Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

A Spatio-Temporal analytical outlook of the exposure to air pollution and COVID-19 mortality in the United States

summary

The world is experiencing a pandemic due to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), also known as COVID-19. The United States is also suffering from a catastrophic death toll from COVID-19. Several studies are providing preliminary evidence that short and long term exposure to air pollution might increase the severity of COVID-19 outcomes, including a higher risk of death. In this study, we develop a spatio-temporal model to estimate the association between exposure to fine particulate matter $PM_{2.5}$ and mortality accounting for several social and environmental factors. More specifically, we implement a Bayesian Zero Inflated Negative Binomial regression model with random effects that vary in time and space. Our goal is to estimate the association between air pollution and mortality accounting for the spatio-temporal variability that remained unexplained by the measured confounders. We applied our model to four US states with weekly data available for each county within each state. We analyze the data separately for each state, because each state shows a very different disease spread pattern. We found a positive association between long-term exposure to the $PM_{2.5}$ and the mortality from the COVID-19 disease for all four states but not all cases are statistically significant. Data and code are available here[add web link].

1 | INTRODUCTION

The world is experiencing an enormous death tool from COVID-19. The number of COVID-19 cases and deaths vary spatially and temporally, and can be affected by many factors, some local some global. There is a large body of literature that investigates the key biological, socioeconomic, and environmental factors that might increase the degree of severity of the health outcomes after having contracted COVID-19. With respect to the environmental factors, it is well-established that short and long term exposure to air pollution increases the risk of several chronic diseases, including respiratory, cardiovascular and respiratory diseases, irrespective of COVID-19 [1,2]. We and others [3,4,5,6,7,8,9] have hypothesized that exposure to air pollution increases the severity of COVID-19 outcomes, because air pollution can affect our immune, respiratory and cardiovascular system. This is a rapidly evolving area of research, see for example [10] for a review of the epidemiological studies on this topic.

In this paper we introduce a Bayesian spatio-temporal model to estimate the association between long-term exposure to $PM_{2.5}$ and COVID-19 health outcomes. To address this scientific question we need to overcome several challenges. These include: 1) large number of zero counts, especially at the beginning of the pandemic; 2) complex spatio-temporal variation that remained unexplained after having accounted for several measured confounders; and 3) computational feasibility. To overcome the challenges listed above and many others, we introduce a Bayesian model with multivariate spatio-temporal distributions of random effects and which also accounts for several measured socio-economic and demographic factors. We modelled the COVID-19 death counts via a zero-inflated negative binomial (ZINB) distribution [11]. Since the frequentist approach to fitting the ZINB model is challenging for longitudinal, spatial and spatio-temporal data, particularly when the model includes multivariate spatial random effects, we have chosen a more tractable Bayesian approach proposed by Neelon (2019)[11].

We apply our Bayesian model to a data set that includes daily and county level number of deaths, air pollution levels and many other potential confounders for several US states for the period 2000 to 2016. We incorporate the spatial and spatio-temporal information into the model by assigning a multivariate intrinsic conditionally autoregressive (ICAR) prior structure[12] to the random effects. Additional time fixed effects are also considered. Then, we analyze the effectiveness of the zero inflated model compared to an ordinary heterogeneous negative binomial model.

Wu *et al.* (2020)[13] also considered a ZINB model at a county level to investigate the association between long-term exposure to $PM_{2.5}$ COVID-19 deaths using an ecological and cross sectional study design. These authors also considered state specific random effects to capture variation between states. While Wu *et al.* (2020) investigated global association by using the data over most of US counties, no spatial dependence nor temporal dependence were assumed in the model, which is the main difference from our modeling. Instead, we decided to analyze spatio-temporal data within each state separately to account for heterogeneous dynamics of the spread of disease across states. For our analysis, we chose four states (Illinois, Florida, California, Georgia) as exemplary states which shows some zero-inflation over counties given the study period. The results with our model show overall positive association between long-term exposure to the ($PM_{2.5}$) and the mortality from the COVID-19 disease, which matches with other previous studies[10].

The rest of the paper is organized as follows. In section 2, we describe and visualize the data set for four states in the US. In section 3 we present the methodology, including how we leverage the spatial and spatio-temporal information to estimate the COVID-19 spread. In this section we also compare different statistical models In section 4, we apply the methods to the data set. In section 5, we discuss the results and comment on future directions.

2 | DATA

2.1 | COVID-19 death counts

We accessed the data from the repository maintained by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). We obtained daily number of deaths in each county for 4 states from the period March 23, 2020 to August 31, 2020. Please note that March 22, 2020 is the start date in this data source. We chose Illinois, Florida, California and Georgia as four examples to illustrate the features of our modelling approach under different scenarios of temporal and spatial dynamic of the COVID-19 deaths. For spatio-temporal models, we computed 23 weekly COVID-19 death counts from March 23 to August 31, 2020 by aggregating daily counts from Monday to Sunday without overlapping. For spatial models, which we consider for comparison study, we used cumulative counts since the beginning of the study period. For the sensitivity analysis, we adjusted starting week of day for deriving weekly outcome in spatio-temporal models, as well as the length of accumulation of an outcome in spatial models. Instead of considering the ratio of a COVID-19 count to a county level population size, the county level population size was considered as an offset variable in the model to remove the obvious effect of the county level population size on the COVID-19 death counts.

2.2 | Exposure to particulate matter ($PM_{2.5}$)

We imported the data from the repository where code and data are publicly available for reproducing analyses in "Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study"[13]. The county-level long-term averaged $PM_{2.5}$ ($\mu g/m^3$) from 2000 to 2016 used in Wu *et al.* (2020)[13] is calculated from an established exposure prediction models[14,15]. Thus, weekly variations of $PM_{2.5}$ during our study period for each county were not considered.

2.3 | Potential confounders

The nine potential risk factors or confounding variables were selected with the benchmark of the previous study[13]. The list of variables includes percent of poverty, population density (sq mi), median house value (thousand $), median household income (thousand $), percent of owner-occupied housing, percent of Hispanic population, percent of Black population, percent of the adult population with less than a high school education, and percent of the adult population older than age 65. These were collected from the 2000 and 2010 Census (`https://www.census.gov`) and the 2005 - 2016 American Community Surveys (`https://www.census.gov/programs-surveys/acs/`) according to Wu *et al.* (2020). Note that these variables do not vary over week as well.

3 | METHODS

In this paper, we will explore spatio-temporal modeling of the count data for our main analysis. Weekly counts of death cases over regions(counties) within a state as spatio-temporal outcomes were considered. Depending on the overdispersion and zero inflation characteristics in the outcome, we shall consider a negative binomial (NB) and a zero inflated negative binomial (ZINB) distribution based modeling approach. We will compare these results with the results from the spatial model with cumulative counts to provide general understanding of our implications of findings.

3.1 | Models

3.1.1 | Spatial Negative Binomial Model

Let y_i represent the death count in the county i at a certain date or the aggregated count for a certain period. We model these cross-sectional spatial count data with a negative binomial distribution[11] specified as,

$$y_i \sim \text{NB}\left(p_i, r\right) \tag{1}$$

where p_i represents the success probability in negative binomial distribution for the county i and r is a dispersion parameter. We model y_i as a generalized linear mixed model with a logistic link function. We assume D fixed-effect covariates including exposure to $PM_{2.5}$, population density, age distribution and several socio-economic variables. County-specific spatial random intercepts, b_i, are introduced to allow spatial dependence among the counties. The model is then defined as

$$\text{logit}\left(p_i\right) = \theta_i = \log(x_{oi}) + \beta_0 + \sum_{d=1}^{D} \beta_d X_d + b_i, \tag{2}$$

where x_{oi} is a population size of the i^{th} county so that $\log(x_{oi})$ indicates an offset variable. $\boldsymbol{\beta} = \left(\beta_0, \dots, \beta_d\right)^T$ represents the coefficient vector for the fixed effect covariates, and $b_1, \dots, b_n$ represent the random intercepts. Note that a positive value of β_d indicates an increase in the expected number of counts. We assign $\{b_i\}$ an intrinsic conditional autoregressive (ICAR) prior[12], which is specified by the following conditional structure:

$$b_i \mid b_{(-i)}, \sigma_r^2 \sim \text{N}\left(\frac{1}{m_i}\sum_{l\in\partial_i} b_l, \frac{\sigma_r^2}{m_i}\right), \tag{3}$$

where m_i is the number of neighbors of the i^{th} county, ∂_i is the set of indices for the neighbors of the i^{th} county and $b_{(-i)}$ is the set of random intercepts except the one for the i^{th} county. σ_r^2/m_i represents the conditional variance given the random intercepts corresponding to the rest of the counties. Note that we assume the first order neighbor structure. This model is similar to the model considered in Wu *et al.* (2020)[13] but we consider spatial random intercepts. This model is used to compare with a spatio-temporal model. In the rest of the paper, we shall refer to this negative binomial spatial model by SNB.

3.1.2 | Spatio-temporal Negative Binomial Model

The NB distribution based modeling of spatio-temporal count data can be expressed as,

$$y_{ij} \sim \text{NB}\left(p_{ij}, r\right), \tag{4}$$

where y_{ij} and p_{ij} are the count and the success probability corresponding to the negative binomial distribution for the i^{th} county at time t_{ij}, $j = 1, \dots, n_i$, respectively. We model p_{ij} as a generalized linear mixed model with the logistic link function in a following way:

$$\text{logit}\left(p_{ij}\right) = \theta_{ij} = \log(x_{oij}) + \beta_0 + \sum_{m=1}^{M} \beta_{1m} T_m + \sum_{d=1}^{D} \beta_{d+M} X_d + b_{i1} + b_{i2} t_{ij}. \tag{5}$$

x_{oij} is a population size of the i^{th} county at the time t_{ij} so that $\log(x_{oij})$ indicates an offset variable. $\sum_{m=1}^{M} \beta_{1m} T_m$ is a flexible non-linear time fixed effect using M cubic B-splines to capture the time trend fully. $\boldsymbol{\beta} = \left(\beta_0, \beta_1, \beta_2, \ldots \beta_{M+d}\right)^T$ represents the coefficient vector for the fixed effect covariates. $\boldsymbol{b}_i = \left(b_{i1}, b_{i2}\right)^T$ represents the spatial bivariate random effects for the i^{th} county. b_{i1} is a random intercept and b_{i2} is a random slope for a linear time trend. In this model we account for time trend both as a fixed effects and also as a random effects because there might be heterogeneity across counties in the temporal dynamic of COVID19 weekly deaths that remained unexplained by the time invariant covariates.

We modelled the county specific random effects vector $\boldsymbol{b}_i$, as a bivariate ICAR prior to incorporate spatial dependence in the intercept and slope of the linear county-specific time trends:

$$\boldsymbol{b}_i \mid \boldsymbol{b}_{(-i)}, \sigma_r^2 \sim \mathrm{N}_2\left(\frac{1}{m_i}\sum_{l\in\partial_i} \boldsymbol{b}_l, \frac{1}{m_i}\boldsymbol{\Gamma}\right), \tag{6}$$

where m_i is the number of neighbors for the i^{th} county and ∂_i is the set of neighbors for the i^{th} county. $\boldsymbol{\Gamma}/m_i$ is a 2×2 conditional covariance matrix of $\boldsymbol{b}_i$ given $\boldsymbol{b}_{(-i)}$, the random effects for the rest of the counties. We shall refer to this model as STNB in the rest of this paper.

3.1.3 | Spatio-temporal Zero Inflated Negative Binomial Model

In order to explain the zero inflation in the count data across different counties over time, we consider a spatio-temporal ZINB model[11] for y_{ij} as,

$$y_{ij} \sim \left(1 - q_{ij}\right) \mathbb{1}_{(\mathrm{w}_{ij}=0 \wedge y_{ij}=0)} + q_{ij}\mathrm{NB}\left(p_{ij}, r\right) \mathbb{1}_{(\mathrm{w}_{ij}=1)}, \tag{7}$$

where q_{ij} represents the probability that the i^{th} county at time t_{ij} belongs to the negative binomial component and w_{ij} represents the corresponding indicator variable. We can interpret q_{ij} as the probability that the i^{th} county at time t_{ij} potentially can have death counts. Thus, we call q_{ij} the probability-at-risk in short. The rest of the parameters are same as defined in (4). To consider both a population size offset and non-linear time fixed effect similar to the model in (5), we model q_{ij} and p_{ij} as,

$$\begin{aligned} \operatorname{logit}\left(q_{ij}\right) &= \operatorname{logit}\left[\Pr\left(\mathrm{w}_{ij} = 1 \mid \boldsymbol{\beta}_1, \boldsymbol{b}_{1i}\right)\right] \\ &= \theta_{1ij} = \log(x_{oij}) + \beta_{10} + \sum_{m=1}^{M} \beta_{1m} T_m + \sum_{d=1}^{D_1} \beta_{1(d+M)} Z_d + b_{1i1} + b_{1i2} t_{ij} \\ \operatorname{logit}\left(p_{ij}\right) &= \theta_{2ij} = \log(x_{oij}) + \beta_{20} + \sum_{m=1}^{M} \beta_{2m} T_m + \sum_{d=1}^{D_2} \beta_{2(d+M)} X_d + b_{2i1} + b_{2i2} t_{ij} \end{aligned} \tag{8}$$

where $\boldsymbol{b}_{1i} = \left(b_{1i1}, b_{1i2}\right)^T$ and $\boldsymbol{b}_{2i} = \left(b_{2i1}, b_{2i2}\right)^T$ represent the random effects corresponding to q_{ij} and p_{ij}, respectively. We impose a multivariate ICAR prior structure[11] on $\boldsymbol{\phi}_i = \left(\boldsymbol{b}_{1i}^T, \boldsymbol{b}_{2i}^T\right)^T$ as,

$$\boldsymbol{\phi}_i \mid \boldsymbol{\phi}_{(-i)}, \boldsymbol{\Gamma} \sim \mathrm{N}_4\left(\frac{1}{m_i}\sum_{l\in\partial_i} \boldsymbol{\phi}_l, \frac{1}{m_i}\boldsymbol{\Gamma}\right), \tag{9}$$

where $\boldsymbol{\Gamma}/m_i$ is a 4×4 conditional covariance matrix. This multivariate ICAR allows spatial dependence and county specific random time trend within count components and at-risk components as well as dependence between them. This model will be referred to as STZINB(NLT) in the rest of this paper, where NLT refers to non-linear fixed time trend for q_{ij}.

