta|\mathbf{Y}_{(-it)}, \mathbf{X}_{(-it)}) \mathrm{d}\Theta\end{aligned}$$

where $\mathbf{Y}_{(-it)}$ and $\mathbf{X}_{(-it)}$ denote the vector of the LBW observations excluding Y_{it} and the vector of the covariates excluding $\mathbf{X}_{it}$, respectively. Using MCMC numerical approximation, an estimate of the CPO is obtained by the harmonic mean of the likelihoods for observations[11]

$$\widehat{\mathrm{CPO}}_{it} = \left[\frac{1}{\mathrm{M}}\sum_{s=1}^{\mathrm{M}}\frac{1}{f(Y_{it}|\Theta^{(s)}, \mathbf{X}_{it})}\right]^{-1}$$

where M is the number of samples after a burn-in period and $\{\Theta^{(s)}\}_{s=1}^{\mathrm{M}}$ are MCMC samples from the posterior distributions $f(\Theta|\mathbf{Y})$. The CPO estimate does not require the computation of the full cross-validation based on IT separate estimations, and it is easily computed by one-time estimation run based on the full sample. Thus, this CPO overcomes the computational difficulties, although it is based on cross-validation methods. A larger value of CPO indicates better prediction based on the model. As a summary measure, the marginal predictive-likelihood (MPL) based on the CPO is obtained by

$$\mathrm{MPL} = \sum_{i,t}\log(\widehat{\mathrm{CPO}}_{it})$$

Thus, a model with a larger value of MPL implies better model fit.[11,21]

Along with the DIC_3 and the MPL, the mean square prediction error (MSPE) is finally considered for the comparison of models in terms of prediction performance

$$\text{MSPE} = \frac{1}{IT}\sum_{i,t}(Y_{it} - \hat{Y}_{it})^2$$

where Y_{it} and $\hat{Y}_{it}$ are the observed value and the corresponding predicted value from the posterior predictive distribution, respectively.

5 Data analysis and results

The results reported in this section are based on posterior samples of 70,000 iterations per MCMC, with 20,000 burn-in iterations. We also keep every 10th sample to avoid the high autocorrelations of the some quantities. Thus, 5000 final samples are used for the estimation of the parameters. To ensure MCMC convergence, several diagnostics such as the Geweke convergence diagnostic,[22] autocorrelation functions, and trace plots are used. Both R (http://www.r-project.org) and WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs) are used to implement the models considered in this article.

We now apply our proposed model to the county-level low birth weight incidence data as described in Section 2. In order to select the best number of spatial groups in the model, the model with a range of the number of groups is fitted. The number of spatial clusters considered here is between 2 and 10, where the number of clusters is assumed to be at least two. The proposed model with the best number of groups is determined using several model comparison measures described in the previous section. In Figure 2, the plots of DIC_3, MPL and MSPE values for the proposed model with the different number of spatial groups are displayed. Overall, we find that DIC_3 and MSPE values decrease and MPL value dramatically increases as the number of spatial clusters (G) increases up to 8. DIC_3 and MSPE values for the proposed model with 7–9 spatial groups are quite similar and small even though the pattern of MPL values for the model with 7–9 groups is a little bit different. In this analysis, we can see that the model with 10 groups (the maximum number of groups considered) does not provide the best improvement among the model with the different number of groups considered in terms of these model selection measures. MPL measure indicates that the model with 8 spatial groups has the largest value among the models considered. In addition, the proposed model with 7 or 8 groups has similar small DIC_3 and MSPE values so it is reasonable that the model with 8 spatial groups provides the most adequate fit.

