Journal of Statistical Planning and Inference 229 (2024) 106098

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A framework of zero-inflated bayesian negative binomial regression models for spatiotemporal data

Qing He, Hsin-Hsiung Huang *

Department of Statistics and Data Science, University of Central Florida, Orlando, 32816, FL, USA

ARTICLE INFO

Article history:
Received 29 July 2022
Received in revised form 20 June 2023
Accepted 9 August 2023
Available online 22 August 2023

MSC:
62F15
62C10
91B72

Keywords:
Nearest-neighbor Gaussian process
Negative binomial model
Pólya-gamma scheme
Social vulnerability
Spatiotemporal effects

ABSTRACT

Spatiotemporal data analysis with massive zeros is widely used in many areas such as epidemiology and public health. We use a Bayesian framework to fit zero-inflated negative binomial models and employ a set of latent variables from Pólya-Gamma distributions to derive an efficient Gibbs sampler. The proposed model accommodates varying spatial and temporal random effects through Gaussian process priors, which have both the simplicity and flexibility in modeling nonlinear relationships through a covariance function. To conquer the computation bottleneck that GPs may suffer when the sample size is large, we adopt the nearest-neighbor GP approach that approximates the covariance matrix using local experts. For the simulation study, we adopt multiple settings with varying sizes of spatial locations to evaluate the performance of the proposed model such as spatial and temporal random effects estimation and compare the result to other methods. We also apply the proposed model to the COVID-19 death counts in the state of Florida, USA from 3/25/2020 through 7/29/2020 to examine relationships between social vulnerability and COVID-19 deaths.

1. Introduction

Zero-inflated models have been widely used for handling count data with excessive zeros (Lewsey and Thomson, 2004; Cheung, 2002). By construction, zero-inflated models assume that the zeros come from one of the cases (Greene, 1994): (1) structural zeros corresponding to individuals who are not at risk for an event, and therefore have no opportunity for a positive count and (2) random zeros which correspond to a latent class of individuals who are at risk for an event but nevertheless have an observed response of zero. A zero-inflated negative binomial (ZINB) model is a popular choice for modeling zero-inflated data as it gives more reliable parameter estimates when the nonzero counts are over-dispersed compared to other models like the zero-inflated Poisson model (Yau et al., 2003). Bayesian approaches to fitting zero-inflated models have gained attention recently (Ghosh et al., 2006; Neelon et al., 2010; Zuur et al., 2012). There is an abundance of advantages of Bayesian zero-inflated models including using prior beliefs with Bayesian inference instead of deriving asymptotically approximate distributions for estimation of parameters (Gelman et al., 1995). Bayesian approaches produce tractable inference for ZINB models and have been implemented through software such as R (R Core Team, 2021), Stan (Stan Development Team, 2021), and WinBUGS (Lunn et al., 2013).

It is important to account for spatial and temporal structures in areas such as epidemiology and public health. Within the Bayesian framework, it is natural to build a Bayesian hierarchical model with a prior distribution for spatial and temporal effects (Kang and Cressie, 2011). Distance-based exponential or Matérn covariance functions are commonly

* Corresponding author.
E-mail address: hsin.huang@ucf.edu (H.-H. Huang).

https://doi.org/10.1016/j.jspi.2023.106098
0378-3758/

adopted to specify spatial correlations in geostatistical data, while conditionally autoregressive models are often used for temporal effects (Cressie and Wikle, 2015; Banerjee et al., 2003; Zhang et al., 2022) through a hierarchical modeling framework (Cressie and Wikle, 2015). Neelon et al. (2022) proposed a cubic B-spline model to fit both temporal (overall time trend) and spatial (county-specific trend) random effects in a Bayesian framework, which depends on the number and locations of the interior knots and specification of the degree of basis functions.

There is limited research with the focus on the Bayesian ZINB model with the consideration of flexible spatial and temporal effects (Wang et al., 2016; Gu et al., 2020). Neelon (2019) proposed a Bayesian ZINB model with spatial and temporal random effects but assumed that temporal effects changed in a fixed increment, which may not reflect the randomness nature of temporal effects. In this study, we focus on proposing a flexible approach to model spatial and temporal effects in a Bayesian ZINB framework. A promising option is using a Gaussian process (GP) prior which characterize spatial or temporal effects through a kernel function and enables tractable nonparametric Bayesian inference. A GP prior is a probability distribution over infinite number of possible functions which make it a powerful tool to model nonlinear patterns. GP accounts for quantification of uncertainty, which relies on the predictive conditional distribution of a new spatial location given a set of observed locations (Cousin et al., 2016). GP can be viewed as a spline in a reproducing kernel Hilbert space with the reproducing kernel of a covariance function (Kimeldorf and Wahba, 1970; Wahba, 1990). It has become a popular modeling tool in multivariate and geostatistical settings. For example, Diana et al. (2021) proposed a Bayesian hierarchical occupancy model using GPs within a logistic regression framework for spatial and temporal random effects. However, the covariance matrix used in GP models may hinder its implementation for large datasets due to computational issues. Recent research has developed various approximation techniques to overcome the computational limitations (Smola and Bartlett, 2001; Quinonero-Candela and Rasmussen, 2005; Snelson and Ghahramani, 2007; Lázaro-Gredilla et al., 2010). There are two broad classes of these techniques, i.e., global approximation methods and local approximation methods (Liu et al., 2020). For global approximation methods, sparsity of the covariance matrix is achieved through global distillation including using a subset of the training data, exploiting the sparse structure of the matrix, and low-rank models (Chalupka et al., 2013; Quinonero-Candela and Rasmussen, 2005; Titsias, 2009). For example, low-rank models embed the original process into a lower-dimensional subspace. However, it usually requires a large subspace when used for large spatial datasets. Local approximation methods are based on multiple sets of local experts to improve the scalability based on the concept of divide-and-conquer (Liu et al., 2020; Gramacy, 2016). This class of approximations has the capability of capturing nonstationary features from the localized experts. Depending on the partition of the input space, there are inductive local experts which employs static partitioning (Vasudevan et al., 2009) and transductive local experts which employ a dynamic partition (Datta et al., 2016a,b). As an example of transductive local experts, Datta et al. (2016a) proposed a nearest-neighbor Gaussian process (NNGP) that uses conditional independence given information from neighboring points for large geostatistical datasets. It has well-defined sparse precision matrices for its finite-dimensional realization and provides fully process-based inference on underlying spatial processes.

In this paper, we adopt the NNGP model (Datta et al., 2016a) to model the spatial and temporal effects because of its ease of use and well-defined sparse precision matrix. The proposed Bayesian zero-inflated negative binomial (ZINB) model uses nearest-neighbor Gaussian process priors to fit spatial and temporal random effects that account for the over-dispersed and zero-inflated count response by incorporating the spatial and temporal covariance matrices. It allows us to obtain more reliable and stable inferences since NNGP yields a lower-dimensional spatial and temporal covariance matrices and parameter space, therefore requires much smaller matrix inversions. The propose framework is a scalable hierarchical ZINB model with flexible spatial and temporal effects through GP. First, the hierarchical model is based on work in Neelon (2019) to estimate regression parameters and the spatial and temporal effects. A key component to the parameter estimator is to introduce a set of latent variables through the Pólya-Gamma (PG) data augmentation, an underlying latent variable methodology (Pillow and Scott, 2012; Polson et al., 2013). With the PG scheme, posterior inference is highly efficient via Gibbs sampling by updating from conditional distributions with conjugate priors and explicit conditional normal distributions. To illustrate Gibbs sampling for generating a sequence of samples, we use an example of a pair of random variables $(\mathcal{X}, \mathcal{Y})$. A sequence of random variables

$$\mathcal{X}'_0, \mathcal{Y}'_0, \mathcal{X}'_1, \mathcal{Y}'_1, \ldots, \mathcal{X}'_t, \mathcal{Y}'_t, \ldots,$$

is generated iteratively as follows: 1. set $\mathcal{X}'_0$ at some initial value; 2. sample the rest from $\mathcal{Y}'_j \sim f(y \mid \mathcal{X}'_j = x'_j)$ and $\mathcal{X}'_{j+1} \sim f(x \mid \mathcal{Y}'_j = y'_j)$ alternately (Casella and George, 1992). It has been shown that the PG method yields superior performance compared to other Bayesian methods in the context of structural equation models with logistic regression for binary variables (Kim et al., 2018) and the Bayesian ZINB (Neelon, 2019). Therefore we propose a more flexible structure for the spatial and temporal random effects through the GP prior in order to achieve fully model-based Bayesian inference for the spatial and temporal effects. More specifically, to facilitate the computation process, we incorporate the framework with a conjugate latent NNGP model (Datta et al., 2016a,b; Zhang et al., 2021a), which is scalable for massive spatial datasets on modest computing environments (Datta et al., 2016a,b; Zhang et al., 2021a).

2. Bayesian inference for ZINB model using NNPG

2.1. The Pólya-Gamma scheme

Let y_i as the outcome of the ith observation, which may be a count $y_i = 0, 1, 2, \ldots$ or a binary indicator $y_i = 0, 1$. We aim to model Y_i and a set of covariates $X_i = [1, x_{i1}, \ldots, x_{iP}]^T$ through $\mathrm{E}[Y_i \mid X_i] = g^{-1}(X_i\beta)$, where $\beta = [\beta_0, \ldots, \beta_P]^T$

are the regression coefficients and g is the canonical link function in generalized linear models (GLMs) (McCullagh and Nelder, 2019). We utilize the efficient PG scheme for estimating the regression coefficients in GLMs. An advantage of the PG scheme is that the posterior distributions of the parameters of interest enable us to make use of Gibbs sampling approach within logistic regression and NB regression frameworks (Polson et al., 2013) through an augmented PG distributed random variable ω. Generally, assuming a prior distribution for regression coefficient $\beta \sim \mathcal{N}(b, \Sigma_0)$, the full conditional distributions induced by PG for the Gibbs sampler are:

$$\omega_i \mid \beta \sim \mathcal{PG}(n_i, X_i\beta), \text{ and } \beta \mid \Omega, z \sim \mathcal{N}(\mu, \Sigma),$$

where

$$\Sigma = (\Sigma_0^{-1} + X^T\Omega X)^{-1}, \text{ and } \mu = \Sigma\left(\Sigma_0^{-1}b + X^T\Omega z\right).$$

Here $\mathcal{PG}$ denotes a PG distribution, $\Omega = \text{diag}(\omega_1, \ldots, \omega_n)$ and z is the latent variable with $z_i = \frac{y_i - n_i/2}{\omega_i}$. Specifically, we use the PG scheme for estimating the regression coefficients in the Bernoulli and NB parts of the ZINB regression model, respectively.