We hypothesize that the assumption of non-linear fixed time trend for q_{ij} might not be necessary in the sense of model parsimony. Thus, we consider a simpler version by assuming linear time fixed effect in a binary component. That is,

$$\operatorname{logit}\left(q_{ij}\right) = \theta_{1ij} = \log(x_{oij}) + \beta_{10} + \beta_{11} t_{ij} + \sum_{d=1}^{D_1} \beta_{1(d+1)} Z_d + b_{1i1} + b_{1i2} t_{ij}. \tag{10}$$

This model will be referred to as STZINB(LT) in the rest of this paper, where LT refers to linear fixed time trend for q_{ij}.

3.2 | Bayesian Inference

3.2.1 | Prior specification and MCMC settings

We illustrate prior and hyper-parameter specifications, and Markov Chain Monte Carlo (MCMC) settings under the STZINB model framework. For the latent at-risk indicators w_{ij}, probability was given as $\exp(\theta_{1ij})/[1+\exp(\theta_{1ij})]$, where θ_{1ij} is defined as either equation (8) or (10). Prior distributions for $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ were assumed to be $N_p\left(\beta_0 = \mathbf{0}, \Sigma_0 = 100\mathcal{I}_p\right)$, respectively, where $\mathcal{I}_p$ is a $p \times p$ identity matrix. For r in negative binomial component, a uniform prior was considered. To construct time-basis functions $T_m, m = 1, \cdots, M$, we standardized time points t_{ij} to be ranged from 0 to 1, i.e. $t_{ij} = \frac{j}{J}$, where $J = 23$ is the number of weeks during the study period, and $j = 1, \cdots, J = 23$ is a week indicator. We set three internal knot points 0.25, 0.50, 0.75 considering 0 and 1 as boundaries, so that $M = 7$ throughout our analysis. DIC is calculated as described in Gelman et al.(2014) [16].

For each MCMC algorithm, we ran 3 chains, 11,000 iterations, and 1,000 burn-in. For each model, convergence of each model was determined by conventional MCMC diagnostics such as trace plots and Geweke z-statistics.

3.2.2 | Conditional Posterior Distribution and Model Fitting

A posterior sampling algorithm of STZINB is adopted from Neelon *et al.* (2019) [11], and it is straightforward to implement the algorithms for the other models as they are simpler models. As outlined in Neelon *et al.* (2019) [11], we need to update at-risk indicators $\boldsymbol{w}$, coefficients for the binary model component $\boldsymbol{\beta}_1$, coefficients for the count model component $\boldsymbol{\beta}_2$, a dispersion parameter for the Negative Binomial distribution r, the set of spatial random effects $\boldsymbol{\phi}$ with Γ. The following illustrate the steps of MCMC to update the parameters.

STEP1 Update at-risk indicators $\boldsymbol{w}$

Given current parameter values, we draw w_{ij} from a Bernoulli distribution with probability η_{ij} such that

$$\eta_{ij} = \frac{\Pr(y_{ij} = 0 | w_{ij} = 1)\Pr(w_{ij} = 1)}{\Pr(y_{ij} = 0 | w_{ij} = 1)\Pr(w_{ij} = 1) + \Pr(y_{ij} = 0 | w_{ij} = 0)\Pr(w_{ij} = 0)} = \frac{q_{ij}^r(1 - p_{ij})}{q_{ij}^r(1 - p_{ij}) + p_{ij}} \tag{11}$$

where q_{ij} and p_{ij} were defined in (8). Note that q_{ij} is the inverse logit of θ_{1ij}, and p_{ij} is the inverse logit of θ_{2ij}. To avoid numerical issue, we adjusted the sampled q_{ij} and p_{ij} to be within $(0.001, 0.999)$, respectively, in practice.

STEP2 Update $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right)$

To update $\boldsymbol{\beta}_1$, we draw a latent variable ξ_{1ij} from a Pólya-Gamma distribution $\text{PG}(1, \theta_{1ij})$ as shown in Polson *et al.* (2013) [17]. Given $\boldsymbol{w}$ and $\boldsymbol{\xi}_1$, the full conditional distribution of $\boldsymbol{\beta}_1$ is

$$\Pr\left(\boldsymbol{\beta}_1 | \boldsymbol{w}, \boldsymbol{\xi}_1\right) \propto \pi(\boldsymbol{\beta}_1) exp\left[-\frac{1}{2}(\boldsymbol{z}_1 - \boldsymbol{X}\boldsymbol{\beta}_1)^T \Omega_1 (\boldsymbol{z}_1 - \boldsymbol{X}\boldsymbol{\beta}_1)\right] \tag{12}$$

where X is a $n\times p$ design matrix, $\pi(\boldsymbol{\beta}_1)$ the prior distribution $N_p\left(\beta_0, \Sigma_0\right)$, $\boldsymbol{z}_1 = \frac{\boldsymbol{w}-1/2}{\xi_1}$, and $\Omega_1 = \text{diag}(\boldsymbol{\xi}_1)$ an $n\times n$ precision matrix. Conditional on $\boldsymbol{z}_1$, we update $\boldsymbol{\beta}_1$ from $N_p(\mu, \Sigma)$ where $\Sigma = \left(\Sigma_0^{-1} + X^T\Omega_1 X\right)^{-1}$, and $\mu = \Sigma\left(\Sigma_0^{-1}\beta_0 + X^T\Omega_1\boldsymbol{z}_1\right)$. We update $\boldsymbol{\beta}_2$ by similar process using the corresponding Pólya-Gamma distribution $\text{PG}(y_{ij} + r, \theta_{2ij})$ as shown in Pillow and Scott (2012) [18].

STEP3 Update r

We can use a random-walk Metropolis-Hastings step illustrated in Neelon *et al.* (2019) [11]. We present Metropolis-Hastings method with uniform prior because of efficiency in computation time.

STEP4 Update $\boldsymbol{\phi}$, Γ

Let $\boldsymbol{\phi}_{11} = (b_{111}, \cdots, b_{1n1})^T$ be the $n\times1$ vectors of random intercepts for the binary component, $\boldsymbol{\phi}_{12} = (b_{112}, \cdots, b_{1n2})^T$ be the $n\times1$ vectors of random slopes for the binary component, $\boldsymbol{\phi}_{21} = (b_{211}, \cdots, b_{2n1})^T$ be the $n\times1$ vectors of random intercepts for the count component, and $\boldsymbol{\phi}_{22} = (b_{212}, \cdots, b_{2n2})^T$ be the $n\times1$ vectors of random slopes for the binary component. Then $\boldsymbol{\phi} = (\boldsymbol{\phi}_{11}, \boldsymbol{\phi}_{12}, \boldsymbol{\phi}_{21}, \boldsymbol{\phi}_{22})^T$ is the $4n\times1$ collection of all random effects by definition. Under the STZINB(NLT) model illustrated in section (3.1.3), the conditional prior for $\boldsymbol{\phi}_{11}$, for instance, is

$$\Pr(\boldsymbol{\phi}_{11} | \boldsymbol{\phi}_{12}, \boldsymbol{\phi}_{21}, \boldsymbol{\phi}_{22}, \Gamma) \propto exp\left[-\frac{1}{2}(\boldsymbol{\phi}_{11} - \boldsymbol{\mu}_{11})^T \Sigma_{11} (\boldsymbol{\phi}_{11} - \boldsymbol{\mu}_{11})\right] \tag{13}$$

where $\Sigma_{11} = \left[\Gamma_{11} - \Gamma_{1,-1}\Gamma_{-1,-1}^{-1}\Gamma_{-1,1}\right]^{-1} Q$, $\boldsymbol{\mu}_{11} = \left[\left(\Gamma_{1,-1}\Gamma_{-1,-1}^{-1}\right) \otimes I_n\right] \boldsymbol{\phi}_{(-1)}$, $\Gamma_{1,-1}$, $\Gamma_1 1$ denotes the first element of Γ, $\Gamma_{(1,-1)}$ is the 1×3 vector comprising the first row of Γ with element 1 removed, $\Gamma_{(-1,-1)}$ is the 3×3 sub-matrix of Γ after removing row 1 and column 1, $Q = M - A$, $M = \text{diag}(m_1, \cdots, m_n)$ an $n \times n$ matrix with diagonal elements equal to the number of neighbors for each spatial unit, A is an $n \times n$ adjacency matrix with $a_{ii} = 0$, $a_{il} = 1$ if county units i and l are neighbors, and $a_{il} = 0$ otherwise. We update each vector of $\boldsymbol{\phi} = (\boldsymbol{\phi}_{11}, \boldsymbol{\phi}_{12}, \boldsymbol{\phi}_{21}, \boldsymbol{\phi}_{22})^T$ from its normal full conditional distribution based on (13) applying sum-to-zero constraints as needed. To update Γ, we use its conjugate inverse-Wishart full conditional.

4 | RESULTS

4.1 | Overall description

TABLE 1 Summary statistics for four states: Illinois, Florida, California, and Georgia. The time period is 2020/03/23-2020/08/31 and the total number of weeks is 23. The average of zero death proportions is the average of proportions of counties with no death count over weeks. The average death counts is the average of death counts over counties and weeks. The remaining rows give averaged quantities over counties. The number in the parenthesis is the corresponding standard deviation.

	Illinois	Florida	California	Georgia
Number of counties	102	67	58	159
Average of zero death proportions (%)	75.5(6.3)	42.1(16.6)	50.5(10.8)	57.5(13.7)
Average of death counts	3.4(28.0)	7.2(19.5)	9.7(35.8)	1.5(3.6)
Cumulative death counts by 2020/08/31	78.7(505.2)	167.0(360.4)	224.5(775.8)	34.6(65.7)
Population size (thousands)	126(526)	297(478)	666(1,455)	63(137)
Long-term ambient $PM_{2.5}$ $(\mu g/m^3)$	10.2(0.9)	9.3(0.8)	8.3(3.4)	11.0(0.9)
Poverty rate (%)	7.5(3.0)	10.3(3.7)	9.6(4.1)	13.3(6.0)
Population density (in sq mi)	339(935)	874(1,254)	1,542(3,227)	272(552)
Median house value[MHV] (in thousand $)	109.2(45.6)	145.0(64.5)	356.9(223.5)	112.9(50.7)
Median household income[MHI] (in thousand $)	53.8(10.5)	46.2(8.7)	60.0(18.0)	42.6(10.9)
Home owners rate (%)	79.5(4.7)	71.8(8.5)	63.2(9.3)	70.1(9.0)
Hispanic (%)	3.5(4.1)	12.2(11.7)	25.4(16.8)	5.5(4.7)
Less than high-school education (%)	17.3(5.8)	19.0(7.4)	18.3(10.7)	28.4(9.8)
Black (%)	2.9(5.3)	13.4(9.8)	2.6(2.6)	26.1(16.9)
Older than age 65 (%)	16.4(2.9)	18.4(6.9)	13.8(3.8)	13.5(3.5)

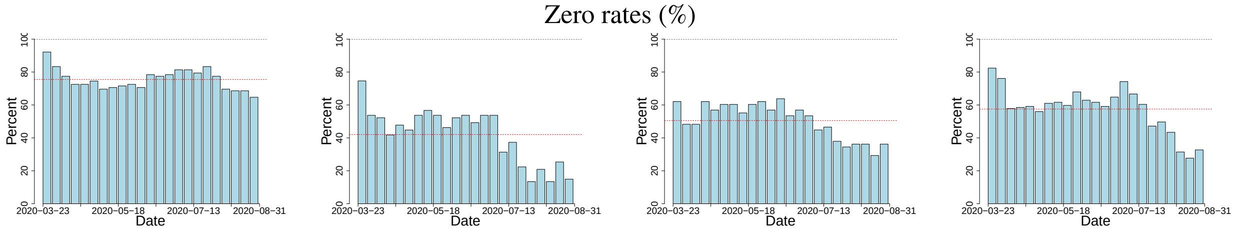
Table 1 shows summary statistics of our data for the four states in the US: Illinois (IL), Florida (FL), California (CA) and Georgia (GA). For example, 75.5 for Illinois in the row of the average of zero death proportions means on average 75.5% of counties in Illinois have zero death. 6.3 in the parenthesis is the standard deviation. 3.4 for Illinois in the row of the average of death counts means on average 3.5 deaths for each county during a week. High proportion of zero death counts implies that zero counts would come from two different sources: structural zero and zero from the negative binomial component. Structural zero component takes care of a possible excess zero proportion over zero proportion explained by the negative binomial distribution component. As evidence in the table, there is evidence of overdispersion because the sample variance of the COVID-19 death counts exceeds the sample mean in all the four states. Therefore, Zero-inflated negative binomial (ZINB) models could be suitable to account for the zero counts and also the overdispersion.