We also compare the proposed model with four competing models in our analysis. The models under consideration are as follows:

(1) **Model 1**: simple log-linear Poisson model with spatio-temporally constant coefficients ($\beta_{itp} = \beta_p \sim \text{N}(0, 10^5)$)
(2) **Model 2**: log-linear Poisson model with temporally varying, spatially constant coefficients ($\beta_{itp} = \beta_{0p} + \beta_{tp}^T$; $\beta_{0p} \sim \text{N}(0, 10^5)$ and $\beta_{tp}^T \sim$ Random Walk process)
(3) **Model 3**: log-linear Poisson model with spatially varying, temporally constant coefficients[2,15] ($\theta_{it} = \mathbf{X}_{it}'\beta_i + u_i + v_i + \zeta_t + \xi_t$ and $\beta_{ip} = \beta_{0p} + \beta_{ip}^{S1} + \beta_{ip}^{S2}$; $u_i \sim \text{N}(0, \sigma_u^2)$, $\zeta_t \sim \text{N}(0, \sigma_\zeta^2)$, $\beta_{0p} \sim \text{N}(0, 10^5)$, $\beta_{ip}^{S1} \sim \text{N}(0, \sigma_{\beta_p^{S1}}^2)$, $\xi_t \sim$ Random Walk process, v_i and $\beta_{ip}^{S2} \sim$ CAR prior)
(4) **Model 4**: log-linear Poisson model with additive spatial and temporal coefficients ($\beta_{itp} = \beta_{0p} + \beta_{tp}^T + \beta_{ip}^S$; $\beta_{0p} \sim \text{N}(0, 10^5)$, $\beta_{tp}^T \sim$ Random Walk process, and $\beta_{ip}^S \sim$ CAR prior)
(5) **Model 5**: the proposed model with 8 spatial groups.

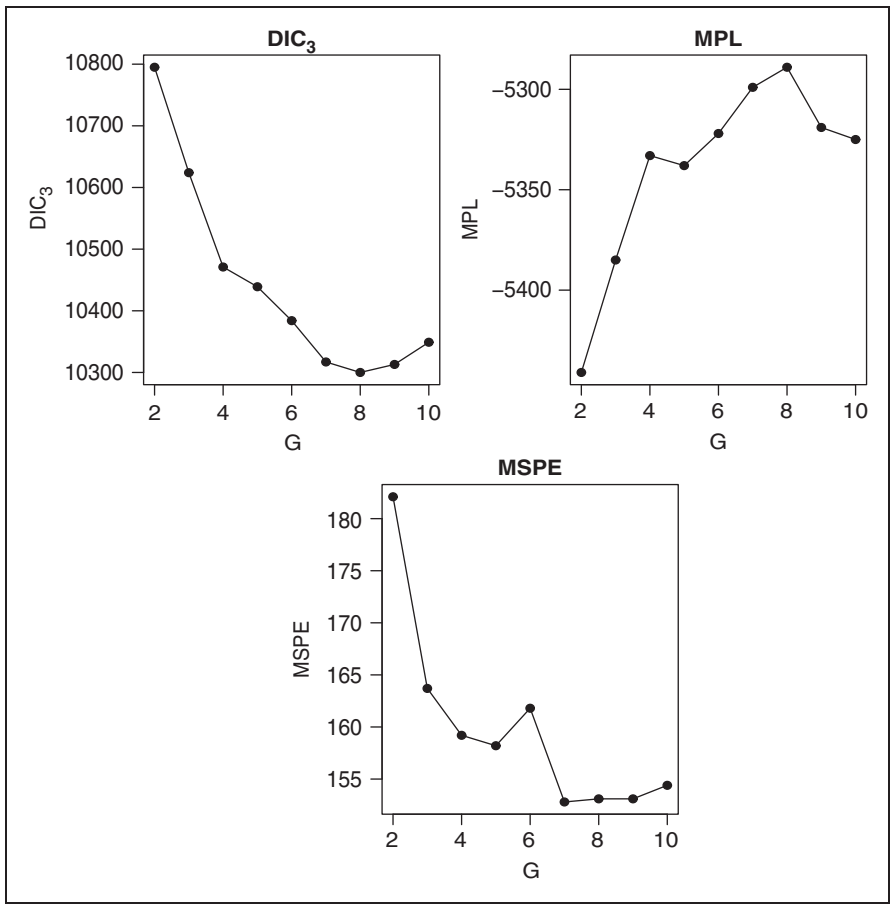

Figure 2. The selection of the number of groups in the proposed model using DIC_3, MPL and MSPE.
DIC: deviance information criterion; MPL: marginal predictive-likelihood; MSPE: mean square prediction error.