In the binomial logistic regression model (logit model), let $n_i = 1\ \forall i$ and $y_i \sim \text{Binom}(1, \frac{1}{1+e^{-\eta_i}})$ where $\eta_i = x_i^T\beta$. Following Polson et al. (2013), the full conditional for β in a binomial regression is

$$p(\beta \mid y, r, \omega) \propto \pi(\beta)\exp\left[-\frac{1}{2}(z - X\beta)^T\Omega(z - X\beta)\right],$$

where $\pi(\beta)$ is the prior distribution, z is an $n \times 1$ vector with $z_i = \frac{y_i - 1/2}{\omega_i}$ and $\Omega = \text{diag}(\omega_i)$. It is readily seen that $z \mid \beta, \Omega \sim \mathcal{N}(X\beta, \Omega)$, which leads to the convenient Gibbs sampler for β. In the negative binomial regression model (Pillow and Scott, 2012), $n_i = y_i + r$ and r can be obtained using either the Metropolis–Hastings algorithm in which a uniform prior is applied with positive candidate values of r are drawn from a zero-truncated normal proposal centered at the current value of r, or the two-stage Gibbs sampling (Zhou and Carin, 2013; Dadaneh et al., 2018). Consider the following model for a count response Y_i,

$$p(Y_i = y_i \mid r, \beta) \overset{d}{=} \frac{\Gamma(y_i + r)}{\Gamma(r)y_i!}(1 - \psi_i)^r(\psi_i)^{y_i},\ r > 0, \quad \text{where}$$
$$\psi_i = \frac{\exp(x_i^T\beta)}{1 + \exp(x_i^T\beta)} = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)},$$

where the NB probability parameter ψ_i is parameterized using the expit (inverse-logit) function, which allows us to apply the same properties of the Pólya-Gamma density as in the logistic case.

The mean and variance of Y_i are

$$\text{E}(Y_i \mid r, \beta) = \frac{\psi_i}{1 - \psi_i} = r\exp(\eta_i) = \mu_i,$$
$$\text{Var}(Y_i \mid r, \beta) = \frac{r\psi_i}{(1 - \psi_i)^2} = r\exp(\eta_i)[1 + \exp(\eta_i)] = \mu_i(1 + \mu_i/r).$$

The parameter r accounts for overdispersion in the data. As $r \to 0$, the counts become increasingly dispersed and the NB distribution converges to the Poisson distribution. The above parameterization also leads to a Gaussian linear model where $z \mid \beta, \Omega \sim \mathcal{N}(X\beta, \Omega)$ where z is an $n \times 1$ vector with $z_i = \frac{y_i - r}{2\omega_i}$, $\omega_i \sim \mathcal{PG}(y_i + r, X_i\beta)$ and $\Omega = \text{diag}(\omega_i)$.

3. Zero-inflated negative binomial model

3.1. The proposed spatiotemporal ZINB-NNGP model

We are interested in a spatiotemporal random intercept model on data with excessive zeros. Assume observations are collected at a fixed collection of distinct locations denoted by $\mathbf{S} = \{1, \ldots, S\}$ and across T distinct time segments (e.g., weeks, months, seasons and years) denoted by $\mathbf{T} = \{1, \ldots, T\}$. We index the observations by $j = 1, \ldots, N$, where N is the total sample size. For a sampling unit $\{s, t\}$ where $s \in \mathbf{S}$ and $t \in \mathbf{T}$, we denote its number of observations by $n_{s,t}$ ($n_{s,t} \geq 0$). In a special case, $n_{s,t} = 1, \forall s \in \mathbf{S}, t \in \mathbf{T}$, that is there is exactly one observation for each sampling unit $\{s, t\}$ and $N = S \times T$. In the proposed framework, it does not require the sampling units to have equal sample sizes or at least one observation. Sparsity and unequal sizes are considered in the simulation study.

Let response Y_j be the observed count of the jth observation, ϕ_j be the probability of belonging to the at-risk group, and ψ_j be the success rate for observations in the at-risk group. The Bayesian ZINB model (Neelon, 2019) with a latent at-risk indicator variable W_j is introduced as follows:

$$Y_j \sim (1 - \phi_j)1_{(W_j=0 \wedge Y_j=0)} + \phi_j\mathcal{NB}(\mu_j, r)1_{(W_j=1)},\ j = 1, \ldots, N,$$

where μ_j is the mean of the NB distribution and r is the overdispersion parameter. A value of 1 for W_j means that the jth observation belongs to the at-risk group, and a value of 0 means the "not at-risk" group. Let s_j and t_j be its spatial

Table 1
The description of the notations used in the proposed model.

Notation	Description
S	The number of spatial locations
T	The number of temporal points
$\mathbf{S} = \{1, 2, \ldots, S\}$	The list of spatial locations
$\mathbf{T} = \{1, 2, \ldots, T\}$	The list of temporal points
N	The number of observations
$Y_j, j = 1, \ldots, N$	The response of the jth observation
$s_j \in \mathbf{S}$	The spatial location of the jth observation
$t_j \in \mathbf{T}$	The temporal point of the jth observation
ϕ_j	The probability of belonging to the at-risk group
ψ_j	The success rate for observations in the at-risk group
W_j	The indicator for belonging to the at-risk group
μ_j	The mean of negative binomial distribution
r	The dispersion parameter of negative binomial distribution
a_{s_j}	The spatial random effect in the logit model
ϵ_{11s_j}	The spatial random noise in the logit model
b_{t_j}	The temporal random effect in the logit model
ϵ_{12t_j}	The temporal random noise in the logit model
c_{s_j}	The spatial random effect in the negative binomial model
ϵ_{21s_j}	The spatial random noise in the negative binomial model
d_{t_j}	The temporal random effect in the negative binomial model
ϵ_{22t_j}	The temporal random noise in the negative binomial model

location and time point, respectively. We model the at-risk probability ϕ_j using a logit model and $Y_j \mid W_j = 1$ using a negative binomial regression model:

$$\begin{aligned} Pr(Y_j = 0) &= (1 - \phi_j) + \phi_j(1 - \psi_j)^r, \\ \text{logit}(\phi_j) &= \text{logit}[Pr(W_j = 1 \mid \alpha, a_{s_j}, b_{t_j}, \epsilon_{11s_j}, \epsilon_{12t_j})] \\ &= x_j^T\alpha + a_{s_j} + b_{t_j} + \epsilon_{11s_j} + \epsilon_{11t_j} \\ &= \eta_{1j}, \end{aligned} \tag{1}$$

$$p(y_j \mid r, \beta, c_{s_j}, d_{t_j}, \epsilon_{21s_j}, \epsilon_{22t_j}, W_j = 1) \overset{d}{=} \frac{\Gamma(y_j + r)}{\Gamma(r)y_j!}(1 - \psi_j)^r\psi_j^{y_j}, \ \forall j \ \text{ s.t. } W_j = 1,$$

$$\begin{aligned} \psi_j &= \frac{\exp(x_j^T\beta + c_{s_j} + d_{t_j} + \epsilon_{21s_j} + \epsilon_{22t_j})}{1 + \exp(x_j^T\beta + c_{s_j} + d_{t_j} + \epsilon_{21s_j} + \epsilon_{22t_j})} \\ &= \frac{\exp(\eta_{2j})}{1 + \exp(\eta_{2j})}, \end{aligned} \tag{2}$$

where a_{s_j} and ϵ_{11s_j} are the spatial random effects and spatial random noise in the logit model, b_{t_j} and ϵ_{12t_j} are the temporal random effects and temporal random noise in the logit model, c_{s_j} and ϵ_{21s_j} are the spatial random effects and spatial random noise in the negative binomial regression model, and d_t and ϵ_{22t_j} are the temporal random effects and temporal random noise in the negative binomial model. The two models are often referred as the binary component (the logit model) and count component (the negative binomial model), which we will use extensively in the paper. The proposed model incorporates additive spatial and temporal random effects as intercepts in the binary and count component. Compared with the existing Bayesian ZINB model (Neelon, 2019), the proposed model is more flexible through the use of GP priors on the spatial and temporal random effects. To elaborate the GP priors, first we define the spatial random effects corresponding to the sampling locations $\mathbf{S} = \{1, \ldots, S\}$ as

$\vec{a} = (a_1, \ldots, a_S)^T$, in the binary component and

$\vec{c} = (c_1, \ldots, c_S)^T$, in the count component.

Similarly, we define the temporal random effects of temporal points $\mathbf{T} = \{1, \ldots, T\}$ as

$\vec{b} = (b_1, \ldots, b_T)^T$, in the binary component and

$\vec{d} = (d_1, \ldots, d_T)^T$, in the count component.

A list of descriptions of the notations is given in Table 1.