The four states have different characteristics of COVID-19 death counts until August. Table 2 illustrates two types of bar plots representing the weekly rates of zeros and the weekly death counts averaged across counties. Illinois has the highest zero rates over time. For the bar plot of the weekly death counts averaged across counties, Illinois is left-skewed while zero rates over time are constantly high compared to the other three states. Florida has the lowest zero rates across counties over time, and zero rates decreased over time in general. The bar plot of the averaged weekly death counts in Florida has a small peak in left-side, and

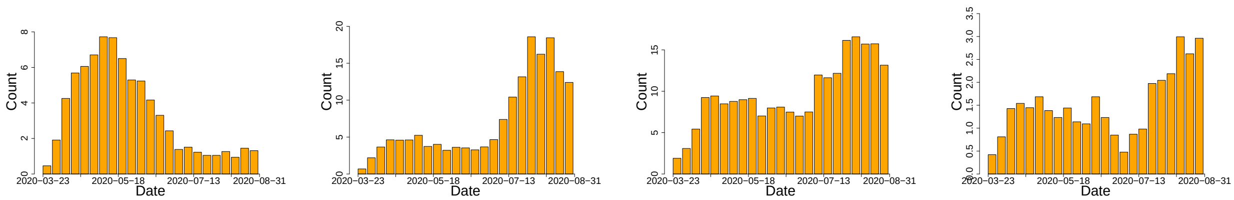
TABLE 2 Bar plots of county-aggregated weekly COVID-19 death counts during the study period (2020/03/23-2020/08/31). In the first row, the height of each bar represents the proportion of counties with no death count (0-100%) for each week, and the red dash line is the global average of zero rates across time (week). In the second row, the height of each bar represents the weekly death counts averaged across all counties. Please note that the the ranges of y-axis are different among the four states.

Illinois	Florida	California	Georgia

Zero rates (%)



Weekly death counts

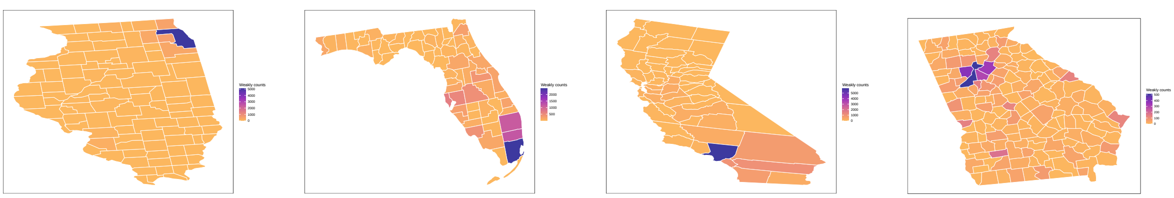


a large peak in the right-side. California and Georgia also have decreasing zero rates across counties over time, and increasing averaged weekly counts over time but the zero rates are overall higher than that of Florida. The bar plots of the averaged weekly death counts for both states also show bi-modal shapes with a higher right peak.

TABLE 3 Spatial visualization of COVID-19 weekly death counts and long-term ambient $PM_{2.5}$ concentration for the four states.

Illinois	Florida	California	Georgia

Weakly counts on the last week of August, 2020



Long-term ambient $PM_{2.5}$

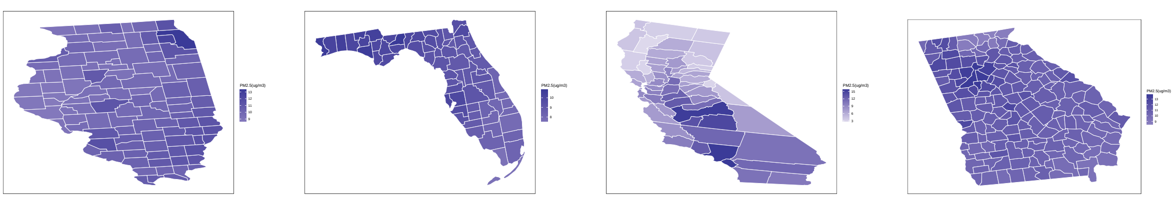


Table 3 provides the spatial visualization of weekly COVID-19 death counts (Last week of August 2020) and long-term ambient $PM_{2.5}$ levels for the four states (IL, FL, CA, GA). For COVID-19 counts, we used a viridis color scale to show the different level of counts across counties. Dark purple color represents a high level of COVID-19 counts, while orange color represents a low level of COVID-19 counts. For $PM_{2.5}$ concentration, we used the default custom color palette. Light purple color represents a low level of $PM_{2.5}$ concentration, while dark purple color represents a high level of COVID-19 counts. We hypothesized that the COVID-19 counts and long-term $PM_{2.5}$ exposure were positively associated. This visualization implies that counties that has relatively a high level of long-term ambient $PM_{2.5}$ concentration show a high level of COVID-19 counts.

4.2 Estimation results of the models on COVID-19 mortality

TABLE 4 Point estimates and 95% credible regions of the fixed effects for the model of the expected weekly number of COVID-19 deaths. Results are obtained by fitting a STZINB(LT) given in (10) for Illinois and Florida and by fitting a STZINB(NLT) given in (8) for California and Georgia.

		Illinois		Florida		California		Georgia	
		STZINB(LT)		STZINB(LT)		STZINB(NLT)		STZINB(NLT)	
$\beta_{2,2}$	$PM_{2.5}$	0.224	(-0.33, 0.86)	0.037	(-0.18, 0.31)	0.266	(-0.06, 0.71)	0.319	(0.00, 0.76)
$\beta_{2,3}$	Poverty	-0.288	(-0.85, 0.14)	-0.070	(-0.29, 0.11)	0.420	(-0.16, 1.22)	-0.164	(-0.38, 0.01)
$\beta_{2,4}$	Population density	0.302	(-1.15, 1.98)	0.119	(-0.28, 0.65)	0.649	(-1.34, 3.49)	-0.185	(-0.48, 0.06)
$\beta_{2,5}$	MHV	-0.313	(-1.12, 0.54)	-0.521	(-1.40, 0.06)	1.359	(0.35, 2.59)	0.160	(-0.02, 0.38)
$\beta_{2,6}$	MHI	0.731	(-0.21, 2.17)	0.448	(-0.07, 1.27)	-0.826	(-2.39, 0.27)	-0.024	(-0.29, 0.28)
$\beta_{2,7}$	Home owners rate	-0.504	(-1.19, 0.02)	-0.299	(-1.16, 0.20)	1.836	(-0.64, 5.03)	-0.336	(-0.67, -0.07)
$\beta_{2,8}$	Hispanic	0.349	(-0.02, 0.78)	0.173	(-0.07, 0.65)	1.775	(-1.04, 5.87)	0.104	(-0.55, 0.33)
$\beta_{2,9}$	Education	0.268	(-0.07, 0.83)	-0.406	(-1.11, 0.13)	-0.552	(-3.25, 1.21)	-0.076	(-0.27, 0.09)
$\beta_{2,10}$	Black	0.375	(0.04, 0.88)	0.005	(-0.16, 0.15)	0.660	(0.12, 1.32)	0.202	(0.01, 0.46)
$\beta_{2,11}$	Older than age 65	-0.310	(-1.07, 0.38)	0.611	(-0.04, 1.69)	-0.613	(-2.35, 0.96)	-0.074	(-0.08, 0.22)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-linear time trend) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear time trend)
MHV : Median House Value
MHI : Median Household Income

Among the Bayesian spatio-temporal models such as STNB, STZINB(NLT), and STZINB(LT), one model that shows the lowest Deviance information criterion (DIC) is chosen for each of the four states. STZINB(LT) is chosen for Illinois and Florida, while STZINB(NLT) is chosen for California and Georgia. Note that we include a non zero-inflated model for comparison and the zero-inflated models were selected for all the states we considered in this study with given study period.

Table 4 describes the estimated coefficients and their 95% credible intervals of the covariates on COVID-19 death counts under the selected model. The results for the four states indicate that long-term exposure to $PM_{2.5}$ is positively associated with the expected COVID-19 death counts but only Georgia shows significance based on 95% credible interval. The point estimate of β_2, 2 that corresponds to $PM_{2.5}$ for Georgia is equal to 0.319, which is the change in the expected weekly COVID-19 death counts in log scale by a unit change in $PM_{2.5}$. That is, we found that an increase of 1 $\mu g/m^3$ in the long-term $PM_{2.5}$ level is associated with a $e^{0.319} - 1 \simeq 37.6\%$ increase in the expected weekly COVID-19 death counts after controlling for all the confounding factors. Similarly, we have 25.1% for Illinois, 3.8% for Florida and 30.5% for California increment in the expected COVID-19 death weekly counts per county.

The results show that the increase in population density is associated with the increase in the expected COVID-19 death counts except Georgia, but none of them are statistically significant given 95% credible intervals. Data from the four states also show that the increase in Hispanic and Black population is associated with the increase in the expected COVID-19 death counts. The effect of Black population is statistically significant except Florida. On the other hand, only Florida shows that the increase in population older than age 65 is associated with the increase in the expected COVID-19 death counts.

Table 5 illustrates the estimated coefficients and their 95% credible intervals of covariates on q_{ij}, the probability-at-risk of COVID-19 death under the selected model. Here also, the results show that long-term ambient $PM_{2.5}$ is positively associated with the probability of COVID19 death. The effect of the long-term ambient $PM_{2.5}$ on the probability of COVID19 death is statistically significant for California and Georgia. The direction of the effect is coherent with the results for the counts component. The directions of the effect for Hispanic and Black population are overall similar. The effect of population older than 65 on the probability-at-risk has the same direction compared to the effect on the expected COVID-19 death counts.

TABLE 5 Point estimates and 95% credible regions of the fixed effects on the probability-at-risk(q_{ij}), the probability belong to the negative binomial component) for the expected weekly number of COVID-19 deaths. Results are obtained by fitting a STZINB(LT) given in (10) for Illinois and Florida and by fitting a STZINB(NLT) given in (8) for California and Georgia.

		Illinois		Florida		California		Georgia	
		STZINB(LT)		STZINB(LT)		STZINB(NLT)		STZINB(NLT)	
$\beta_{1,2}$	$PM_{2.5}$	0.419	(-0.44, 1.48)	0.871	(-0.26, 2.60)	0.490	(0.22, 0.83)	0.572	(0.12, 1.20)
$\beta_{1,3}$	Poverty	-0.011	(-0.96, 0.86)	-0.238	(-1.01, 0.69)	-0.045	(-0.05, 0.42)	0.113	(-0.13, 0.48)
$\beta_{1,4}$	Population density	3.172	(-1.45, 9.30)	2.632	(-0.91, 7.20)	-0.381	(-1.52, 0.96)	-0.018	(-0.55, 0.58)
$\beta_{1,5}$	MHV	1.064	(-0.61, 3.36)	0.127	(-1.14, 2.08)	0.973	(-0.34, 2.18)	0.235	(-0.25, 0.70)
$\beta_{1,6}$	MHI	-0.924	(-3.59, 0.97)	0.022	(-1.26, 1.29)	0.317	(-0.73, 1.50)	-0.020	(-0.41, 0.44)
$\beta_{1,7}$	Home owners rate	-0.273	(-1.72, 1.27)	0.630	(-0.60, 2.05)	-0.680	(-1.98, 0.41)	0.096	(-0.26, 0.64)
$\beta_{1,8}$	Hispanic	0.660	(-0.34, 1.74)	1.479	(0.39, 3.22)	-0.308	(-2.01, 1.08)	0.026	(-0.26, 0.26)
$\beta_{1,9}$	Education	-0.472	(-1.37, 0.30)	-0.433	(-1.62, 0.66)	0.619	(-0.41, 1.71)	-0.095	(-0.34, 0.20)
$\beta_{1,10}$	Black	-0.075	(-1.11, 0.96)	0.781	(0.06, 1.94)	0.673	(-0.01, 1.54)	0.090	(-0.20, 0.39)
$\beta_{1,11}$	Older than age 65	-0.309	(-2.00, 1.63)	0.598	(-0.60, 1.86)	-0.133	(-1.30, 1.20)	-0.084	(-0.36, 0.20)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-linear time trend) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear time trend)
MHV : Median House Value
MHI : Median Household Income