Table 1 reports the model assessment measures for the models considered. Model 1 has the largest DIC_3 and MSPE values and the smallest MPL value, which suggests that a model with constant coefficients within their space and time domains is inappropriate for this LBW data set. Adding the temporal-varying coefficients over space to the simple linear Poisson model, the DIC_3, MPL and MSPE measures favor Model 2 over Model 1. Similarly, the Poisson linear model with spatially varying and temporally constant coefficients (Model 3) is better than Model 1 in terms of the comparison measures. In addition, the Poisson linear model including additive space-time varying coefficients (Model 4) has smaller DIC_3 and MSPE values and larger MPL values than Models 1–3. However, the proposed model with 8 spatial groups provides the smallest DIC_3 and MPL values and the largest MSPE value among the models considered. This means that the proposed model with 8 spatial groups is better than the Poisson linear models with various space-time structures of the coefficients in terms of the model assessment criteria. Therefore, the proposed model with 8 spatial groups is the best fit and from now, all the results of estimates are based on the proposed model with 8 spatial clusters.

The map of the spatial cluster indicator including 8 groups and the histogram of the number of counties by group are presented in Figure 3. The fifth and seventh groups have the largest number of counties in Georgia, which is 39 counties (24.5%) for each spatial group. In addition, the downtown of Atlanta is assigned to the seventh group. The eighth spatial group has the third largest number of counties and it includes 27 counties (17.0%). Many counties in the Atlanta suburbs are especially allocated to the eighth group. On the other hand, 1 county (Charlton county) is only assigned to the first group and six counties located in east areas of Georgia are assigned to the sixth group. The remaining groups (groups 2, 3 and 4) include between 12 and 20 counties.

We find that the temporal profiles of the covariate effects on LBW vary across spatial clusters. For example, the coefficient corresponding to the proportion of mothers with low education in the fourth spatial group dramatically increases over time as presented in Figure 4. This coefficient in group 3 is also a little increasing over time. Even though the coefficient in groups 7 and 8 has a stable temporal trend, the estimated coefficient in group 7 is smaller than that in group 8 over time.

Table 1. Model comparison using DIC_3, MPL and MSPE for LBW data

Model	$\overline{D}$	pD_3	DIC_3	MPL	MSPE
Model 1 (simple)	11520	22	11542	−5773	306
Model 2 (temporal)	11300	94	11394	−5714	275
Model 3 (spatial)	10360	230	10590	−5355	180
Model 4 (additive)	10320	238	10558	−5348	176
Model 5 (proposed)	9991	309	10300	−5289	153

DIC: deviance information criterion; MPL: marginal predictive-likelihood; MSPE: mean square prediction error; LBW: low birth weight.

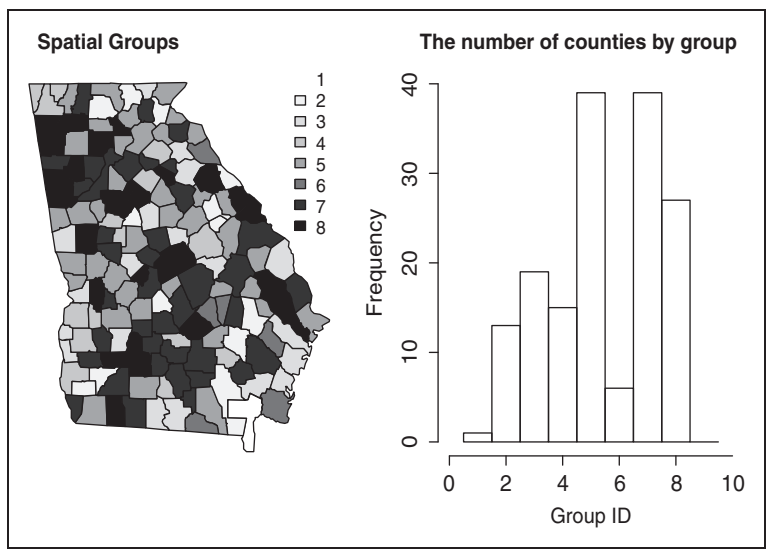

Figure 3. The map of 8 spatial groups (left) and the histogram of the number of counties by group (right).