To account for correlations between spatial locations or temporal points, we employ GPs for both the spatial and temporal effects. Specifically, we assume that $\vec{a} = (a_1, \ldots, a_S)^T$ in the binary component is distributed according to a GP with parameters (σ_{11}, l_{11}). Let h_s be the specific location information of site s (e.g., latitude and longitude) which can be used to compute distance between different locations. It corresponds to assuming that $(a_1, \ldots, a_S)^T \sim$

$\mathcal{N}(0, K_{\sigma_{11},l_{11}}(h_1,\ldots,h_S))$, where $K_{\sigma_{11},l_{11}}(h_1,\ldots,h_S)_{i,j} = \sigma_{11}^2 e^{-\frac{|h_i-h_j|^2}{l_{11}^2}}$, $i,j = 1,\ldots,S$. Parameter σ_{11} determines the overall variability of the GP, while parameter l_{11} measures the correlation between location i and j. Similarly, the autocorrelated temporal random effect $\vec{b} = (b_1,\ldots,b_T)^T$ in the count component is distributed according to a GP with parameters (σ_{12}, l_{12}). Let w_t be the index of temporal point t. For example, if the data are at monthly level spanning from January to December, then the corresponding indices could be $(1,\ldots,12)$. The GP prior can be written as $(b_1,\ldots,b_T)^T \sim \mathcal{N}(0, K_{\sigma_{12},l_{12}}(w_1,\ldots,w_T))$.

Let V_1 and V_2 be the $N \times S$ and $N \times T$ design matrices for the spatial and temporal random effects for the observed data. The following hierarchical structure completes the definition of the binary component with spatial random effect $\vec{a} = (a_1,\ldots,a_S)^T$, temporal random effect $\vec{b} = (b_1,\ldots,b_T)^T$, spatial random noise $\vec{\epsilon}_{11} = (\epsilon_{111},\ldots,\epsilon_{11S})^T$, and temporal random noise $\vec{\epsilon}_{12} = (\epsilon_{121},\ldots,\epsilon_{12T})^T$ including the prior distributions of relevant parameters. The hierarchical structure for the negative binomial regression with random effects $\vec{c} = (c_1,\ldots,c_S)^T$, $\vec{d} = (d_1,\ldots,d_T)^T$, $\vec{\epsilon}_{21} = (\epsilon_{211},\ldots,\epsilon_{21S})^T$, and $\vec{\epsilon}_{22} = (\epsilon_{221},\ldots,\epsilon_{22T})^T$ are in the same fashion.

$$
\begin{aligned}
(a_1,\ldots,a_S)^T &\sim \mathcal{N}(0, K_{\sigma_{11},l_{11}}(h_1,\ldots,h_S)),\\
\sigma_{11}^2 &\sim \mathcal{IG}(a_{\sigma_1}, b_{\sigma_1}),\ l_{11} \sim Gamma(a_{l_1}, b_{l_1}),\\
(b_1,\ldots,b_T)^T &\sim \mathcal{N}(0, K_{\sigma_{12},l_{12}}(w_1,\ldots,w_T)),\\
\sigma_{12}^2 &\sim \mathcal{IG}(a_{\sigma_2}, b_{\sigma_2}),\ l_{12} \sim Gamma(a_{l_2}, b_{l_2}),\\
\epsilon_{11s} &\overset{iid}{\sim} \mathcal{N}(0, \sigma_{\epsilon_{11}}^2),\ \sigma_{\epsilon_{11}}^2 \sim \mathcal{IG}(a_\epsilon, b_\epsilon), s = 1,\ldots,S,\\
\epsilon_{12t} &\overset{iid}{\sim} \mathcal{N}(0, \sigma_{\epsilon_{12}}^2),\ \sigma_{\epsilon_{12}}^2 \sim \mathcal{IG}(a_\epsilon, b_\epsilon), t = 1,\ldots,T.
\end{aligned}
$$

3.1.1. Latent nearest-neighbor Gaussian process

In a spatial regression model at location $s \in \mathbf{S}$ and time $t \in \mathbf{T}$ in a spatiotemporal domain $\mathcal{D}$, assume that

$$y(s,t) = \mu + w_1(s) + w_2(t) + \epsilon_1(s) + \epsilon_2(t),\ (s,t) \in \mathcal{D},$$

where μ is the mean function that is itself modeled as a linear combination of known covariates, $w_1(s)$ follows a latent GP with mean zero and a positive-definite cross-covariance matrix $C_\phi(s,s')$, and $\epsilon_1(s)$ follows a normal distribution with mean zero and variance σ^2. In the same fashion, we can define the temporal random effect $w_2(t)$ and the temporal random noise $\epsilon_2(t)$. This type of regression model is called a latent GP model since the GP is used as a prior in the latent process $w(s)$ and $w(t)$, while a response GP model imposes a GP on the outcome $y(s,t)$. In the proposed Bayesian spatiotemporal ZINB model (2), we use the latent GP model with both spatial and temporal random effects.

In NNGP (Datta et al., 2016b), the joint density of $(w_1,\ldots,w_S)$ was approximated by a series of conditional densities of size at most m, where $m \leq S$ and S is the size of distinct spatial locations. It was shown that the approximation was essentially a multivariate Gaussian density with covariance matrix $\tilde{C}_\phi(\mathbf{S},\mathbf{S})$. Using the spatial random effects $w_1(\mathbf{S})$ described above as an example, by selecting a list of nearest neighbors with size m for each location in $\mathbf{S}$, the Gaussian density $p(w_1(\mathbf{S}))$ can be approximated as follows (Datta et al., 2016b; Zhang et al., 2021a):

$$p(w_1(\mathbf{S})) = N(w_1(\mathbf{S}) \mid \mathbf{0}, C_\phi(\mathbf{S},\mathbf{S})) \approx N(w_1(\mathbf{S}) \mid \mathbf{0}, \tilde{C}_\phi(\mathbf{S},\mathbf{S})). \tag{3}$$

The approximation $\tilde{C}_\phi(\mathbf{S},\mathbf{S})$ is computationally efficient because its inverse $\tilde{C}_\phi(\mathbf{S},\mathbf{S})^{-1}$ is sparse and can be written as:

$$\tilde{C}_\phi(\mathbf{S},\mathbf{S})^{-1} = (I - A_S)^T D_S^{-1}(I - A_S),$$

where A_S is a lower triangular matrix and D_S is a diagonal matrix. To compute A_S, suppose $N_m(i)$ be the set of column indices of the m nearest neighbors that contain nonzero entries in the ith row of A_S, where $m = 1,\ldots,S-1$ is a hyperparameter controlling the size of nearest neighbors. Let $A_S = [a_1 : \cdots : a_S]^T$ and $D_S = \mathrm{diag}(d_1, d_2, \ldots, d_S)$. The first row of A_S has all elements equal to 0 and $d_1 = C_\phi(1,1)$. For $i = 2,\ldots,S$, we obtain the nonzero entries at column positions indexed by $N_m(i)$ in A_S and the diagonal elements in D_S as follows:

$$a_i(N_m(i)) = C_\phi(i, N_m(i)) C_\phi(N_m(i), N_m(i))^{-1} \text{ and} \tag{4}$$

$$d_i = C_\phi(i,i) - C_\phi(i, N_m(i)) C_\phi(N_m(i), N_m(i))^{-1} C_\phi(N_m(i), i). \tag{5}$$

3.1.2. Spatio-temporal correlations via NNGP in ZINB

In this section, we describe how to embed spatiotemporal correlations through NNGP in the proposed Bayesian ZINB framework. As we assume GP priors for both the spatial and temporal random effects in the proposed model, the latent NNGP can be used for one random process or both when needed. Since the latent NNGP is designed for a spatial process initially, we use the spatial random effects to demonstrate the application of latent NNGP for easier understanding. When the number of spatial locations S is large, the computation bottleneck in the conjugate Gaussian regression model lies in computing $K_{\sigma_{11},l_{11}}^{-1}$ and $K_{\sigma_{21},l_{21}}^{-1}$ when updating σ_{11} and σ_{21}. To facilitate it, we can find a sparse alternative for

Table 2
Parameter estimates and 95% credible intervals for the proposed model, the spatiotemporal ZINB model by Neelon (2019), and the traditional ZINB model in simulation 1. Note that the comparative models have different model assumptions and therefore do not have posterior results for hyperparameters used in the paper.

Para.	True value	Proposed model	Bayesian ZINB by Neelon (2019)	Traditional ZINB
α_0	−0.25	−0.07 (−1.68, 1.57)	0.25 (−0.05, 0.68)	0.62 (0.43, 0.81)
α_1	0.25	0.17 (0.06, 0.29)	0.20 (0.07, 0.35)	−0.24 (−0.32, −0.16)
β_0	0.50	0.82 (−0.72, 2.47)	0.92 (0.69, 1.14)	−0.45 (−0.91, 0.02)
β_1	−0.25	−0.20 (−0.27, −0.12)	−0.20 (−0.28, −0.12)	−0.22 (−0.37, −0.07)
l_{11}	0.35	1.05 (0.64, 1.71)		
σ_{11}	0.50	1.81 (1.40, 2.31)		
l_{12}	1.00	3.45 (0.45, 4.92)		
σ_{12}	0.20	0.77 (0.43, 1.19)		
l_{21}	0.35	0.32 (0.25, 1.44)		
σ_{21}	0.50	1.21 (0.87, 2.15)		
l_{22}	1.00	2.55 (0.15, 4.63)		
σ_{22}	0.20	0.81 (0.42, 1.28)		
$\sigma_{\epsilon_{11}}$	0.05	0.05 (0.03, 0.08)		
$\sigma_{\epsilon_{12}}$	0.05	0.05 (0.03, 0.08)		
$\sigma_{\epsilon_{21}}$	0.05	0.05 (0.03, 0.08)		
$\sigma_{\epsilon_{22}}$	0.05	0.05 (0.03, 0.08)		
r	1.00	0.95 (0.63, 1.27)	0.73 (0.52, 1.14)	

the Cholesky decomposition of $K^{-1}_{\sigma_{11},l_{11}}$ and $K^{-1}_{\sigma_{21},l_{21}}$. We use the covariance matrix for the spatial random effects from the binary component as an example to illustrate the methodology. First, we rewrite the covariance matrix $C_\phi(\mathbf{S},\mathbf{S}) = K_{\sigma_{11},l_{11}}(h_1,\ldots,h_S) = \sigma_{11}\rho_{l_{11}}$. Let $\tilde{\rho}_{l_{11}}$ be the NNGP approximation of $\rho_{l_{11}}$ with its inverse $\tilde{\rho}^{-1}_{l_{11}} = (I - A_S)^T D_S^{-1}(I - A_S)$, as described in Section 3.1.1. Nonzero entries in the ith row where $i = 2,\ldots,N$ of A_S and D_S in (4) and (5) become

$$a_i(N_m(i)) = \rho_{l_{11}}(i, N_m(i))(\rho_{l_{11}}(N_m(i), N_m(i)))^{-1} \tag{6}$$

and

$$d_i = 1 - a_i(N_m(i))\rho_{l_{11}}(N_m(s_i), i). \tag{7}$$

In this same fashion, we model the temporal correlation. The latent NNGP is applied to facilitate to computation bottleneck when updating the hyperparameters for the spatial and temporal random effects. Hence, when the spatial size and/or temporal size are large, the latent NNGP can be used. Otherwise, one may just compute the inverse of the covariance matrix. The step-by-step posterior sampling procedure is described in the appendix.