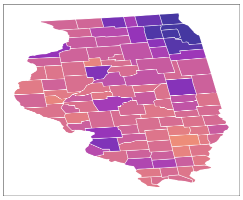
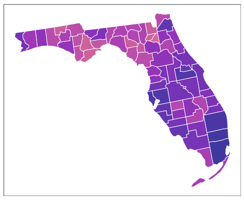
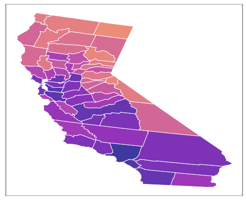
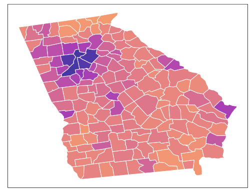
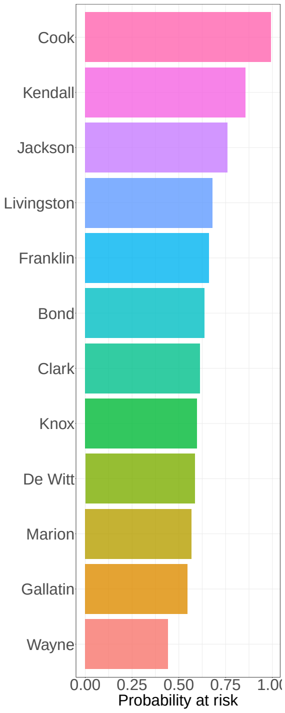
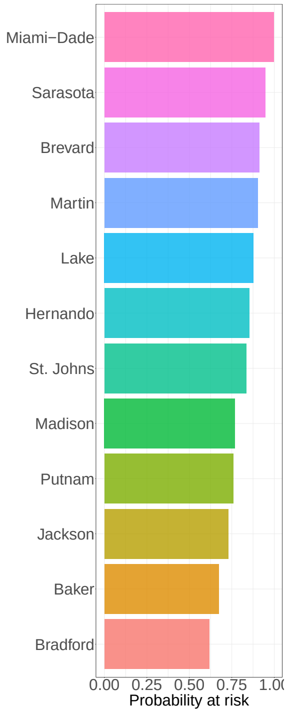
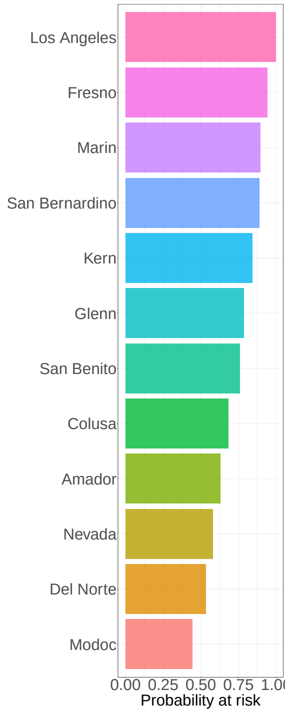
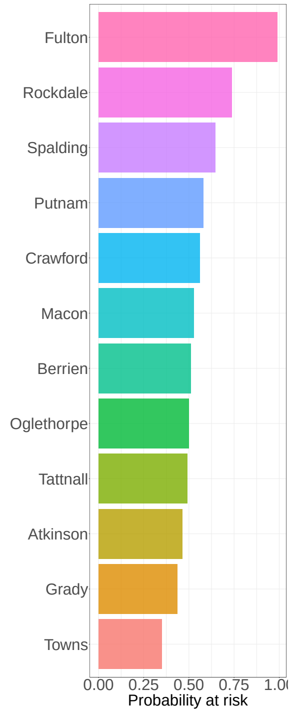
If the long-term ambient $PM_{2.5}$ increases in 1 $\mu g/m^3$ adjusting the other risk factors in California, for example, the log odds ratio for the probability-at-risk (q_{ij}) increases in 0.490. This implies if the long-term ambient $PM_{2.5}$ changes from 8 $\mu g/m^3$ to 9 $\mu g/m^3$, the probability-at-risk will increase in $\exp(9 \cdot 0.490)/[1+\exp(9 \cdot 0.490)] - \exp(8 \cdot 0.490)/[1+\exp(8 \cdot 0.490)] \simeq 0.007$ in California. On the last week of August, COVID-19 probability -at-risk over counties in California are ranged from 0.442 in Modic County to 0.997 in Los Angeles County.

Table 6 shows the time-averaged probability-at-risk. Note that q_{ij} represents the probability of the i^{th} county being at risk (Negative Binomial component) at j-th week which we can obtain during the MCMC procedures. The dark colored counties for each state indicates a high time-averaged probability. In the second row, we shows 12 representative counties out of all counties after sorting all counties by the estimated probability-at-risk with consideration of equidistant rank. The difference between the first and the second county is larger in Illinois and Georgia than that in the other two states. The decreasing rate over counties is slower in Florida compared to the other states.

The estimated Γ, the covariance matrix of four spatial random effects in each state shows that off-diagonal entries are close to zero (Figure 9 in the Appendix). Recall that these random effects are random intercepts and random slopes for time in the count component and at-risk probability component. Thus, estimated zero implies these random effects are not correlated to each other, although each random effects are spatially dependent.

Table 7 shows the estimated nonlinear mixed time effects in the Negative Binomial component (the COVID-19 death counts). Each of the four states shows different nonlinear temporal effect patterns in the Negative Binomial component. We selected five representative counties out of twelve appeared in Table 6 considering the rank of exposed COVID-19 probability-at-risk, and they show distinctive patterns over study periods since we allow county specific random effects in intercept and linear component.

TABLE 6 Visualization of time-averaged COVID-19 death probability-at-risk using the selected STZINB models for the four states. (IL, FL, CA, GA)

Illinois	Florida	California	Georgia
			
			

4.3 | Comparison by modeling techniques

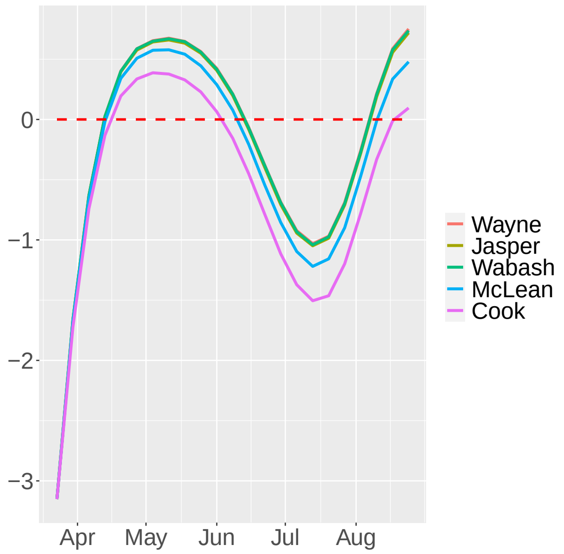
A spatial negative binomial model (SNB) with cumulative death counts is considered to compare with spatio-temporal models. The results are provided in Table 8. Since the models are different as well as the response variables are different, we can not directly compare the results between the SNB model and spatio-temporal models. However, we can see whether direction/uncertainty of association is consistent or not between these models. The long-term exposure to $PM_{2.5}$ is positively associated with the expected cumulative death counts for COVID-19 for all the four states from the SNB model, which is consistent with the results from the spatio-temporal models provided in the previous subsection. This supports the hypothesis of positive association between the long-term ambient $PM_{2.5}$ and the expected COVID-19 death counts.

On the other hand, some results show different uncertainty. For example, The long-term ambient $PM_{2.5}$ is not significantly associated with the expected COVID-19 death counts using the STZINB model for California, while the association is significant using the SNB model. Another example is that the percentage of Hispanic population is negatively associated with the expected COVID-19 death counts in Illinois from the SNB model, while it is positively associated with the expected COVID-19 death counts in Illinois from the STZINB(LT) model. Although there are some differences, overall direction/uncertainty of the estimated effects are similar in both SNB model and spatio-temporal models.

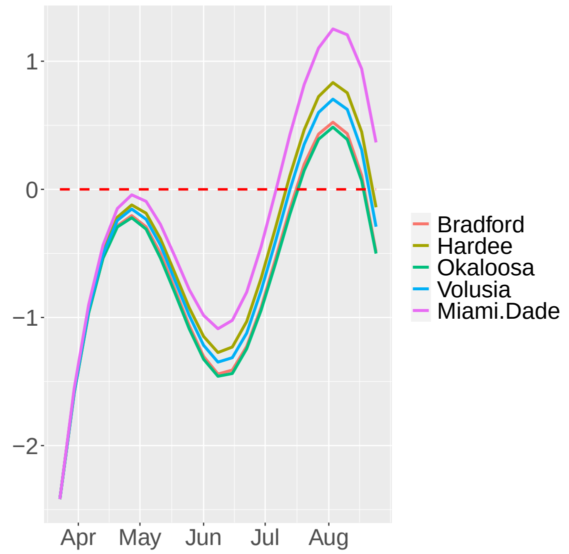
The estimated effects are rather sensitive to the length of the period for aggregating the outcome variable, which could be an issue to consider a SNB model for COVID-19 death counts. The estimated effects using the data with different lengths are provided in Figure 20 in the Appendix. In addition to this issue, a SNB model ignores temporal characteristics of the data. Thus,

TABLE 7 The median value of estimated nonlinear mixed time effect in the Negative Binomial component (the COVID-19 death counts) using the selected model structure. The x-axis represents timeline (weeks). We show five representative counties in each state for visualization. A red dash line is a zero reference. The ranges of y-axis are different among the four states.

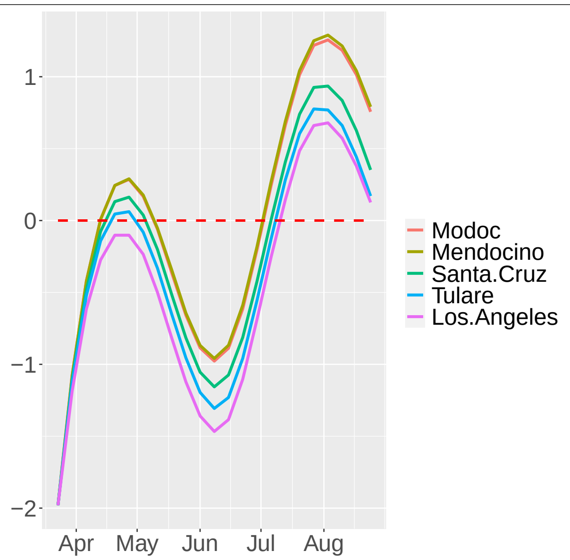
Illinois



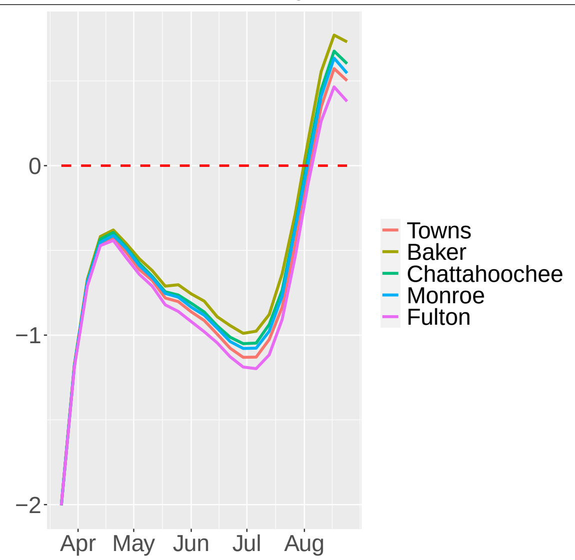
Florida



California



Georgia



it is natural to consider a spatio-temporal model, but we should be careful in interpreting the results from the spatio-temporal models since the model complexity can result in an over-fitting or unstable estimation.

5 | CONCLUSION AND DISCUSSION

We explored the relationship between long-term exposure to $PM_{2.5}$ and county level COVID-19 weekly death counts while adjusting for several social and environmental factors using spatio-temporal negative binomial models with/without a zero-inflated component. County random effects that incorporate spatial dependence by a ICAR model were considered to capture spatial

TABLE 8 The estimation results of effects on the cumulative death counts of COVID-19 using the spatial negative binomial model (section 3.1.1) for the four states (Illinois, Florida, California, and Georgia). The counts are aggregated from March 23, 2020 to August 31, 2020.

Model/Variable		Illinois		Florida		California		Georgia	
$\beta_{2,2}$	$PM_{2.5}$	1.226	(-3.47, 5.46)	0.228	(-0.14, 0.61)	0.306	(0.15, 0.45)	0.330	(0.08, 0.56)
$\beta_{2,3}$	Poverty	0.115	(-0.32, 0.55)	-0.135	(-0.41, 0.16)	0.156	(-0.18, 0.48)	-0.077	(-0.26, 0.10)
$\beta_{2,4}$	Population density	-0.051	(-0.37, 0.24)	0.536	(-0.47, 1.59)	-0.356	(-0.89, 0.20)	-0.049	(-0.43, 0.36)
$\beta_{2,5}$	MHV	-0.366	(-1.85, 1.19)	-0.348	(-0.68, -0.02)	1.088	(0.30, 2.03)	0.169	(-0.04, 0.39)
$\beta_{2,6}$	MHI	-0.090	(-1.04, 0.88)	0.392	(-0.05, 0.81)	-0.466	(-1.29, 0.30)	-0.179	(-0.46, 0.12)
$\beta_{2,7}$	Home owners rate	0.652	(-0.22, 1.50)	-0.193	(-0.59, 0.21)	-0.036	(-0.69, 0.53)	-0.120	(-0.33, 0.09)
$\beta_{2,8}$	Hispanic	-0.455	(-0.88, -0.03)	0.440	(0.11, 0.76)	0.419	(-0.29, 1.10)	0.054	(-0.11, 0.21)
$\beta_{2,9}$	Education	0.367	(-0.08, 0.79)	-0.112	(-0.41, 0.18)	-0.013	(-0.68, 0.57)	-0.095	(-0.30, 0.11)
$\beta_{2,10}$	Black	0.111	(-0.25, 0.53)	0.170	(-0.06, 0.41)	0.485	(0.11, 0.90)	0.120	(-0.06, 0.21)
$\beta_{2,11}$	Older than age 65	0.246	(-0.19, 0.73)	0.647	(0.34, 0.98)	-0.096	(-0.58, 0.39)	-0.028	(-0.25, 0.19)

MHV : Median House Value

MHI : Median Household Income

variation and spatio-temporal interaction in both structural zero component and negative binomial component. We considered possible non-linear time effects in both components as well.

The effects of long-term exposure to $PM_{2.5}$ and other social and environmental factors on the COVID-19 weekly death counts can be heterogeneous by state since socio-cultural, behavioral and healthcare system as well as COVID-19 policies could vary by state. Thus, we applied the spatio-temporal models to each state of consideration in this study (Illinois, Florida, California and Georgia). Based on model comparison by DIC, the zero-inflated models were selected for all the four states. Within zero-inflated models, the linear time trend model for the probability-at-risk was chosen for Illinois and Florida while the non-linear time trend model was chosen for the other two states. Note that we assumed non-linear time trend for negative binomial component for all the four states. These results were compared with the results by the spatial negative binomial model without temporal information. The spatial model was applied to the cumulative death counts until the date we considered (August 31, 2020), which is different from weekly death counts so that we cannot directly compare the estimated coefficients corresponding to risk factors. However, we can see how the results change after adjusting temporal information as well as zero-inflated characteristics. Also, note that a zero-inflated component for a pure spatial model with cumulative death counts may not be necessary anymore given that the chance of zero count for each county is getting low as the time period is increasing.