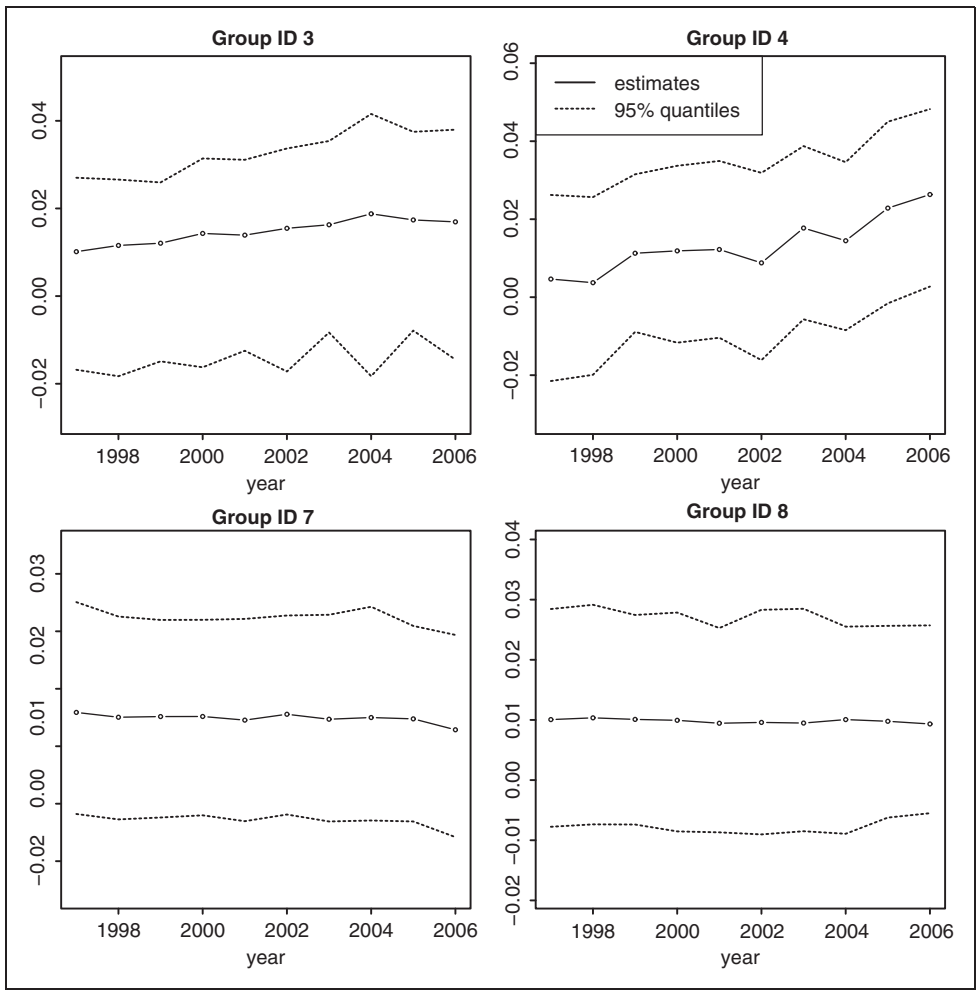

Figure 4. Temporal plots of the coefficient corresponding to the proportion of mothers with low education for the selected spatial groups (Group IDs 3, 4, 7 and 8). The solid lines indicate the posterior mean and the dotted lines indicate the 95% confidence intervals.