4. Simulated and real data analyses

In this section, we apply the proposed Bayesian ZINB-NNGP model to both simulated and COVID-19 fatality data. For the simulation data, a total number of N observations are generated from S locations across T time points. First, we consider exactly one repetition for each sampling unit (s,t), we vary the spatial sizes to show if the proposed model can be generalized well to a large spatial dimension. We set $S = 200$ and $T = 20$ in simulation 1 and set $S = 500$ and $T = 20$ in simulation 2. Second, we relax the constraint on the number of repetitions for each sampling unit by simulating a dynamic number of repetitions for each sampling unit from a Poisson distribution. This is shown in simulation 3.

For the COVID-19 analysis, the daily level data on the total number of COVID-19 deaths in Florida counties in the early stage from 3/25/2020 through 7/29/2020 are analyzed. There are 69 counties ($S = 69$) and 127 days ($T = 127$) in the dataset. The reported death count of county s at day t, denoted by $y_{s,t}$, is the response variable. We apply the proposed model to analyze the association of COVID-19 death at-risk and count rates with social vulnerability, other sociodemographic characteristics (i.e., health insurance coverage, urbanicity), population health care resources (i.e., primary care physicians), population health measures (i.e., life expectancy, obesity), and population density.

4.1. Simulation data

4.1.1. Simulation 1: a moderate spatial dimension and one repetition in each sampling unit

In simulation 1, we have simulated a dataset with $S = 200$ and $T = 20$, which results in a total sample size of $N = 4000$. For the fixed effect, we sample $x \sim \mathcal{N}(0,1)$, set the intercept and slope in the binary component at $\alpha_0 = -0.25$ and $\alpha_1 = 0.25$, and the intercept and slope in the count component as $\beta_0 = 0.5$ and $\beta_1 = -0.25$. The spatial locations are generated from a unit square. The latent NNGP approximation is used when updating σ_{11} and σ_{21} for the spatial random effects. For the size of nearest neighbors used in the latent NNGP model, we have tested different values from 8 to 15 and found that the results are relatively stable. In the results reported for both the simulations and COVID analysis, we set the size of nearest neighbors as 13. The posterior means and 95% credible intervals as well as the true values of the parameters and hyperparameters are presented in Table 2. We also use the Bayesian spatiotemporal ZINB model with random effects

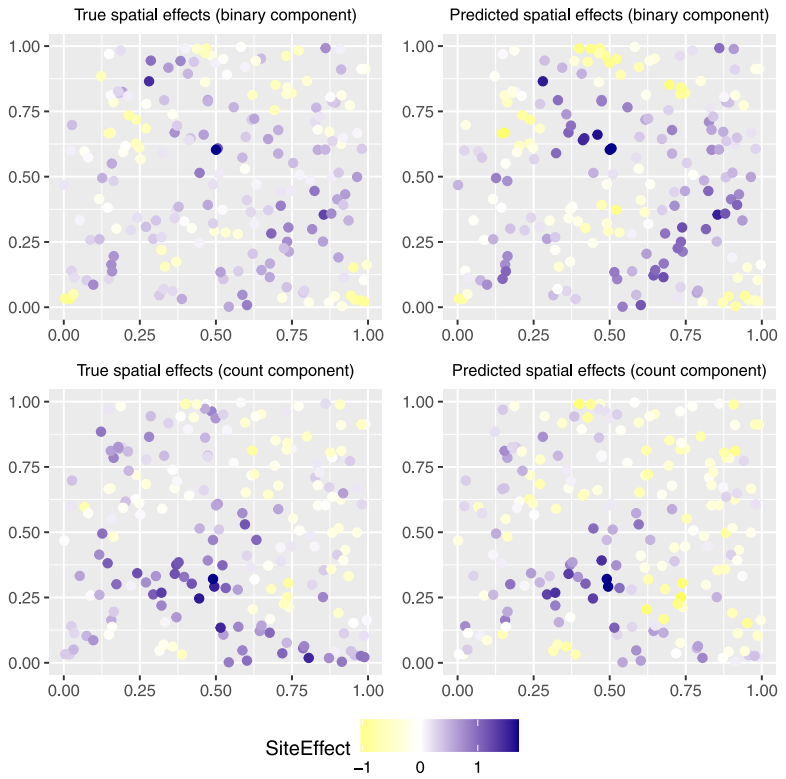

Fig. 1. The true and fitted spatial effects in simulation 1 using the proposed model. Each dot represents a location and the color represent the scale of the spatial random effects. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

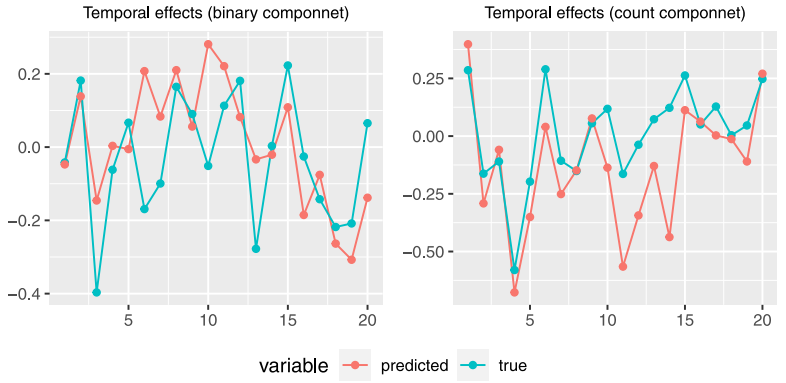

Fig. 2. The true and fitted temporal effects in simulation 1 using the proposed model.

proposed by Neelon (2019) and ZINB regression models from R package 'pscl' (Jackman et al., 2015) for comparison. Due to the difference in model structure, both comparative models do not have results for the hyperparameters used in the proposed framework.

The results in Table 2 show that the 95% credible intervals of the proposed model have covered the true values for all the parameters (i.e., $\alpha_0, \alpha_1, \beta_0, \beta_1$), while the intervals of the Bayesian ZINB model by Neelon (2019) do not cover the true values of α_0 and β_0 and the intervals of the traditional ZINB do not cover the true values for parameters α_0, α_1 and β_0. Fig. 1 presents the map of the true and posterior mean spatial random effects in the binary component (in the top panel) and the count component (in the bottom panel). In both components, the fitted spatial pattern closely mirrors the true distribution, suggesting the proposed model recovers the underlying spatial pattern in the data. Fig. 2 presents the true and fitted temporal effects for each temporal point. The predicted temporal patterns capture the trend of the true patterns in general.

Table 3
Parameter estimates and 95% credible intervals for the proposed model and the traditional ZINB model in simulation 2.

Para.	True value	Proposed model	Traditional ZINB
α_0	−0.25	0.00 (−1.94, 1.84)	0.32 (0.14, 0.50)
α_1	0.25	0.20 (0.13, 0.28)	−0.23 (−0.28, −0.17)
β_0	0.50	0.07 (−2.05, 2.14)	−0.20 (−0.59, 0.20)
β_1	−0.25	−0.23 (−0.28, −0.17)	−0.23 (−0.34, −0.12)
l_{11}	0.35	3.02 (2.13, 3.54)	
σ_{11}	0.50	2.87 (2.34, 3.45)	
l_{12}	1.00	3.92 (2.37, 4.94)	
σ_{12}	0.20	0.77 (0.44, 1.22)	
l_{21}	0.35	0.39 (0.20, 2.89)	
σ_{21}	0.50	1.74 (1.22, 3.41)	
l_{22}	1.00	3.05 (1.16, 4.79)	
σ_{22}	0.20	0.83 (0.43, 1.33)	
$\sigma_{\epsilon_{11}}$	0.05	0.05 (0.03, 0.07)	
$\sigma_{\epsilon_{12}}$	0.05	0.05 (0.03, 0.08)	
$\sigma_{\epsilon_{21}}$	0.05	0.05 (0.03, 0.08)	
$\sigma_{\epsilon_{22}}$	0.05	0.05 (0.03, 0.08)	
r	1.00	1.08 (0.80, 1.29)	