The long-term exposure to $PM_{2.5}$, even after adjusting the non-linear time trend with a county-specific random slope, spatial dependence as well as the other confounders, is positively associated with the COVID-19 weekly death counts for all the four states. The direction of association, although not all of them are significant, is consistent with the results of the previous study [13], which did not consider spatial and temporal dependence in the model, as well as the spatial only negative binomial model. These findings add another evidence of the increase in risk of death for COVID-19 by the long-term exposure to air pollution into the literature.

By applying the models to each state, separately, we are able to see different patterns among states. Although the effect of the long-term exposure to $PM_{2.5}$ on the COVID-19 weekly death counts is in the same direction for all the four states, the size of the effects are different. Also, the effects of the some other confounders show different direction by state. The proposed model, spatio-temporal zero-inflated negative binomial model with nonlinear time effects, spatial random effects and spatio-temporal interaction random effects is very flexible to capture the spatio-temporal characteristics of the data. On the other hand, complexity of the model could lead to the increased variability in estimation due to many parameters to estimate. This issue could be alleviated by controlling the prior distribution with the information from the previous study. We have developed a user friendly R tool which can be used to get inference from any state of the US. All model related R codes and software can be freely downloaded from GitHub repository. We are also in the process of developing a R-Shiny based application for cloud based deployment and interactive interface for non-statistician's easy use and access.

As the spread of COVID-19 is on-going, each state has different dynamics of the disease spread and a spatio-temporal model is flexible enough to handle different types of dynamics. As the surge of COVID-19, the zero-inflated model might not be suitable for many states anymore. However, by investigating and comparing several spatio-temporal models including non zero-inflated models, we can find a reasonable model that explain the characteristics of the data. The current risk factor and confounders we used are not temporally varying so that we introduced non-linear time effects with county specific random slopes to handle temporal variations due to unavailable spatio-temporal covariates such as policy changes by county health department on COVID-19, policy changes for school opening and business operation. Once this additional information is available, we can easily accommodate it into the model.

One can consider applying the spatio-temporal model to county level COVID-19 death counts for several states together, or the entire U.S. This can be doable but a single model may not be able to capture heterogeneous effects that we found in this study. Also, handling a spatial dependence model with a large number of spatial regions bring an additional computational burden. On the other hand, we can extend the current model with a spatio-temporally varying coefficient model to capture heterogeneous effects by states or by county so that we can investigate the whole data into one model framework. A modified spatial dependence modeling such as allowing spatial dependence only within state to increase computational efficiency can be also considered. We plan to investigate such extension as a future study.

References

1. Jiang X, Mei X, Feng D. Air pollution and chronic airway diseases: what should people know and do?. *Journal of Thoracic Disorder* 2016; 8: 31 - 40.

2. Lelieveld J, Munzel T. Air pollution, chronic smoking, and mortality. *European Heart Journal*; 38(40): 3204.

3. Chakrabarty RK, Beeler P, Liu P, et al. Ambient PM2. 5 exposure and rapid spread of COVID-19 in the United States. *Science of The Total Environment* 2020: 143391.

4. Liu P, Beeler P, Chakrabarty RK. Dynamic interplay between social distancing duration and intensity in reducing COVID-19 US hospitalizations: A "law of diminishing returns". *Chaos: An Interdisciplinary Journal of Nonlinear Science* 2020; 30(7): 071102. doi: 10.1063/5.0013871

5. Ogen Y. Assessing nitrogen dioxide (NO2) levels as a contributing factor to the coronavirus (COVID-19) fatality rate. *Science of The Total Environment* 2020: 138605.

6. Jiang Y, Xu J. The association between COVID-19 deaths and short-term ambient air pollution/meteorological condition exposure: a retrospective study from Wuhan, China. *Air Quality, Atmosphere & Health* 2020: 1–5.

7. Yongjian Z, Jingu X, Fengming H, Liqing C. Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Science of the total environment* 2020: 138704.

8. Conticini E, Frediani B, Caro D. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy?. *Environmental pollution* 2020: 114465.

9. Comunian S, Dongo D, Milani C, Palestini P. Air pollution and Covid-19: the role of particulate matter in the spread and increase of Covid-19's morbidity and mortality. *International journal of environmental research and public health* 2020; 17(12): 4487.

10. Bhaskar A, Chandra J, Braun D, Cellini J, Dominici F. Air pollution, SARS-CoV-2 transmission, and COVID-19 outcomes: A state-of-the-science review of a rapidly evolving research area. *medRxiv* 2020. doi: 10.1101/2020.08.16.20175901

11. Neelon B. Bayesian Zero-Inflated Negative Binomial Regression Based on Pólya-Gamma Mixtures. *Bayesian Anal.* 2019; 14(3): 829–855. doi: 10.1214/18-BA1132

12. Banerjee S, Carlin B, Gelfand A. *Hierarchical Modeling and Analysis of Spatial Data*. 101. ; 2004

13. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *medRxiv* 2020. doi: 10.1101/2020.04.05.20054502

14. Atmospheric Composition Analysis Group. http://fizz.phys.dal.ca/~atmos/martin/; .

15. Van Donkelaar A, Martin R, Li C, Burnett R. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental science & technology.* 2019; 53: 2595-2611.

16. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian data analysis.* CRC press . 2013.

17. Polson G, Scott G, Windle J. Bayesian Inference for Logistic Models Using Pólya–Gamma Latent Variables. *Journal of the American Statistical Association* 2013; 108: 1339-1349. doi: 10.1080/01621459.2013.829001

18. Pillow W, Scott G. Fully Bayesian inference for neural models with negative-binomial spiking. In: Bartlett P, Pereira F, Burges C, Bottou L, Weinberger K., eds. *Advances in Neural Information Processing Systems*. 25. MIT Press. 2012 (pp. 1907-1915).

APPENDIX

5.1 | Result: Random effects

Table 9 illustrates the estimated covariance elements and their 95% credible intervals under the selected model. This implies that the four spatial random effects of COVID-19 death counts were not significantly correlated to each other.

TABLE 9 The covariance of spatial random effects of COVID-19 death counts from March 23 to August 31 among the four states (Illinois, Florida, and California and Georgia).

Model/Variable	Illinois		Florida		California		Georgia	
Model	STZINB(LT)		STZINB(LT)		STZINB(NLT)		STZINB(NLT)	
$\Gamma_{1,1}$	0.010	(0.0073, 0.0127)	0.015	(0.0105, 0.0208)	0.018	(0.0114, 0.0245)	0.006	(0.0050, 0.0078)
$\Gamma_{1,2}$	0.000	(-0.0021, 0.0018)	0.000	(-0.0037, 0.0038)	0.000	(-0.0046, 0.0046)	0.000	(-0.0010, 0.0010)
$\Gamma_{1,3}$	0.000	(-0.002, 0.002)	0.000	(-0.0036, 0.0041)	0.000	(-0.0048, 0.0049)	0.000	(-0.0010, 0.0010)
$\Gamma_{1,4}$	0.000	(-0.002, 0.0019)	0.000	(-0.0037, 0.0037)	0.000	(-0.0045, 0.0051)	0.000	(-0.0010, 0.0010)
$\Gamma_{2,2}$	0.010	(0.0073, 0.0127)	0.015	(0.0106, 0.0212)	0.018	(0.0121, 0.0249)	0.006	(0.0050, 0.0078)
$\Gamma_{2,3}$	0.000	(-0.0021, 0.0019)	0.000	(-0.0037, 0.0038)	0.000	(-0.0051, 0.0047)	0.000	(-0.0009, 0.0010)
$\Gamma_{2,4}$	0.000	(-0.0021, 0.0018)	0.000	(-0.0036, 0.004)	0.000	(-0.0046, 0.0051)	0.000	(-0.0010, 0.0010)
$\Gamma_{3,3}$	0.010	(0.0074, 0.0127)	0.015	(0.0107, 0.0209)	0.018	(0.012, 0.0246)	0.006	(0.0050, 0.0078)
$\Gamma_{3,4}$	0.000	(-0.0019, 0.002)	0.000	(-0.0037, 0.0041)	0.000	(-0.0049, 0.0048)	0.000	(-0.0010, 0.0010)
$\Gamma_{4,4}$	0.010	(0.0074, 0.013)	0.015	(0.0101, 0.0207)	0.018	(0.0119, 0.0248)	0.006	(0.0050, 0.0078)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-linear time trend) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear time trend) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

5.2 Result: Illinois

TABLE 10 Parameter estimates and 95% credebile intervals for the spatio-temporal models in the Illinois study

Model		Variable	STZINB(NLT)		STZINB(LT)		STNB	
			Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
Binary	$\beta_{1,0}$	intercept	-2.484	(-12.53, 7.35)	0.090	(-12.97, 13.5)		
	$\beta_{1,1}$	time	-	-	-0.320	(-15.63, 14.21)		
	$\beta_{1,2}$	$PM_{2.5}$	0.309	(-0.14, 0.78)	0.419	(-0.44, 1.48)		
	$\beta_{1,3}$	poverty	-0.179	(-0.77, 0.31)	-0.011	(-0.96, 0.86)		
	$\beta_{1,4}$	Population density	1.857	(-0.68, 4.69)	3.172	(-1.45, 9.30)		
	$\beta_{1,5}$	MHV	0.774	(-0.39, 1.89)	1.064	(-0.61, 3.36)		
	$\beta_{1,6}$	MHI	-0.330	(-1.53, 0.66)	-0.924	(-3.59, 0.97)		
	$\beta_{1,7}$	Home owners rate	-0.428	(-1.31, 0.48)	-0.273	(-1.72, 1.27)		
	$\beta_{1,8}$	hispanic	0.719	(0.13, 1.40)	0.660	(-0.34, 1.74)		
	$\beta_{1,9}$	Education	-0.241	(-0.67, 0.14)	-0.472	(-1.37, 0.30)		
	$\beta_{1,10}$	Black	0.051	(-0.45, 0.53)	-0.075	(-1.11, 0.96)		
	$\beta_{1,11}$	Older than age 65	-0.754	(-1.51, -0.04)	-0.309	(-2.00, 1.63)		
Count	$\beta_{2,0}$	intercept	-2.576	(-13.95, 7.26)	-2.968	(-13.24, 7.98)	-2.838	(-12.74, 7.48)
	$\beta_{2,1}$	time	-	-	-	-	-	-
	$\beta_{2,2}$	$PM_{2.5}$	0.247	(-0.38, 0.97)	0.224	(-0.33, 0.86)	0.217	(-0.12, 0.64)
	$\beta_{2,3}$	poverty	-0.246	(-0.78, 0.22)	-0.288	(-0.85, 0.14)	-0.277	(-0.70, 0.06)
	$\beta_{2,4}$	Population density	0.478	(-1.04, 2.16)	0.302	(-1.15, 1.98)	0.003	(-1.00, 1.18)
	$\beta_{2,5}$	MHV	-0.249	(-1.12, 0.70)	-0.313	(-1.12, 0.54)	-0.374	(-0.97, 0.12)
	$\beta_{2,6}$	MHI	0.603	(-0.79, 1.81)	0.731	(-0.21, 2.17)	0.719	(0.16, 1.37)
	$\beta_{2,7}$	Home owners rate	-0.368	(-0.99, 0.14)	-0.504	(-1.19, 0.02)	-0.507	(-0.99, -0.02)
	$\beta_{2,8}$	hispanic	0.331	(-0.01, 0.80)	0.349	(-0.02, 0.78)	0.352	(0.08, 0.63)
	$\beta_{2,9}$	Education	0.322	(-0.11, 0.90)	0.268	(-0.07, 0.83)	0.201	(-0.06, 0.51)
	$\beta_{2,10}$	Black	0.398	(0.03, 0.91)	0.375	(0.04, 0.88)	0.275	(0.02, 0.60)
	$\beta_{2,11}$	Older than age 65	-0.062	(-0.71, 0.54)	-0.31	(-1.07, 0.38)	-0.331	(-0.88, 0.17)
	r	dispersion	1.728	(0.04, 4.59)	1.900	(0.02, 5.00)	2.187	(0.07, 5.54)
Random Effects	Γ_{11}		0.010	(0.0072, 0.0127)	0.010	(0.0073, 0.0127)		
	Γ_{12}		0.000	(-0.002, 0.002)	0.000	(-0.0021, 0.0018)		
	Γ_{13}		0.000	(-0.0019, 0.0019)	0.000	(-0.002, 0.002)		
	Γ_{14}		0.000	(-0.002, 0.0019)	0.000	(-0.002, 0.0019)		
	Γ_{22}		0.010	(0.0075, 0.0129)	0.010	(0.0073, 0.0127)		
	Γ_{23}		0.000	(-0.002, 0.002)	0.000	(-0.0021, 0.0019)		
	Γ_{24}		0.000	(-0.002, 0.002)	0.000	(-0.0021, 0.0018)		
	Γ_{33}		0.010	(0.0073, 0.0127)	0.010	(0.0074, 0.0127)	0.010	(0.0074, 0.0128)
	Γ_{34}		0.000	(-0.0019, 0.002)	0.000	(-0.0019, 0.002)	0.000	(-0.002, 0.0019)
	Γ_{44}		0.010	(0.0074, 0.0129)	0.010	(0.0074, 0.013)	0.010	(0.0075, 0.0129)
DIC			2998.091		2748.671		3246.834	