This indicates that some of southern-east and southern-west areas (groups 3 and 4) have increasing temporal effects of the proportion of low educated mothers over time while the downtown of Atlanta and some of the Atlanta suburbs have stable temporal effects. It is also found that the estimates of the coefficient are positive even though they are not significant. They thus suggest that a higher proportion of low educated mothers is associated with increased risk of the LBW incidence and the effects of maternal low education level in rural areas are higher than those in urban areas. These findings are consistent with previous analyses.[23,24]

To evaluate the spatial prediction performance of the proposed model, we conduct calibration analysis. First, we randomly selected 16 counties (about 10% of the data) and fitted the proposed model 16 times by removing all observations from the selected county. For the second analysis, we removed the observations from all 16 counties simultaneously at a single time point, repeating this

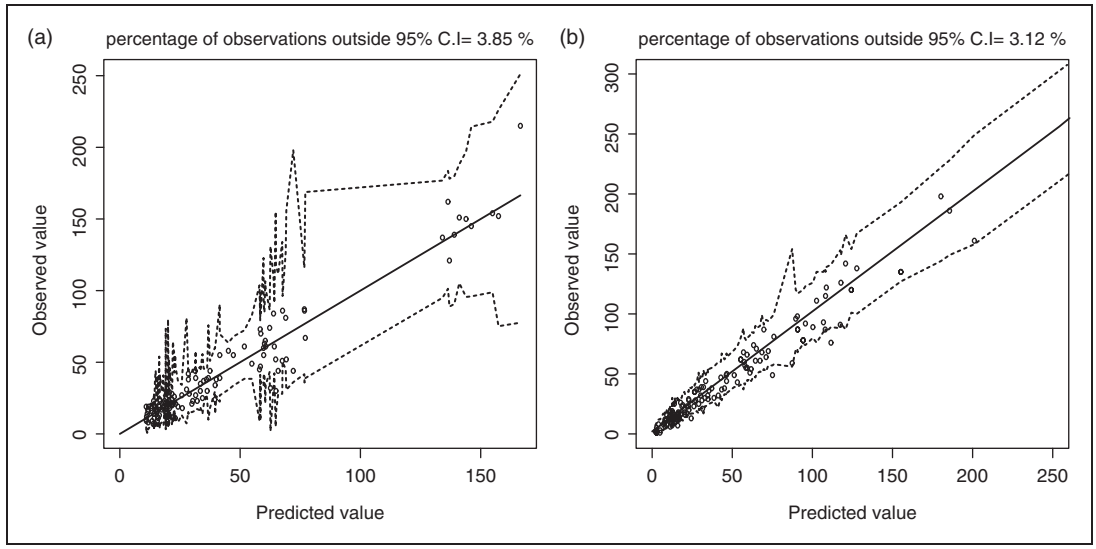

Figure 5. Model diagnostics for LBW data. The dotted lines show the 95% prediction intervals. (a) First design, (b) Second design.
LBW: Low birth weight.

analysis for each of the 10 observation times. For each case, the 160 estimated LBW counts $\hat{Y}_{it}$ are obtained and compared with the observed counts Y_{it}. Figure 5 presents the calibration plots for the LBW. The percentage of the observations that are outside the 95% intervals is 3.85% for the first calibration design (Figure 5(a)) and 3.12% for the second one (Figure 5(b)). We thus concluded our proposed model in this application performed well in terms of the spatial forecasting ability.

To explore whether the random walk process assumption in the coefficients considered is reasonable in this data set, we refit the proposed model with 8 spatial groups where the coefficients have an autoregressive distribution with order 1, AR(1), ($\beta_{gtp} \sim N(\rho_{gp}\beta_{g,t-1,p}, \sigma^2_{gp})$, $0 < \rho_{gp} < 1$). The proposed model with an AR(1) distribution has larger DIC (10330) and MSPE (154) and smaller MPL (−5338) than the model with a random walk process. Thus, a random walk process is more appropriate for the coefficient time-dependent structure in this data set.

6 Conclusion

In this article, we introduced a novel and flexible Bayesian latent model for space-time health data that accounts for the spatio-temporally varying coefficients. The proposed model captures the locally varying temporal behavior of covariate effects. We show, using low birth weight incidence data, that the latent model with spatio-temporally varying coefficients proposed in this paper provides a substantial improvement over other competing models that do not have such properties in terms of a range of model comparison measures.