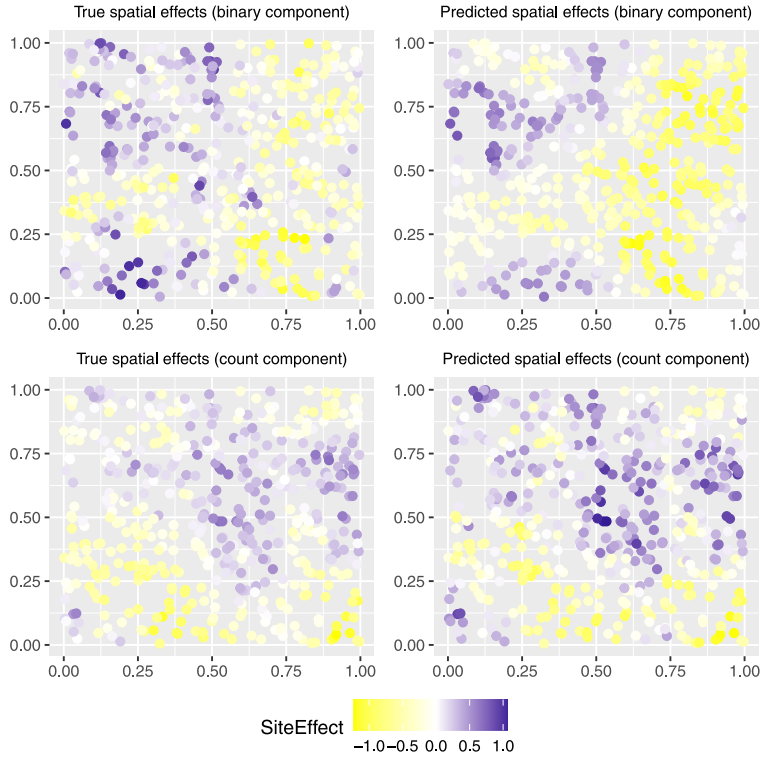

Fig. 3. The true and fitted spatial effects in simulation 2 using the proposed model.

4.1.2. Simulation 2: a large spatial dimension and one repetition in each sampling unit

In simulation 2, we use the same setting as simulation 1 and increase the spatial dimension from 200 to 500, which results in a total sample size of $N = 10{,}000$. The latent NNGP approximation is used to update σ_{11} and σ_{21} for the spatial random effects. The true and fitted results are shown in Table 3. The 95% credible intervals of the proposed model have covered the true values for all the parameters, while the intervals of the traditional ZINB model do not cover the true values for most parameters. We find that the Bayesian ZINB algorithm proposed by Neelon (2019) breaks in the first few iterations due to singularity issues which are caused by the high spatial dimension.

The true and posterior mean spatial and temporal random effects of simulation 2 are shown in Figs. 3 and 4. Both show that the fitted random effects are consistent with the true random effects.

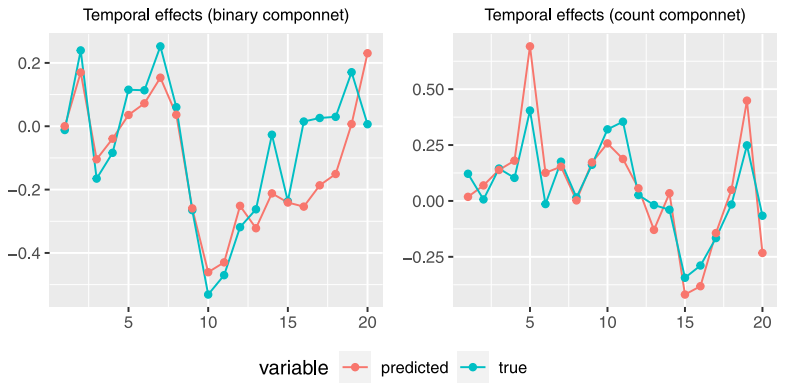

Fig. 4. The true and fitted temporal effects in simulation 2 using the proposed model.

Table 4
Parameter estimates and 95% credible intervals in simulation 3 using the proposed model and the spatiotemporal ZINB model.

Para.	True value	Proposed method	Bayesian ZINB by Neelon (2019)	Traditional ZINB
α_0	−0.25	−0.27 (−2.00, 1.44)	−0.34 (−0.48, −0.18)	0.45 (0.30, 0.60)
α_1	0.25	0.18 (0.11, 0.26)	0.19 (0.11, 0.26)	−0.20 (−0.26, −0.15)
β_0	0.50	0.31 (−1.68, 2.25)	0.38 (0.2, 0.55)	−0.05 (−0.34, 0.25)
β_1	−0.25	−0.20 (−0.26, −0.15)	−0.20 (−0.25, −0.15)	−0.21 (−0.31, −0.12)
l_{11}	0.35	1.32 (0.98, 1.86)		
σ_{11}	0.50	1.95 (1.53, 2.44)		
l_{12}	1.00	3.91 (2.04, 4.96)		
σ_{12}	0.20	0.74 (0.43, 1.14)		
l_{21}	0.35	0.39 (0.26, 1.77)		
σ_{21}	0.50	1.37 (0.86, 2.59)		
l_{22}	1.00	3.29 (0.99, 4.88)		
σ_{22}	0.20	0.80 (0.47, 1.24)		
$\sigma_{\epsilon_{11}}$	0.05	0.05 (0.03, 0.08)		
$\sigma_{\epsilon_{12}}$	0.05	0.05 (0.03, 0.08)		
$\sigma_{\epsilon_{21}}$	0.05	0.05 (0.03, 0.07)		
$\sigma_{\epsilon_{22}}$	0.05	0.05 (0.03, 0.08)		
r	1.00	1.19 (0.93, 1.46)	1.13 (0.90, 0.55)	

4.1.3. Simulation 3: dynamic number of repetitions in each sampling unit

In simulation 3, we run the simulation with a dynamic number of repetitions in each sampling unit. We set $S = 200$ and $T = 20$ like simulation 1. For each sampling unit (s, t) the number of repetitions $n_{s,t}$ is randomly sampled from a Poisson distribution with mean 2, which results in a total sample size of $N = 7867$. The posterior estimates and 95% credible regions are presented in Table 4. As the spatial size is relatively small and there are repetitions in each sampling unit in simulation 3, the Bayesian ZINB (Neelon, 2019) is comparable to the proposed model. Figs. 5 and 6 show the true and fitted spatial and temporal random effects using the proposed model. The fitted random effects are better than those in simulation 1 and 2 where there is only one repetition per sampling unit.

4.2. COVID-19 fatality rates modeling

COVID-19 has impacted populations around the world, with the fatality rate varying dramatically across counties of Florida (Karmakar et al., 2021; Khedhiri, 2021; Zhang et al., 2021b). Using statistical methods to identify factors which are associated with COVID-19 fatality/mortality rates is one of the most important topics. In the early stage of COVID-19, many counties were at low risk of having positive cases and did not have reported deaths. Hence, the proposed ZINB-NNGP model is a good option to model the mortality rate with excessive zeros. In this section, we apply the proposed model to understand the association of COVID-19 mortality at-risk and count rates with social vulnerability, adjusting other sociodemographic characteristics (i.e., health insurance coverage, urbanicity), population health care resources (i.e., primary care physicians), population health measures (i.e., life expectancy, obesity), and population density. We are not able to include the results of the Bayesian ZINB by Neelon (2019) for the COVID analysis because this Bayesian ZINB algorithm cracks after few iterations due to a singularity issue in high-dimensional spatial/temporal covariance matrix inversions, which also happened in simulation 2.

We consider the daily COVID-19 deaths in Florida at the county level reported from 3/25/2020 through 7/29/2020. More specifically, because of the variation in the schedule on which deaths are reported by days of week, we analyze the 7-day average deaths from the New York Times GitHub repository (https://github.com/nytimes/covid-19-data). The

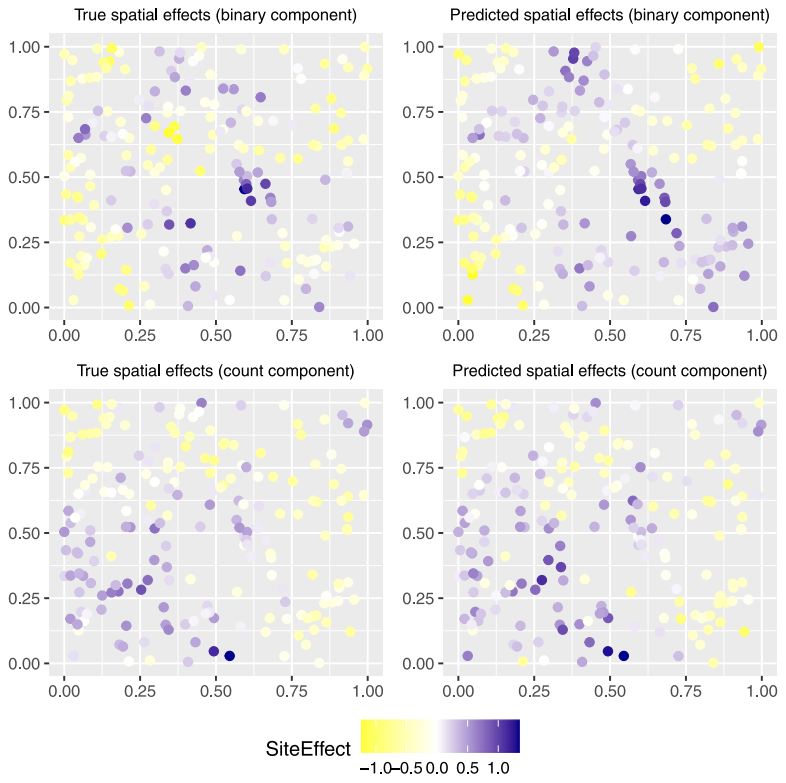

Fig. 5. The true and fitted spatial effects in simulation 3 using the proposed model.