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 11 Comparison by both spatio-temporal models with weekly counts and spatial models with cumulative counts in the Illinois study

	Variable	STZINB (NLT)		STZINB (LT)		STNB		SNB	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.247	(-0.38, 0.97)	0.224	(-0.33, 0.86)	0.217	(-0.12, 0.64)	1.226	(-3.47, 5.46)
$\beta_{2,3}$	poverty	-0.246	(-0.78, 0.22)	-0.288	(-0.85, 0.14)	-0.277	(-0.70, 0.06)	0.115	(-0.32, 0.55)
$\beta_{2,4}$	Population density	0.478	(-1.04, 2.16)	0.302	(-1.15, 1.98)	0.003	(-1.00, 1.18)	-0.051	(-0.37, 0.24)
$\beta_{2,5}$	MHV	-0.249	(-1.12, 0.70)	-0.313	(-1.12, 0.54)	-0.374	(-0.97, 0.12)	-0.366	(-1.85, 1.19)
$\beta_{2,6}$	MHI	0.603	(-0.79, 1.81)	0.731	(-0.21, 2.17)	0.719	(0.16, 1.37)	-0.090	(-1.04, 0.88)
$\beta_{2,7}$	Home owners rate	-0.368	(-0.99, 0.14)	-0.504	(-1.19, 0.02)	-0.507	(-0.99, -0.02)	0.652	(-0.22, 1.50)
$\beta_{2,8}$	Hispanic	0.331	(-0.01, 0.80)	0.349	(-0.02, 0.78)	0.352	(0.08, 0.63)	-0.455	(-0.88, -0.03)
$\beta_{2,9}$	Education	0.322	(-0.11, 0.90)	0.268	(-0.07, 0.83)	0.201	(-0.06, 0.51)	0.367	(-0.08, 0.79)
$\beta_{2,10}$	Black	0.398	(0.03, 0.91)	0.375	(0.04, 0.88)	0.275	(0.02, 0.60)	0.111	(-0.25, 0.53)
$\beta_{2,11}$	Older than age 65	-0.062	(-0.71, 0.54)	-0.310	(-1.07, 0.38)	-0.331	(-0.88, 0.17)	0.246	(-0.19, 0.73)
r	dispersion	1.728	(0.04, 4.59)	1.900	(0.02, 5.00)	2.187	(0.07, 5.54)	0.736	(0.49, 0.97)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 12 Sensitivity analysis for spatial Negative Binomial model to the definition of cumulative counts in the Illinois study: Total-cumulative count means accumulated counts since beginning of this study; 1-Month-cumulative count means accumulated counts since 1-month before; 1-week-cumulative count means accumulated counts since 1-week before

	Variable	Total-cumulative		1 Month-cumulative		1 week-cumulative	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	1.226	(-3.47, 5.46)	-0.821	(-6.62, 4.69)	-0.685	(-8.78, 7.21)
$\beta_{2,3}$	poverty	0.115	(-0.32, 0.55)	0.218	(-0.29, 0.78)	0.165	(-0.62, 0.94)
$\beta_{2,4}$	Population density	-0.051	(-0.37, 0.24)	-0.383	(-0.89, 0.07)	-0.635	(-1.49, 0.20)
$\beta_{2,5}$	MHV	-0.366	(-1.85, 1.19)	-0.319	(-2.06, 1.51)	-0.330	(-3.22, 2.21)
$\beta_{2,6}$	MHI	-0.090	(-1.04, 0.88)	-0.275	(-1.43, 0.99)	-0.582	(-2.41, 1.42)
$\beta_{2,7}$	Home owners rate	0.652	(-0.22, 1.50)	0.152	(-0.88, 1.23)	0.557	(-1.31, 2.42)
$\beta_{2,8}$	Hispanic	-0.455	(-0.88, -0.03)	-0.550	(-1.06, -0.04)	-0.374	(-1.35, 0.50)
$\beta_{2,9}$	Education	0.367	(-0.08, 0.79)	0.527	(-0.09, 1.10)	0.273	(-0.60, 1.25)
$\beta_{2,10}$	Black	0.111	(-0.25, 0.53)	-0.010	(-0.47, 0.51)	0.160	(-0.66, 0.97)
$\beta_{2,11}$	Older than age 65	0.246	(-0.19, 0.73)	0.114	(-0.42, 0.64)	0.292	(-0.63, 1.27)
r	dispersion	0.736	(0.49, 0.97)	0.553	(0.32, 0.81)	0.301	(0.16, 0.51)

MHV : Median House Value
MHI : Median Household Income

TABLE 13 Sensitivity analysis for the STZINB(NLT) model to the definition of weekly counts in the Illinois study: Mon-Sun means accumulated counts from Monday to the next Sunday; Thu-Wed means accumulated counts from Thursday to the next Wednesday; Sun-Mon means accumulated counts from Sunday to the next Monday;

	Variable	Mon-Sun		Thu-Wed		Sun-Mon	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.247	(-0.38, 0.97)	0.071	(-0.35, 0.53)	0.123	(-0.38, 0.62)
$\beta_{2,3}$	poverty	-0.246	(-0.78, 0.22)	-0.095	(-0.61, 0.3)	-0.125	(-0.61, 0.31)
$\beta_{2,4}$	Population density	0.478	(-1.04, 2.16)	0.280	(-1, 1.92)	0.315	(-1.08, 1.91)
$\beta_{2,5}$	MHV	-0.249	(-1.12, 0.7)	-0.172	(-0.94, 0.69)	-0.124	(-0.89, 0.78)
$\beta_{2,6}$	MHI	0.603	(-0.79, 1.81)	0.657	(-0.18, 1.61)	0.487	(-0.45, 1.44)
$\beta_{2,7}$	Home owners rate	-0.368	(-0.99, 0.14)	-0.480	(-1.06, 0.02)	-0.435	(-1.01, 0.08)
$\beta_{2,8}$	hispanic	0.331	(-0.01, 0.8)	0.405	(0.04, 0.79)	0.407	(0.06, 0.8)
$\beta_{2,9}$	education	0.322	(-0.11, 0.9)	0.302	(-0.1, 0.73)	0.319	(-0.09, 0.81)
$\beta_{2,10}$	Black	0.398	(0.03, 0.91)	0.367	(0.04, 0.78)	0.327	(0.03, 0.71)
$\beta_{2,11}$	Older than age 65	-0.062	(-0.71, 0.54)	-0.069	(-0.7, 0.42)	-0.086	(-0.63, 0.45)
r	dispersion	1.728	(0.04, 4.59)	1.902	(0.09, 4.59)	1.614	(0.02, 4.13)

MHV : Median House Value
MHI : Median Household Income

[Result: Illinois ends here]

5.3 Result: Florida

TABLE 14 Parameter estimates and 95% credebile intervals for the spatio-temporal models in Florida study

Model		Variable	STZINB(NLT)		STZINB(LT)		STNB	
			Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
Binary	$\beta_{1,0}$	intercept	-1.542	(-11.36, 8.29)	-0.677	(-12.13, 12.05)		
	$\beta_{1,1}$	time	-	-	-0.138	(-15.67, 13.48)		
	$\beta_{1,2}$	$PM_{2.5}$	0.390	(-0.08, 0.89)	0.871	(-0.26, 2.60)		
	$\beta_{1,3}$	poverty	-0.416	(-0.80, -0.05)	-0.238	(-1.01, 0.69)		
	$\beta_{1,4}$	Population density	1.213	(-0.33, 2.85)	2.632	(-0.91, 7.20)		
	$\beta_{1,5}$	MHV	-0.279	(-0.95, 0.28)	0.127	(-1.14, 2.08)		
	$\beta_{1,6}$	MHI	0.231	(-0.51, 0.93)	0.022	(-1.26, 1.29)		
	$\beta_{1,7}$	Home owners rate	0.215	(-0.43, 0.89)	0.630	(-0.60, 2.05)		
	$\beta_{1,8}$	hispanic	1.067	(0.49, 1.67)	1.479	(0.39, 3.22)		
	$\beta_{1,9}$	education	-0.630	(-1.29, -0.13)	-0.433	(-1.62, 0.66)		
	$\beta_{1,10}$	Black	0.530	(0.19, 0.90)	0.781	(0.06, 1.94)		
	$\beta_{1,11}$	Older than age 65	0.571	(-0.11, 1.27)	0.598	(-0.60, 1.86)		
Count	$\beta_{2,0}$	intercept	-4.159	(-13.86, 5.57)	-4.254	(-14.12, 5.91)	-4.279	(-13.75, 5.54)
	$\beta_{2,1}$	time	-	-	-	-	-	-
	$\beta_{2,2}$	$PM_{2.5}$	0.052	(-0.16, 0.33)	0.037	(-0.18, 0.31)	0.032	(-0.14, 0.28)
	$\beta_{2,3}$	poverty	-0.054	(-0.30, 0.13)	-0.070	(-0.29, 0.11)	-0.046	(-0.24, 0.10)
	$\beta_{2,4}$	Population density	0.151	(-0.37, 0.69)	0.119	(-0.28, 0.65)	0.165	(-0.28, 0.70)
	$\beta_{2,5}$	MHV	-0.633	(-1.51, -0.02)	-0.521	(-1.40, 0.06)	-0.626	(-1.52, -0.02)
	$\beta_{2,6}$	MHI	0.543	(-0.05, 1.28)	0.448	(-0.07, 1.27)	0.525	(-0.03, 1.32)
	$\beta_{2,7}$	Home owners rate	-0.359	(-1.21, 0.23)	-0.299	(-1.16, 0.20)	-0.364	(-1.18, 0.22)
	$\beta_{2,8}$	Hispanic	0.211	(-0.03, 0.64)	0.173	(-0.07, 0.65)	0.172	(-0.05, 0.50)
	$\beta_{2,9}$	education	-0.481	(-1.14, 0.09)	-0.406	(-1.11, 0.13)	-0.453	(-1.04, 0.08)
	$\beta_{2,10}$	Black	-0.017	(-0.24, 0.15)	0.005	(-0.16, 0.15)	-0.046	(-0.25, 0.10)
	$\beta_{2,11}$	Older than age 65	0.692	(0.02, 1.77)	0.611	(-0.04, 1.69)	0.689	(0.01, 1.74)
	r	dispersion	47.416	(25.58, 80.37)	78.809	(8.78, 134.76)	54.572	(33.35, 76.82)
Random Effects	Γ_{11}		0.015	(0.0104, 0.021)	0.015	(0.0105, 0.0208)		
	Γ_{12}		0.000	(-0.004, 0.0037)	0.000	(-0.0037, 0.0038)		
	Γ_{13}		0.000	(-0.004, 0.0036)	0.000	(-0.0036, 0.0041)		
	Γ_{14}		0.000	(-0.0038, 0.0037)	0.000	(-0.0037, 0.0037)		
	Γ_{22}		0.015	(0.0106, 0.0211)	0.015	(0.0106, 0.0212)		
	Γ_{23}		0.000	(-0.0037, 0.0039)	0.000	(-0.0037, 0.0038)		
	Γ_{24}		0.000	(-0.0037, 0.0039)	0.000	(-0.0036, 0.004)		
	Γ_{33}		0.015	(0.0103, 0.0208)	0.015	(0.0107, 0.0209)	0.015	(0.0107, 0.0209)
	Γ_{34}		0.000	(-0.0039, 0.0038)	0.000	(-0.0037, 0.0041)	0.000	(-0.0037, 0.004)
	Γ_{44}		0.015	(0.01, 0.0203)	0.015	(0.0101, 0.0207)	0.015	(0.0103, 0.0209)
DIC			3907.058		3690.912		3794.056	

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 15 Comparison by both spatio-temporal models with weekly counts and spatial models with cumulative counts in Florida study

	Variable	STZINB (NLT)		STZINB (LT)		STNB		SNB	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.052	(-0.16, 0.33)	0.037	(-0.18, 0.31)	0.032	(-0.14, 0.28)	0.228	(-0.14, 0.61)
$\beta_{2,3}$	poverty	-0.054	(-0.30, 0.13)	-0.070	(-0.29, 0.11)	-0.046	(-0.24, 0.10)	-0.135	(-0.41, 0.16)
$\beta_{2,4}$	Population density	0.151	(-0.37, 0.69)	0.119	(-0.28, 0.65)	0.165	(-0.28, 0.70)	0.536	(-0.47, 1.59)
$\beta_{2,5}$	MHV	-0.633	(-1.51, -0.02)	-0.521	(-1.40, 0.06)	-0.626	(-1.52, -0.02)	-0.348	(-0.68, -0.02)
$\beta_{2,6}$	MHI	0.543	(-0.05, 1.28)	0.448	(-0.07, 1.27)	0.525	(-0.03, 1.32)	0.392	(-0.05, 0.81)
$\beta_{2,7}$	Home owners rate	-0.359	(-1.21, 0.23)	-0.299	(-1.16, 0.20)	-0.364	(-1.18, 0.22)	-0.193	(-0.59, 0.21)
$\beta_{2,8}$	Hispanic	0.211	(-0.03, 0.64)	0.173	(-0.07, 0.65)	0.172	(-0.05, 0.50)	0.440	(0.11, 0.76)
$\beta_{2,9}$	education	-0.481	(-1.14, 0.09)	-0.406	(-1.11, 0.13)	-0.453	(-1.04, 0.08)	-0.112	(-0.41, 0.18)
$\beta_{2,10}$	Black	-0.017	(-0.24, 0.15)	0.005	(-0.16, 0.15)	-0.046	(-0.25, 0.10)	0.170	(-0.06, 0.41)
$\beta_{2,11}$	Older than age 65	0.692	(0.02, 1.77)	0.611	(-0.04, 1.69)	0.689	(0.01, 1.74)	0.647	(0.34, 0.98)
r	dispersion								