The spatio-temporal latent model introduced here is the first step to illustrate the locally different temporal profiles of coefficients using a hierarchical framework. However, our approach has some limitations. In the real data analysis, the estimated number of groups in the proposed model is quite

large (8). It would be because the model assumed that the spatial groups are fixed across covariates. However, the spatial groups could vary with different covariates. If we would consider different number of spatial groups depending on covariates, then we would expect that the model would produce a smaller number of groups. Thus, the spatial grouping structure could be extended to the different spatial groups depending on covariates in the future. We also found that covariates in some spatial groups have little effects on LBW. So, the subset of covariates that have significant effects on health data could vary across spatial groups. As future work, we will consider the development of spatial variable selection approaches within a spatial latent model framework, which allows the subset of important covariates to vary across spatial groups. This research is particularly useful and valuable in spatial health effects studies because covariate effects are significantly different depending on areas as well as the subset of covariates can be identified within a subset of geographical areas.

Funding

This research was supported by NIH Grant R21 R21HL088654-01A2.

References

1. Knorr-Held L and Besag J. Modelling risk from a disease in time and space. *Stat Med* 1998; **17**: 2045–2060.
2. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. *Stat Med* 2000; **19**: 2555–2567.
3. Mugglin AS, Cressie N and Gemmell I. Hierarchical statistical modelling of influenza epidemic dynamics in space and time. *Stat Med* 2002; **21**: 2703–2721.
4. Tzala T and Best N. Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality. *Stat Meth Med Res* 2008; **17**: 97–118.
5. Gelfand AE, Banerjee S and Gamerman D. Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics* 2005; **16**: 465–479.
6. Lawson AB, Song HR, Cai B, et al. Space-time latent component modeling of geo-referenced health data. *Stat Med* 2010; **29**: 2012–2027.
7. Escobar MD and West M. Bayesian density estimation and inference using mixtures. *J Am Stat Assoc* 1995; **90**: 577–588.
8. Reich BJ and Bondell HD. A spatial Dirichlet process mixture model for clustering population genetics data. *Biometrics* 2011; **67**: 381–390.
9. Choi J, Lawson AB, Cai B, et al. Evaluation of Bayesian spatial-temporal latent models in small area health data. *Environmetrics* 2011; **22**: 1008–1022.
10. Celeux G, Forbes F, Robert C, et al. Deviance information criteria for missing data models. *Bayesian Anal* 2006; **1**: 651–674.
11. Congdon P. *Bayesian models for categorical data*. New York: John Wiley and Sons, 2005.
12. Kirby RS, Liu J, Lawson AB, et al. Spatio-temporal patterning of small area low birth weight incidence and its correlates: a latent spatial structure approach. *J Spatial Spatio-temporal Epidemiol* 2011; **2**: 265–271.
13. Kotelchuck M. An evaluation of the Kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index. *Am J Public Health* 1994; **84**: 1414–1420.
14. Banerjee S, Carlin BP and Gelfand AE. *Hierarchical modeling and analysis for spatial data*. New York: Chapman and Hall, 2004.
15. Besag J, York J and Mollie A. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Ann Inst Stat Math* 1991; **4**: 1–59.
16. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 2006; **1**: 515–533.
17. Stephens M. Dealing with label switching in mixture models. *J Roy Stat Soc B* 2000; **62**: 795–809.
18. Spiegelhalter DJ, Best N, Carlin BP, et al. Bayesian measures of model complexity and fit (with discussion). *J Roy Stat Soc B* 2002; **64**: 583–639.
19. Geisser S and Eddy W. A Predictive approach to model selection. *J Am Stat Assoc* 1979; **74**: 153–160.
20. Gelfand AE and Dey DK. Bayesian model choice: asymptotics and exact calculations. *J Roy Stat Soc B* 1994; **56**: 501–514.
21. Ibrahim J, Chen MH and Sinha D. *Bayesian survival analysis*. New York, NY: Springer, 2001.
22. Geweke J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernado JM, Berger JO, Dawid AP and Smith AFM (eds) *Bayesian statistics 4*. Oxford: Oxford University Press, 1992, pp.169–193.
23. Cogswell ME and Yip R. The influence of fetal and maternal factors on the distribution of birthweight. *Semin Perinatol* 1995; **19**: 222–240.
24. Bushy A. Health issues of women in rural environments: an overview. *J Am Med Women Assoc* 1998; **53**: 53–56.