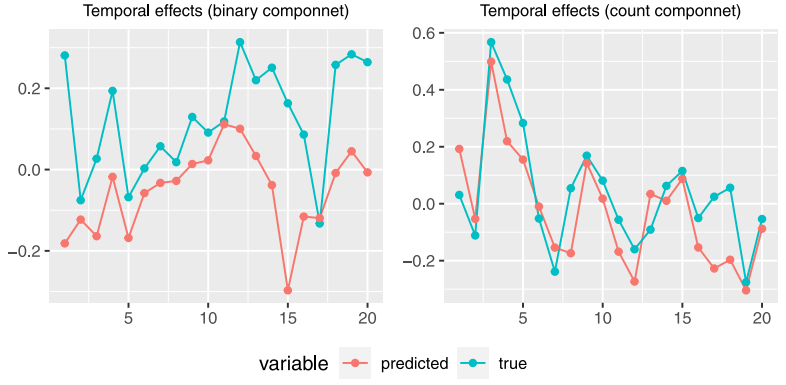

Fig. 6. The true and fitted temporal effects in simulation 3 using the proposed model.

final analytic sample comprises 8351 observations (67 counties across 127 study days). The county level deaths in Florida during the selected time period are shown in the left plot of Fig. 11. There were approximately 74% counties with zero deaths reported in total, 14% counties with an average of one death per day and 12% counties with 2 ore more deaths per day.

Social vulnerability is measured by the social vulnerability index (SVI), which is a national and state-specific county ranking system and designed to assist public health officials in identifying communities in need of support and resources in an event of pandemic (Centers for Disease Control and Prevention, 2018). The SVI has four different themes, namely, socioeconomic status, household composition, race/ethnicity/language, and housing/transportation, which comprise the overall SVI. We use the overall SVI in the analysis and categorize it to three levels: high (>75 percentile), moderate (between 25 percentile and 75 percentile), and low (<25 percentile). A higher level of overall SVI indicates more vulnerability, as suggested in the literature (Hughes et al., 2021; Bruckhaus et al., 2022). We also adjust the following county-level characteristics in the model: other sociodemographic characteristics (i.e., health insurance coverage, urbanicity); population health care resources (i.e., primary care physicians); population health measures (i.e., life expectancy,

Table 5
Data sources.

Variable	Year	Source
COVID-19 deaths	2020	The New York Times
Social Vulnerability Index and component measures		
Overall Social Vulnerability Index	2014–2018	Centers for Disease Control and Prevention
Socioeconomic status index	2014–2018	Centers for Disease Control and Prevention
Poverty rate	2014–2018	Centers for Disease Control and Prevention
Household characteristics and disability index	2014–2018	Centers for Disease Control and Prevention
65 years or older	2014–2018	Centers for Disease Control and Prevention
Minority status and language index	2014–2018	Centers for Disease Control and Prevention
Minority	2014–2018	Centers for Disease Control and Prevention
Housing type and transportation index	2014–2018	Centers for Disease Control and Prevention
No health insurance coverage	2015–2019	US Census Bureau
Public transportation to commute to work	2015–2019	US Census Bureau
Urbanicity	2013	US Department of Agriculture
Total primary care physicians	2019	Health Resources & Services Administration
Life expectancy	2016–2018	Robert Wood Johnson Foundation
Obesity	2016	Robert Wood Johnson Foundation
Population Density	2014–2018	Centers for Disease Control and Prevention

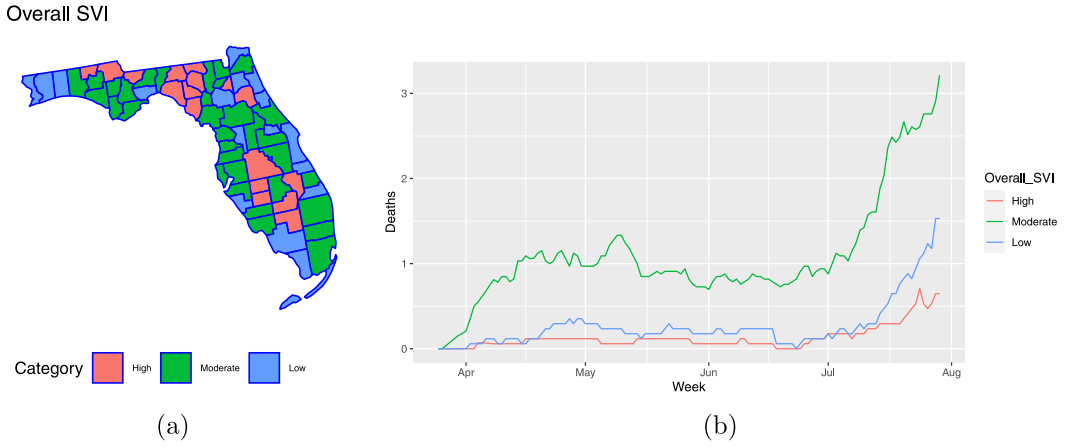

Fig. 7. (a) The map of Florida with colors according to the county's overall SVI category. (b) The daily average deaths by overall SVI in Florida from 3/25/2020 to 7/29/2020. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

obesity); and population density. The variables used in the analysis and the data sources are listed in Table 5 (Dasgupta et al., 2020).

A map of the overall SVI among Florida counties (panel (a)) and a time series plot of the reported daily average deaths for each SVI category (panel (b)) are shown in Fig. 7. In panel (a), it is shown that most counties have moderate overall SVI (green color in the map). In panel (b), the counties with moderate overall SVI have more deaths than the high and low group. It is different from the finding that counties with high SVI had more COVID-19 deaths in other research focused on the U.S. or other locations (Karmakar et al., 2021; Neelon et al., 2022). To further understand how the COVID-19 deaths spread across the three SVI categories, we show the map of the daily average deaths in four different months in Fig. 8. Counties at a higher risk such as Miami-Dade and Hillsborough have more COVID-19 deaths and they have moderate overall SVI scores.

Table 6 shows the estimated coefficients as well as the 95% credible intervals in the binary and count component. Counties with higher overall SVI levels are more likely to report deaths (binary component). This is consistent with the previous research (Karmakar et al., 2021). Meanwhile, the percentage of people under age 65 without health insurance is associated with a higher risk of reporting deaths, which is consistent with finding from other studies (Mountantonakis et al., 2021; DuPre et al., 2021). It is also found that obesity is associated with a lower risk of reporting deaths, although studies have suggested that obesity is a risk factor for death from COVID-19 (Tartof et al., 2020; Rottoli et al., 2020). This may be due to potentially confounding effects of obesity with the overall SVI (An and Xiang, 2015). Population density is associated with a higher risk of reporting deaths and more deaths, which is consistent with other research (Neelon et al., 2021; Foo et al., 2021).

The fitted daily average death by overall SVI category is shown in panel (a) of Fig. 9. The fitted trend (solid line) is consistent with the actual (dotted line). In the earlier time (March 2020 through the middle of July 2020), there is

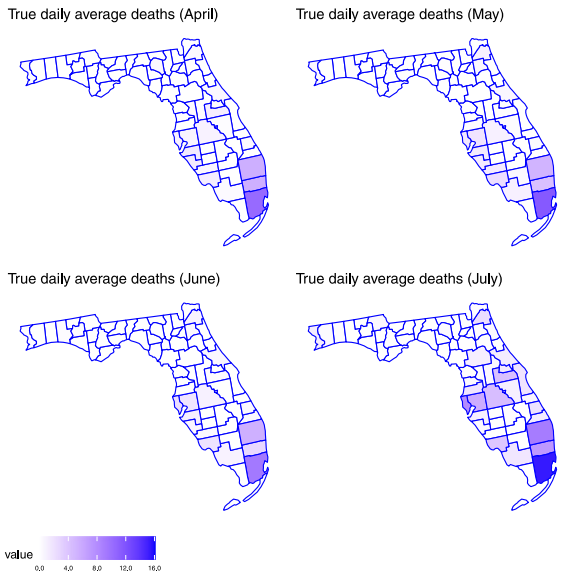

Fig. 8. The true daily average deaths by month in the study period.

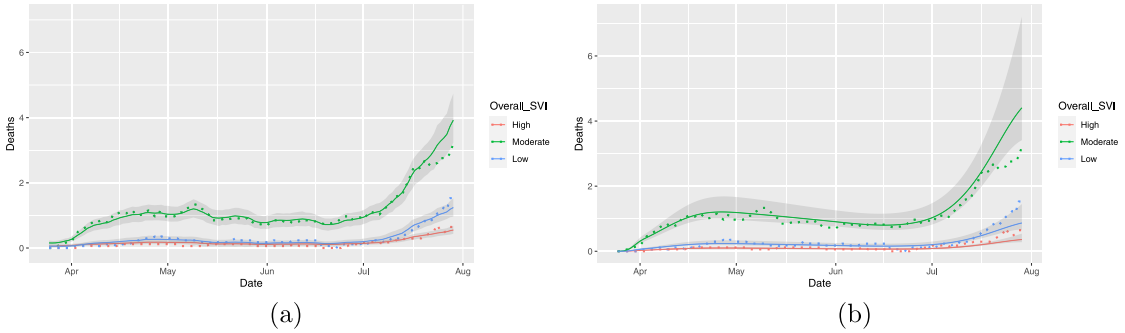

Fig. 9. The true and fitted daily average deaths in Florida by SVI category using the proposed model in panel (a) and the B-splines spatiotemporal model proposed by Neelon et al. (2022) in panel (b). The dotted lines are the true daily average, solid lines are the fitted daily average, and the grey ribbons are the 95% credible region.

Table 6
The parameter estimates and 95% credible intervals of the Florida counties' sociodemographic and health characteristics with COVID-19 deaths.