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 16 Sensitivity analysis for spatial Negative Binomial model to the definition of cumulative counts in the Florida study: Total-cumulative count means accumulated counts since beginning of this study; 1-Month-cumulative count means accumulated counts since 1-month before; 1-week-cumulative count means accumulated counts since 1-week before

	Variable	Total-cumulative		1 Month-cumulative		1 week-cumulative	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.228	(-0.14, 0.61)	0.210	(-0.26, 0.66)	0.128	(-0.47, 0.73)
$\beta_{2,3}$	poverty	-0.135	(-0.41, 0.16)	0.072	(-0.33, 0.51)	0.141	(-0.41, 0.69)
$\beta_{2,4}$	Population density	0.536	(-0.47, 1.59)	0.443	(-0.62, 1.42)	0.242	(-0.82, 1.43)
$\beta_{2,5}$	MHV	-0.348	(-0.68, -0.02)	-0.426	(-0.86, 0.03)	-0.565	(-1.14, 0.07)
$\beta_{2,6}$	MHI	0.392	(-0.05, 0.81)	0.512	(-0.02, 1.05)	0.710	(0.02, 1.42)
$\beta_{2,7}$	Home owners rate	-0.193	(-0.59, 0.21)	-0.354	(-0.85, 0.14)	-0.568	(-1.20, 0.05)
$\beta_{2,8}$	Hispanic	0.440	(0.11, 0.76)	0.283	(-0.18, 0.74)	-0.102	(-0.81, 0.53)
$\beta_{2,9}$	education	-0.112	(-0.41, 0.18)	-0.066	(-0.42, 0.30)	0.125	(-0.38, 0.63)
$\beta_{2,10}$	Black	0.170	(-0.06, 0.41)	0.083	(-0.18, 0.39)	-0.013	(-0.45, 0.42)
$\beta_{2,11}$	Older than age 65	0.647	(0.34, 0.98)	0.654	(0.28, 1.04)	0.848	(0.31, 1.32)
r	dispersion	2.132	(1.50, 2.94)	1.373	(0.94, 1.9)	0.999	(0.56, 1.59)

MHV : Median House Value
MHI : Median Household Income

TABLE 17 Sensitivity analysis for the STZINB(NLT) model to the definition of weekly counts in the Florida study: Mon-Sun means accumulated counts from Monday to the next Sunday; Thu-Wed means accumulated counts from Thursday to the next Wednesday; Sun-Mon means accumulated counts from Sunday to the next Monday;

	Variable	Mon-Sun		Thu-Wed		Sun-Mon	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.052	(-0.16, 0.33)	0.000	(-0.11, 0.13)	0.03	(-0.15, 0.26)
$\beta_{2,3}$	poverty	-0.054	(-0.30, 0.13)	-0.112	(-0.37, 0.1)	-0.075	(-0.31, 0.15)
$\beta_{2,4}$	Population density	0.151	(-0.37, 0.69)	0.140	(-0.12, 0.48)	0.062	(-0.34, 0.57)
$\beta_{2,5}$	MHV	-0.633	(-1.51, -0.02)	-0.484	(-1.29, 0.13)	-0.679	(-1.63, 0.06)
$\beta_{2,6}$	MHI	0.543	(-0.05, 1.28)	0.465	(-0.24, 1.25)	0.672	(-0.14, 1.6)
$\beta_{2,7}$	Home owners rate	-0.359	(-1.21, 0.23)	-0.283	(-0.8, 0.22)	-0.45	(-1.22, 0.19)
$\beta_{2,8}$	hispanic	0.211	(-0.03, 0.64)	0.178	(-0.08, 0.6)	0.227	(-0.06, 0.73)
$\beta_{2,9}$	education	-0.481	(-1.14, 0.09)	-0.331	(-0.93, 0.15)	-0.488	(-1.27, 0.09)
$\beta_{2,10}$	Black	-0.017	(-0.24, 0.15)	-0.005	(-0.22, 0.19)	-0.038	(-0.28, 0.13)
$\beta_{2,11}$	Older than age 65	0.692	(0.02, 1.77)	0.617	(-0.07, 1.58)	0.848	(0.06, 1.88)

MHV : Median House Value

MHI : Median Household Income

[Result: Florida ends here]

5.4 Result: California

TABLE 18 Parameter estimates and 95% credebile intervals for the spatio-temporal models in California study

Model		Variable	STZINB(NLT)		STZINB(LT)		STNB	
			Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
Binary	$\beta_{1,0}$	intercept	-2.024	(-11.76, 7.25)	1.182	(-10.59, 12.8)		
	$\beta_{1,1}$	time	-	-	0.547	(-14.0, 14.96)		
	$\beta_{1,2}$	$PM_{2.5}$	0.490	(0.22, 0.83)	0.496	(0.01, 1.12)		
	$\beta_{1,3}$	poverty	-0.045	(-0.05, 0.42)	-0.105	(-1.27, 1.00)		
	$\beta_{1,4}$	Population density	-0.381	(-1.52, 0.96)	-0.390	(-2.57, 2.00)		
	$\beta_{1,5}$	MHV	0.973	(-0.34, 2.18)	0.790	(-1.63, 3.70)		
	$\beta_{1,6}$	MHI	0.317	(-0.73, 1.50)	0.619	(-1.65, 3.70)		
	$\beta_{1,7}$	Home owners rate	-0.680	(-1.98, 0.41)	-1.046	(-3.25, 0.77)		
	$\beta_{1,8}$	Hispanic	-0.308	(-2.01, 1.08)	-0.800	(-3.76, 1.88)		
	$\beta_{1,9}$	education	0.619	(-0.41, 1.71)	0.897	(-1.19, 3.47)		
	$\beta_{1,10}$	Black	0.673	(-0.01, 1.54)	0.724	(-0.41, 2.08)		
	$\beta_{1,11}$	Older than age 65	-0.133	(-1.30, 1.20)	0.544	(-1.31, 3.10)		
Count	$\beta_{2,0}$	intercept	-0.531	(-10.79, 9.43)	-1.950	(-11.84, 7.71)	-1.911	(-11.3, 7.57)
	$\beta_{2,1}$	time	-	-	-	-	-	-
	$\beta_{2,2}$	$PM_{2.5}$	0.266	(-0.06, 0.71)	0.306	(0.01, 0.71)	0.223	(0.03, 0.52)
	$\beta_{2,3}$	poverty	0.420	(-0.16, 1.22)	0.354	(-0.14, 1.02)	0.210	(-0.21, 0.67)
	$\beta_{2,4}$	Population density	0.649	(-1.34, 3.49)	0.492	(-1.06, 2.54)	0.331	(-0.37, 1.20)
	$\beta_{2,5}$	MHV	1.359	(0.35, 2.59)	1.355	(0.46, 2.52)	1.148	(0.36, 1.90)
	$\beta_{2,6}$	MHI	-0.826	(-2.39, 0.27)	-0.682	(-2.05, 0.24)	-0.651	(-1.40, -0.04)
	$\beta_{2,7}$	Home owners rate	1.836	(-0.64, 5.03)	1.557	(-0.51, 4.09)	1.152	(-0.09, 2.71)
	$\beta_{2,8}$	Hispanic	1.775	(-1.04, 5.87)	1.568	(-0.64, 4.58)	1.178	(0.15, 2.51)
	$\beta_{2,9}$	education	-0.552	(-3.25, 1.21)	-0.438	(-2.30, 1.24)	-0.328	(-1.28, 0.54)
	$\beta_{2,10}$	Black	0.660	(0.12, 1.32)	0.746	(0.16, 1.36)	0.592	(0.27, 1.05)
	$\beta_{2,11}$	Older than age 65	-0.613	(-2.35, 0.96)	-0.789	(-2.20, 0.48)	-0.779	(-1.87, 0.06)
	r	dispersion	1.630	(0.02, 4.58)	2.140	(0.03, 5.42)	3.449	(0.15, 9.02)
Random Effects	Γ_{11}		0.018	(0.0114, 0.0245)	0.018	(0.0119, 0.0247)		
	Γ_{12}		0.000	(-0.0046, 0.0046)	0.000	(-0.0045, 0.0049)		
	Γ_{13}		0.000	(-0.0048, 0.0049)	0.000	(-0.0046, 0.0047)		
	Γ_{14}		0.000	(-0.0045, 0.0051)	0.000	(-0.005, 0.0045)		
	Γ_{22}		0.018	(0.0121, 0.0249)	0.018	(0.0114, 0.0244)		
	Γ_{23}		0.000	(-0.0051, 0.0047)	0.000	(-0.0051, 0.0044)		
	Γ_{24}		0.000	(-0.0046, 0.0051)	0.000	(-0.0047, 0.005)		
	Γ_{33}		0.018	(0.012, 0.0246)	0.018	(0.012, 0.0252)	0.018	(0.0115, 0.0245)
	Γ_{34}		0.000	(-0.0049, 0.0048)	0.000	(-0.005, 0.0044)	0.000	(-0.0051, 0.0046)
	Γ_{44}		0.018	(0.0119, 0.0248)	0.018	(0.0121, 0.0251)	0.018	(0.0119, 0.0248)
DIC			10980.87		12984.49		20179.82	

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 19 Comparison by both spatio-temporal models with weekly counts and spatial models with cumulative counts in the California study

	Variable	STZINB (NLT)		STZINB (LT)		STNB		SNB	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.266	(-0.06, 0.71)	0.306	(0.01, 0.71)	0.223	(0.03, 0.52)	0.306	(0.15, 0.45)
$\beta_{2,3}$	poverty	0.420	(-0.16, 1.22)	0.354	(-0.14, 1.02)	0.210	(-0.21, 0.67)	0.156	(-0.18, 0.48)
$\beta_{2,4}$	Population density	0.649	(-1.34, 3.49)	0.492	(-1.06, 2.54)	0.331	(-0.37, 1.20)	-0.356	(-0.89, 0.20)
$\beta_{2,5}$	MHV	1.359	(0.35, 2.59)	1.355	(0.46, 2.52)	1.148	(0.36, 1.90)	1.088	(0.30, 2.03)
$\beta_{2,6}$	MHI	-0.826	(-2.39, 0.27)	-0.682	(-2.05, 0.24)	-0.651	(-1.40, -0.04)	-0.466	(-1.29, 0.30)
$\beta_{2,7}$	Home owners rate	1.836	(-0.64, 5.03)	1.557	(-0.51, 4.09)	1.152	(-0.09, 2.71)	-0.036	(-0.69, 0.53)
$\beta_{2,8}$	Hispanic	1.775	(-1.04, 5.87)	1.568	(-0.64, 4.58)	1.178	(0.15, 2.51)	0.419	(-0.29, 1.10)
$\beta_{2,9}$	education	-0.552	(-3.25, 1.21)	-0.438	(-2.30, 1.24)	-0.328	(-1.28, 0.54)	-0.013	(-0.68, 0.57)
$\beta_{2,10}$	Black	0.660	(0.12, 1.32)	0.746	(0.16, 1.36)	0.592	(0.27, 1.05)	0.485	(0.11, 0.90)
$\beta_{2,11}$	Older than age 65	-0.613	(-2.35, 0.96)	-0.789	(-2.20, 0.48)	-0.779	(-1.87, 0.06)	-0.096	(-0.58, 0.39)
r	dispersion	1.630	(0.02, 4.58)	2.140	(0.03, 5.42)	3.449	(0.15, 9.02)	1.388	(0.94, 1.96)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 20 Sensitivity analysis for spatial Negative Binomial model to the definition of cumulative counts in the California study: Total-cumulative count means accumulated counts since beginning of this study; 1-Month-cumulative count means accumulated counts since 1-month before; 1-week-cumulative count means accumulated counts since 1-week before

	Variable	Total-cumulative		1 Month-cumulative		1 week-cumulative	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.306	(0.15, 0.45)	0.354	(0.18, 0.54)	0.232	(0.04, 0.43)
$\beta_{2,3}$	poverty	0.156	(-0.18, 0.48)	0.119	(-0.29, 0.54)	0.183	(-0.26, 0.64)
$\beta_{2,4}$	Population density	-0.356	(-0.89, 0.20)	-0.402	(-1.07, 0.29)	-0.286	(-1.12, 0.43)
$\beta_{2,5}$	MHV	1.088	(0.30, 2.03)	0.946	(-0.13, 2.02)	0.774	(-0.53, 2.07)
$\beta_{2,6}$	MHI	-0.466	(-1.29, 0.30)	-0.615	(-1.68, 0.40)	-0.828	(-2.00, 0.34)
$\beta_{2,7}$	Home owners rate	-0.036	(-0.69, 0.53)	-0.197	(-0.95, 0.57)	-0.38	(-1.22, 0.50)
$\beta_{2,8}$	Hispanic	0.419	(-0.29, 1.10)	0.524	(-0.38, 1.52)	0.398	(-0.65, 1.42)
$\beta_{2,9}$	education	-0.013	(-0.68, 0.57)	-0.295	(-1.06, 0.53)	-0.513	(-1.38, 0.34)
$\beta_{2,10}$	Black	0.485	(0.11, 0.90)	0.394	(-0.10, 0.89)	0.366	(-0.13, 0.93)
$\beta_{2,11}$	Older than age 65	-0.096	(-0.58, 0.39)	0.081	(-0.59, 0.68)	-0.135	(-0.85, 0.52)
r	dispersion	1.388	(0.94, 1.96)	1.028	(0.63, 1.44)	0.978	(0.52, 1.43)