Variable	Binary component	Count component
Overall SVI	1.06 (0.011, 2.109)	0.418 (−0.412, 1.229)
No health insurance coverage	0.308 (0.114, 0.492)	0.014 (−0.118, 0.148)
Urbanicity	0.021 (−0.471, 0.484)	−0.261 (−0.706, 0.153)
Total primary care physicians	−0.017 (−0.038, 0.004)	−0.005 (−0.019, 0.009)
Life expectancy	−0.049 (−0.14, 0.051)	−0.009 (−0.181, 0.154)
Obesity	−0.248 (−0.399, −0.104)	−0.062 (−0.187, 0.062)
Population density	1.387 (0.658, 2.109)	1.415 (0.883, 1.993)

approximately one death each day for counties with moderate SVI and zero deaths for counties with low or high SVI. Starting from July 2020, the death counts has increased for all counties in Florida, and counties in the moderate category had the most significant increase than others. We also apply a recently developed spatiotemporal model (not in a zero-inflated setting) (Neelon et al., 2022) which used B-splines to model temporal effects for the COVID-19 analysis. We notice that it could be difficult to set the location and number of knots appropriately. Its fitted daily average deaths by SVI category are shown in panel (b) of Fig. 9. There is a higher predictive uncertainty at the end of the temporal period for all three overall SVI categories.

To further compare the death counts across the three SVI categories, risk ratios (RRs) for the high and moderate SVI versus the low SVI are shown in Fig. 10, the counties in moderate SVI category have a higher risk of reporting deaths than

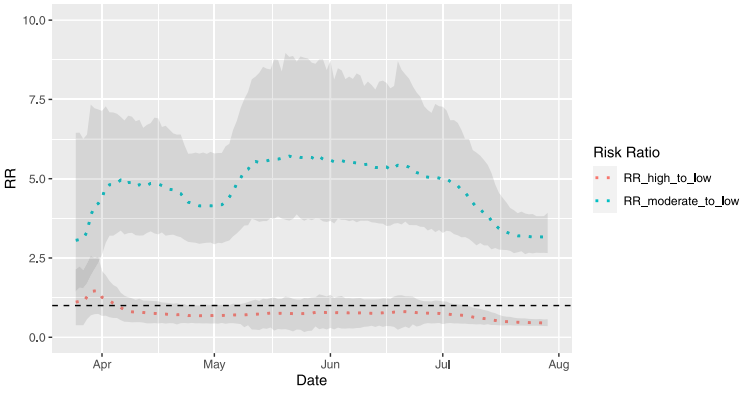

Fig. 10. The RRs for counties with high and moderate social vulnerability versus low social vulnerability. The dotted lines are posterior mean RRs and the grey ribbons were the 95% credible regions. RRs > 1 indicate higher risk. The black dashed line is the reference line with risk ratio $= 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

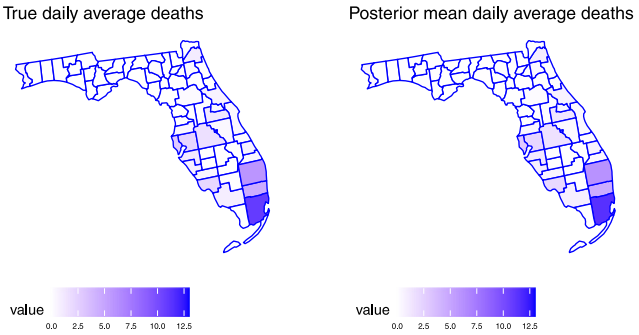

Fig. 11. The true and fitted daily average deaths in Florida by counties.

the counties in the low SVI category (the blue dotted line). The counties in the high SVI category have a similar risk of reporting death with counties in the low SVI category between March 2020 and June 2020 and lower risk in July 2020 (red dotted line).

The county level observed average death counts and posterior mean average death counts in the overall SVI model is shown in Fig. 11. The proposed model identifies counties with high death counts such as Miami-Dade.

In Fig. 12, we select four counties from the three SVI categories to show the individual fitted daily deaths: Hernando county (moderate SVI), Broward county (high SVI), Polk county (low SVI), and Miami-Dade county (moderate SVI). The fitted curves capture the overall patterns but are generally more smooth than the observed death curves which are more wiggly. The model does not capture the downward trend in Miami-Dade (see panel (d) of Fig. 12) in the later period of July 2020. This may be related to emergency orders of closing certain indoor spaces and outdoor spaces where groups of people congregated without physical distancing in July 2020, which are not included in the analysis. For comparisons, we show the fitting results using the spatiotemporal model proposed in Neelon et al. (2022) in Fig. 13. The posterior mean daily deaths do not capture the trends very well and the 95% CIs fail to cover most the true values for Hernando county in panel (a), Polk county in panel (c), and Miami-Dade county in panel (d). The result for Broward county in panel (b) is similar to that of the proposed model.

5. Discussion

We have introduced a Bayesian framework for zero-inflated negative binomial regression models with spatiotemporal effects. The proposed model is able to simultaneously incorporate a flexible structure for the spatial effects and temporal for the binary and count component. The Gibbs sampler is efficient for posterior inference via the Pólya-Gamma data augmentation and latent NNGP conditioned on these latent variables through the Bayesian hierarchical inference. The use of the latent NNGP to approximate the covariance inverse matrix empowers the proposed to be feasible when the spatial or temporal dimension is large. Our simulations suggest that the proposed model is comparable to existing methods when such comparisons are available and more applicable in a higher dimension. It also outperforms other methods when there is only one repetition in each sampling unit.

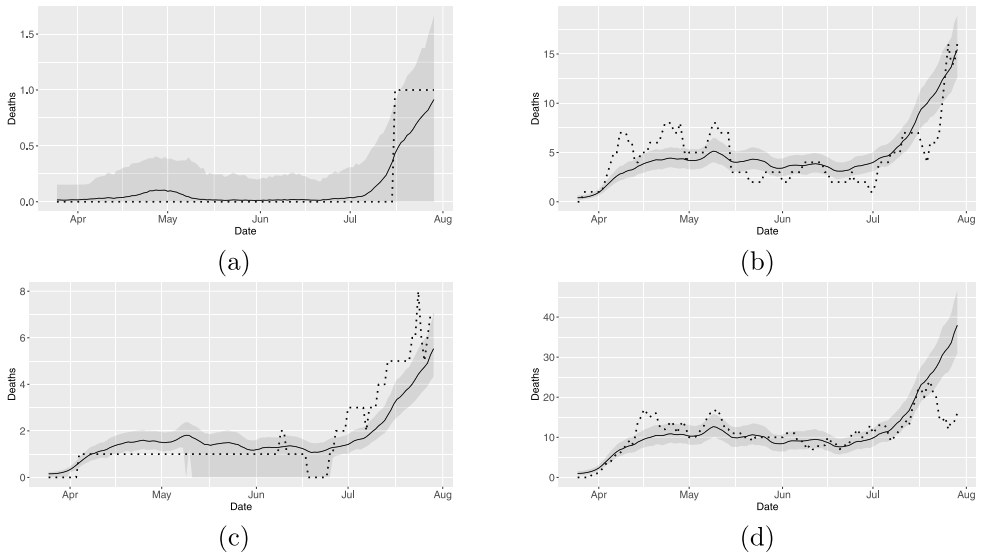

Fig. 12. The true and posterior mean daily deaths in Hernando county (a), Broward county (b), Polk county (c), and Miami-Dade county (d) from 3/25/2020 to 7/29/2020 using our proposed method. The dotted lines are the true deaths and the solid lines are the fitted deaths. The grey ribbons are the 95% credible intervals.

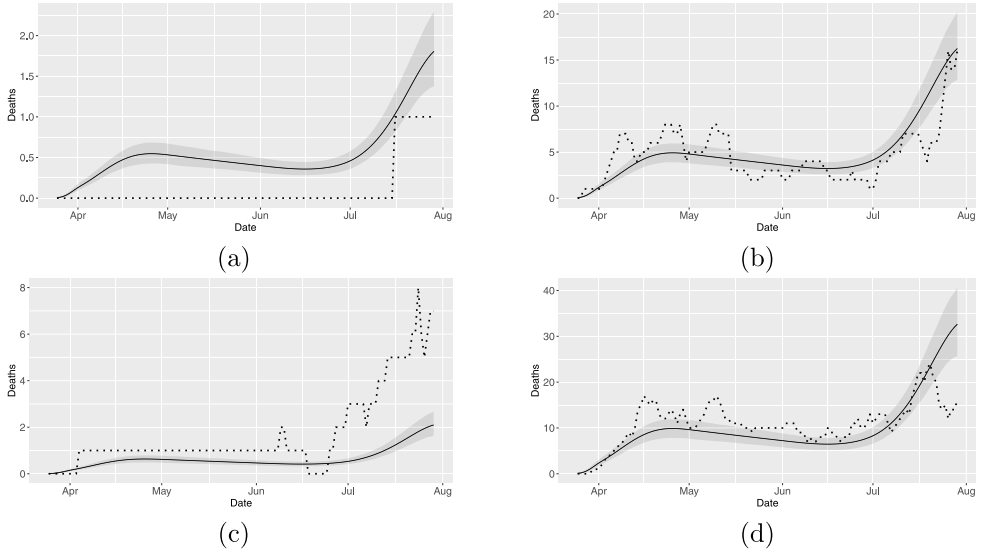

Fig. 13. The true and posterior mean daily deaths in Hernando county (a), Broward county (b), Polk county (c), and Miami-Dade county (d) from 3/25/2020 to 7/29/2020 using the B-splines spatiotemporal model proposed by Neelon et al. (2022). The dotted lines are the true deaths and the solid lines are the fitted deaths. The grey ribbons are the 95% credible intervals.

There are several potential improvements for future work. There is potential for selecting important variables in this Bayesian framework. For example, by using multivariate Bayesian models with polynomial-tailed shrinkage priors (Wang et al., 2023) for the covariance matrix, one can derive variable selection methods. Such methods would be useful to model ultrahigh dimensional spatiotemporal data. Additionally, the proposed model assumes a common temporal effect across different spatial locations and a common spatial effect across different temporal points. Future study could be focused on dynamic settings that capture the interaction between spatial and temporal effects.

6. Software

Software in the form of R code, together with a sample input data set and complete documentation is available on request from the corresponding author (hsin.huang@ucf.edu).