MHV : Median House Value
MHI : Median Household Income

TABLE 21 Sensitivity analysis for the STZINB(NLT) model to the definition of weekly counts in the California study: Mon-Sun means accumulated counts from Monday to the next Sunday; Thu-Wed means accumulated counts from Thursday to the next Wednesday; Sun-Mon means accumulated counts from Sunday to the next Monday;

	Variable	Mon-Sun		Thu-Wed		Sun-Mon	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.266	(-0.06, 0.71)	0.419	(-0.02, 1.12)	0.448	(-0.01, 1.18)
$\beta_{2,3}$	poverty	0.42	(-0.16, 1.22)	0.375	(-0.26, 1.24)	0.487	(-0.2, 1.34)
$\beta_{2,4}$	Population density	0.649	(-1.34, 3.49)	0.485	(-0.55, 1.82)	0.496	(-0.6, 1.81)
$\beta_{2,5}$	MHV	1.359	(0.35, 2.59)	2.253	(0.36, 4.74)	2.411	(0.36, 5.41)
$\beta_{2,6}$	MHI	-0.826	(-2.39, 0.27)	-1.437	(-3.14, -0.01)	-1.834	(-4.24, -0.14)
$\beta_{2,7}$	Home owners rate	1.836	(-0.64, 5.03)	2.204	(-0.09, 4.86)	2.189	(-0.36, 4.96)
$\beta_{2,8}$	Hispanic	1.775	(-1.04, 5.87)	1.737	(-0.39, 4.39)	1.667	(-0.24, 3.92)
$\beta_{2,9}$	education	-0.552	(-3.25, 1.21)	-0.672	(-2.18, 0.61)	-0.937	(-2.63, 0.62)
$\beta_{2,10}$	Black	0.66	(0.12, 1.32)	0.864	(0.14, 1.75)	1.013	(0.17, 2.09)
$\beta_{2,11}$	Older than age 65	-0.613	(-2.35, 0.96)	-0.976	(-3.56, 0.87)	-1.092	(-3.26, 0.67)

MHV : Median House Value

MHI : Median Household Income

[Result: California ends here]

5.5 | Result: Georgia

TABLE 22 Parameter estimates and 95% credebile intervals for the spatio-temporal models in Georgia study

Model		Variable	STZINB(NLT)		STZINB(LT)		STNB	
			Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
Binary	$\beta_{1,0}$	intercept	-4.133	(-14.8, 6.43)	0.023	(-13.6, 13.96)		
	$\beta_{1,1}$	time	-	-	-0.091	(-15.46, 14.73)		
	$\beta_{1,2}$	$PM_{2.5}$	0.572	(0.12, 1.20)	0.502	(-0.36, 1.35)		
	$\beta_{1,3}$	poverty	0.113	(-0.13, 0.48)	0.332	(-0.23, 1.19)		
	$\beta_{1,4}$	Population density	-0.018	(-0.55, 0.58)	0.26	(-1.03, 1.86)		
	$\beta_{1,5}$	MHV	0.235	(-0.25, 0.70)	0.197	(-0.71, 1.02)		
	$\beta_{1,6}$	MHI	-0.020	(-0.41, 0.44)	0.024	(-0.87, 0.99)		
	$\beta_{1,7}$	Home owners rate	0.096	(-0.26, 0.64)	0.182	(-0.6, 1.3)		
	$\beta_{1,8}$	Hispanic	0.026	(-0.26, 0.26)	-0.072	(-0.61, 0.5)		
	$\beta_{1,9}$	education	-0.095	(-0.34, 0.2)	-0.106	(-0.64, 0.45)		
	$\beta_{1,10}$	Black	0.090	(-0.20, 0.39)	-0.184	(-1.08, 0.43)		
	$\beta_{1,11}$	Older than age 65	-0.084	(-0.36, 0.2)	-0.051	(-0.65, 0.59)		
Count	$\beta_{2,0}$	intercept	-3.515	(-13.7, 6.85)	-4.508	(-14.97, 6.02)	-3.504	(-13.87, 6.5)
	$\beta_{2,1}$	time	-	-	-	-	-	-
	$\beta_{2,2}$	$PM_{2.5}$	0.319	(0.00, 0.76)	0.436	(0.06, 0.98)	0.348	(0.06, 0.77)
	$\beta_{2,3}$	poverty	-0.164	(-0.38, 0.01)	-0.172	(-0.4, 0)	-0.151	(-0.38, 0.01)
	$\beta_{2,4}$	Population density	-0.185	(-0.48, 0.06)	-0.192	(-0.55, 0.06)	-0.192	(-0.49, 0.07)
	$\beta_{2,5}$	MHV	0.160	(-0.02, 0.38)	0.202	(0.01, 0.44)	0.189	(0.02, 0.39)
	$\beta_{2,6}$	MHI	-0.024	(-0.29, 0.28)	-0.027	(-0.28, 0.26)	-0.075	(-0.32, 0.2)
	$\beta_{2,7}$	Home owners rate	-0.336	(-0.67, -0.07)	-0.316	(-0.66, -0.06)	-0.301	(-0.61, -0.07)
	$\beta_{2,8}$	Hispanic	0.104	(-0.05, 0.33)	0.117	(-0.03, 0.34)	0.101	(-0.05, 0.32)
	$\beta_{2,9}$	education	-0.076	(-0.27, 0.09)	-0.083	(-0.28, 0.09)	-0.076	(-0.26, 0.08)
	$\beta_{2,10}$	Black	0.202	(0.01, 0.46)	0.245	(0.04, 0.53)	0.214	(0.02, 0.48)
	$\beta_{2,11}$	Older than age 65	0.074	(-0.08, 0.22)	0.053	(-0.1, 0.21)	0.063	(-0.08, 0.22)
	r	dispersion	3.213	(0.02, 7.8)	2.490	(0.02, 6.38)	2.688	(0.05, 7.07)
Random Effects	Γ_{11}		0.006	(0.005, 0.008)	0.006	(0.005, 0.008)		
	Γ_{12}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)		
	Γ_{13}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)		
	Γ_{14}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)		
	Γ_{22}		0.006	(0.005, 0.008)	0.006	(0.005, 0.008)		
	Γ_{23}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)		
	Γ_{24}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)		
	Γ_{33}		0.006	(0.005, 0.008)	0.006	(0.005, 0.008)	0.006	(0.005, 0.008)
	Γ_{34}		0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)	0.000	(-0.001, 0.001)
	Γ_{44}		0.006	(0.005, 0.008)	0.006	(0.005, 0.008)	0.006	(0.005, 0.008)
DIC			3206.319		3957.589		5242.302	

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 23 Comparison by both spatio-temporal models with weekly counts and spatial models with cumulative counts in the Georgia study

	Variable	STZINB (NLT)		STZINB (LT)		STNB		SNB	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.319	(0.00, 0.76)	0.436	(0.06, 0.98)	0.348	(0.06, 0.77)	0.330	(0.08, 0.56)
$\beta_{2,3}$	poverty	-0.164	(-0.38, 0.01)	-0.172	(-0.4, 0.0)	-0.151	(-0.38, 0.01)	-0.077	(-0.26, 0.1)
$\beta_{2,4}$	Population density	-0.185	(-0.48, 0.06)	-0.192	(-0.55, 0.06)	-0.192	(-0.49, 0.07)	-0.049	(-0.43, 0.36)
$\beta_{2,5}$	MHV	0.160	(-0.02, 0.38)	0.202	(0.01, 0.44)	0.189	(0.02, 0.39)	0.169	(-0.04, 0.39)
$\beta_{2,6}$	MHI	-0.024	(-0.29, 0.28)	-0.027	(-0.28, 0.26)	-0.075	(-0.32, 0.2)	-0.179	(-0.46, 0.12)
$\beta_{2,7}$	Home owners rate	-0.336	(-0.67, -0.07)	-0.316	(-0.66, -0.06)	-0.301	(-0.61, -0.07)	-0.120	(-0.33, 0.09)
$\beta_{2,8}$	hispanic	0.104	(-0.05, 0.33)	0.117	(-0.03, 0.34)	0.101	(-0.05, 0.32)	0.054	(-0.11, 0.21)
$\beta_{2,9}$	education	-0.076	(-0.27, 0.09)	-0.083	(-0.28, 0.09)	-0.076	(-0.26, 0.08)	-0.095	(-0.3, 0.11)
$\beta_{2,10}$	Black	0.202	(0.01, 0.46)	0.245	(0.04, 0.53)	0.214	(0.02, 0.48)	0.120	(-0.06, 0.32)
$\beta_{2,11}$	Older than age 65	0.074	(-0.08, 0.22)	0.053	(-0.1, 0.21)	0.063	(-0.08, 0.22)	-0.028	(-0.25, 0.19)
r	dispersion	3.213	(0.02, 7.8)	2.490	(0.02, 6.38)	2.688	(0.05, 7.07)	1.566	(1.22, 1.85)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 24 Sensitivity analysis for spatial Negative Binomial model to the definition of cumulative counts in the Georgia study: Total-cumulative count means accumulated counts since beginning of this study; 1-Month-cumulative count means accumulated counts since 1-month before; 1-week-cumulative count means accumulated counts since 1-week before

	Variable	Total-cumulative		1 Month-cumulative		1 week-cumulative	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.330	(0.08, 0.56)	0.235	(-0.03, 0.48)	0.074	(-0.24, 0.4)
$\beta_{2,3}$	poverty	-0.077	(-0.26, 0.1)	0.092	(-0.12, 0.27)	0.078	(-0.2, 0.37)
$\beta_{2,4}$	Population density	-0.049	(-0.43, 0.36)	0.149	(-0.25, 0.52)	0.298	(-0.21, 0.8)
$\beta_{2,5}$	MHV	0.169	(-0.04, 0.39)	0.176	(-0.01, 0.36)	0.026	(-0.26, 0.32)
$\beta_{2,6}$	MHI	-0.179	(-0.46, 0.12)	-0.183	(-0.46, 0.13)	-0.088	(-0.5, 0.33)
$\beta_{2,7}$	Home owners rate	-0.120	(-0.33, 0.09)	-0.117	(-0.36, 0.12)	0.009	(-0.31, 0.32)
$\beta_{2,8}$	hispanic	0.054	(-0.11, 0.21)	-0.047	(-0.21, 0.13)	-0.040	(-0.27, 0.18)
$\beta_{2,9}$	education	-0.095	(-0.3, 0.11)	-0.161	(-0.37, 0.06)	-0.191	(-0.47, 0.11)
$\beta_{2,10}$	Black	0.120	(-0.06, 0.32)	-0.195	(-0.41, 0.03)	-0.125	(-0.41, 0.16)
$\beta_{2,11}$	Older than age 65	-0.028	(-0.25, 0.19)	-0.098	(-0.34, 0.13)	-0.109	(-0.40, 0.20)
r	dispersion	1.566	(1.22, 1.85)	1.484	(1.15, 1.88)	1.070	(0.80, 1.46)

STNB : Spatio-temporal Negative Binomial (Section 3.1.2)
STZINB(NLT) : Spatio-temporal Zero-Inflated Negative Binomial (Non-Linear Time version) (Section 3.1.3)
STZINB(LT) : Spatio-temporal Zero-Inflated Negative Binomial (Linear Time version) (Section 3.1.3)
MHV : Median House Value
MHI : Median Household Income

TABLE 25 Sensitivity analysis for the STZINB(NLT) model to the definition of weekly counts in the Georgia study: Mon-Sun means accumulated counts from Monday to the next Sunday; Thu-Wed means accumulated counts from Thursday to the next Wednesday; Sun-Mon means accumulated counts from Sunday to the next Monday

	Variable	Mon-Sun		Thu-Wed		Sun-Mon	
		Posterior Mean	95% HDI	Posterior Mean	95% HDI	Posterior Mean	95% HDI
$\beta_{2,2}$	$PM_{2.5}$	0.590	(0.11, 1.17)	0.628	(0.11, 1.19)	0.59	(0.09, 1.21)
$\beta_{2,3}$	poverty	0.118	(-0.15, 0.47)	0.014	(-0.27, 0.33)	0.155	(-0.15, 0.56)
$\beta_{2,4}$	Population density	-0.041	(-0.52, 0.55)	0.014	(-0.6, 0.73)	0.048	(-0.55, 0.69)
$\beta_{2,5}$	MHV	0.248	(-0.22, 0.7)	0.335	(-0.11, 0.9)	0.218	(-0.28, 0.74)
$\beta_{2,6}$	MHI	0.054	(-0.35, 0.46)	-0.161	(-0.66, 0.22)	0.022	(-0.39, 0.46)
$\beta_{2,7}$	Home owners rate	0.019	(-0.37, 0.48)	0.148	(-0.25, 0.7)	0.113	(-0.26, 0.77)
$\beta_{2,8}$	hispanic	0.022	(-0.25, 0.27)	0.094	(-0.14, 0.37)	0.083	(-0.24, 0.36)
$\beta_{2,9}$	education	-0.112	(-0.38, 0.13)	-0.087	(-0.37, 0.19)	-0.085	(-0.36, 0.26)
$\beta_{2,10}$	Black	0.069	(-0.24, 0.32)	0.164	(-0.1, 0.44)	0.085	(-0.27, 0.43)
$\beta_{2,11}$	Older than age 65	-0.067	(-0.38, 0.19)	-0.094	(-0.42, 0.24)	-0.001	(-0.37, 0.35)

MHV : Median House Value

MHI : Median Household Income