CRediT authorship contribution statement

Qing He: Data collection and analysis, Writing – original draft, Software, Validation. **Hsin-Hsiung Huang:** Conceptualization, Methodology, Supervision, Writing – reviewing & editing.

Declaration of competing interest

We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

Acknowledgments

The authors thank the reviewers for providing valuable comments. We would like to acknowledge support for this project from the National Science Foundation, United States (NSF grants DMS-1924792 and DMS-2318925).

Appendix

The step-by-step posterior sampling procedure

The update algorithm for parameters of interests as well as random effects and hyperparameters are shown below. In each component, we update parameters and random effects together as the posterior distribution are both Gaussian, which can facilitate convergence of the MCMC algorithm in simulation studies.

Step 1: Update the latent at-risk indicators

The update for W_j depends on the value of y_j. If $y_j > 0$, then the jth subject belongs to the at-risk class and $W_j = 1$. If $y_j = 0$, W_j is updated by drawing from a Bernoulli distribution with probability

$$Pr(W_j = 1 \mid y_j = 0, \text{rest}) = \frac{\pi_j \nu_j^r}{1 - \pi_j(1 - \nu_j^r)},$$

where $\pi_j = \frac{\exp(\eta_{1j})}{1+\exp(\eta_{1j})}$ is the unconditional probability that $W_j = 1$ and $\nu_j = 1 - \psi_j$.

Step 2: Update $\alpha, \vec{a}, \vec{b}$

Denote $[\alpha, \vec{a}, \vec{b}]^T$ by ϕ_1. Assume a $\mathcal{N}(\phi_0, \Sigma_0)$ prior, we update ϕ_1 from the posterior distribution in two steps: $\omega_{1j} \mid \eta_{1j} \sim \mathcal{PG}(1, \eta_{1j})$ and $\phi_1 \mid \Omega_1, z_1 \sim \mathcal{N}(\mu, \Sigma)$, where

$$\begin{aligned}\Sigma &= (\Sigma_0^{-1} + V^T \Omega_1 V)^{-1},\\ \mu &= \Sigma\left(\Sigma_0^{-1}\phi_0 + V^T \Omega_1 (z_1 - V_1\vec{\epsilon}_{11} - V_2\vec{\epsilon}_{12})\right),\end{aligned}$$

and $z_{1j} = \frac{W_j - 1/2}{\omega_{1j}}$ is the latent variable, $\Omega_1 = \text{diag}(\omega_{1j})$, and $V = [X, V_1, V_2]$ is the augmented design matrix.

Step 3: Update $\vec{\epsilon}_{11}$

Conditional on $\alpha, \vec{a}, \vec{b}, \vec{\epsilon}_{12}$, assuming a $\mathcal{N}(0, \sigma^2_{\epsilon_{11}})$ prior, update $\vec{\epsilon}_{11}$ from a $\mathcal{N}(\mu, \Sigma)$ distribution where

$$\begin{aligned}\Sigma &= \left(\frac{1}{\sigma^2_{\epsilon_{11}}} I + V_1^T \Omega_1 V_1\right)^{-1},\\ \mu &= \Sigma\left(V_1^T \Omega_1 (z_1 - X\alpha - V_1\vec{a} - V_2\vec{b} - V_2\vec{\epsilon}_{12})\right).\end{aligned}$$

Step 4: Update $\vec{\epsilon}_{12}$

Conditional on $\alpha, \vec{a}, \vec{b}, \vec{\epsilon}_{11}$, assuming a $\mathcal{N}(0, \sigma^2_{\epsilon_{12}})$ prior, update $\vec{\epsilon}_{12}$ from a $\mathcal{N}(\mu, \Sigma)$ distribution, where

$$\begin{aligned}\Sigma &= \left(\frac{1}{\sigma^2_{\epsilon_{12}}} I + V_2^T \Omega_1 V_2\right)^{-1},\\ \mu &= \Sigma\left(V_2^T \Omega_1 (z_1 - X\alpha - V_1\vec{a} - V_2\vec{b} - V_1\vec{\epsilon}_{11})\right).\end{aligned}$$

Step 5: Update l_{11}, σ_{11}

First, we use the Metropolis-Hasting algorithm to update l_{11}. Using a normal proposal distribution which is symmetric, the acceptance probability can be written as follows:

$$\begin{aligned}A &= \min\left(1, \frac{\pi(l^*_{11})}{\pi(l_{11})}\right)\\ &= \min\left(1, \frac{p(\vec{a} \mid K_{\sigma_{11}, l^*_{11}}(h_1, \ldots, h_S))p(l^*_{11} \mid a_{l_1}, b_{l_1})}{p(\vec{a} \mid K_{\sigma_{11}, l_{11}}(h_1, \ldots, h_S))p(l_{11} \mid a_{l_1}, b_{l_1})}\right).\end{aligned}$$

Then, σ_{11} can be sampled from the full conditional as

$$\sigma_{11}^2 \sim \mathcal{IG}\left(a_{\sigma_1} + \frac{S}{2}, b_{\sigma_1} + \frac{\vec{a}^T (K_{\sigma_{11},l_{11}}(h_1, \ldots, h_S))^{-1}\vec{a}}{2}\right)$$

The posterior distribution of σ_{11} requires the inverse of $K_{\sigma_{11},l_{11}}(h_1, \ldots, h_S)$, which may be a computation bottleneck when the size of spatial locations is large. Hence, we apply the latent conjugate NNGP to approximate the inverse matrix.

Step 6: Update l_{12}, σ_{12}^2

The update for l_{12}, σ_{12}^2 is similar to the one for l_{11}, σ_{11}^2. Depending on the size of temporal points, the latent conjugate NNGP can also be applied in the update for σ_{12}^2 if T is large. First, we use the Metropolis-Hasting algorithm to update l_{12}. Using a normal proposal distribution which is symmetric, the acceptance probability can be written as follows:

$$\begin{aligned} A &= \min\left(1, \frac{\pi(l_{12}^*)}{\pi(l_{12})}\right) \\ &= \min\left(1, \frac{p(\vec{b} \mid K_{\sigma_{12},l_{12}^*}(w_1, \ldots, w_T))p(l_{12}^* \mid a_{l_2}, b_{l_2})}{p(\vec{b} \mid K_{\sigma_{12},l_{12}}(w_1, \ldots, w_T))p(l_{12} \mid b_{l_2}, b_{l_2})}\right). \end{aligned}$$

Then, σ_{12} can be sampled from the full conditional as

$$\sigma_{12}^2 \sim \mathcal{IG}\left(a_{\sigma_1} + \frac{T}{2}, b_{\sigma_2} + \frac{\vec{b}^T (K_{\sigma_{12},l_{12}}(w_1, \ldots, w_T))^{-1}\vec{b}}{2}\right).$$

Step 7: Update $\sigma_{\epsilon_{11}}$

Assuming a $\mathcal{IG}(a_\epsilon, b_\epsilon)$ prior, draw $\sigma_{\epsilon_{11}}$ from the posterior distribution:

$$\sigma_{\epsilon_{11}}^2 \sim \mathcal{IG}\left(a_\epsilon + \frac{S}{2}, b_\epsilon + \frac{\sum_{s=1}^S \epsilon_{11s}^2}{2}\right).$$

Step 8: Update $\sigma_{\epsilon_{12}}$

Assuming a $\mathcal{IG}(a_\epsilon, b_\epsilon)$ prior, draw $\sigma_{\epsilon_{12}}$ from the posterior distribution:

$$\sigma_{\epsilon_{12}}^2 \sim \mathcal{IG}\left(a_\epsilon + \frac{T}{2}, b_\epsilon + \frac{\sum_{t=1}^T \epsilon_{12t}^2}{2}\right).$$

Step 9: Update $\beta, \vec{c}, \vec{d}$

Denote $[\beta, \vec{c}, \vec{d}]^T$ by ϕ_2. Assume a $\mathcal{N}(\phi_0, \Sigma_0)$ prior, we update ϕ_2 in two steps:

$$\begin{aligned} \omega_{2j}|\eta_{2j} &\sim \mathcal{PG}(y_j + r, \eta_{2j}), \forall j \text{ s.t. } W_j = 1 \\ \phi_2|\Omega_2, z_2 &\sim \mathcal{N}(\mu, \Sigma), \end{aligned}$$

where

$$\begin{aligned} \Sigma &= (\Sigma_0^{-1} + V^{*T}\Omega_2 V^*)^{-1} \\ \mu &= \Sigma\left(\Sigma_0^{-1}\phi_0 + V^{*T}\Omega_2(z_2 - V_1^*\vec{\epsilon}_{21} - V_2^*\vec{\epsilon}_{22})\right), \end{aligned}$$

where z_2 is the latent variable with $z_{2j} = \frac{Y_j - r}{2\omega_{2j}}$, $\Omega_2 = \text{diag}(\omega_{2j})$, and $V^* = [X^*, V_1^*, V_2^*]$ is the augmented design matrix for observations at risk ($W_j = 1$).

Step 10: Update $\vec{\epsilon}_{21}$

Conditional on $\beta, \vec{c}, \vec{d}, \vec{\epsilon}_{22}$, assuming a $\mathcal{N}(0, \sigma_{\epsilon_{21}}^2)$ prior, update $\vec{\epsilon}_{21}$ from a $\mathcal{N}(\mu, \Sigma)$ distribution where

$$\Sigma = \left(\frac{1}{\sigma_{\epsilon_{2S}}^2} I + (V_1^*)^T \Omega_2 V_1^*\right)^{-1}, \tag{8}$$

$$\mu = \Sigma\left((V_1^*)^T \Omega_2(z_2 - X\beta - V_1^*\vec{c} - V_2^*\vec{d} - V_2^*\vec{\epsilon}_{22})\right). \tag{9}$$

Step 11: Update $\vec{\epsilon}_{22}$

Conditional on $\beta, \vec{c}, \vec{d}, \vec{\epsilon}_{21}$, assuming a $\mathcal{N}(0, \sigma_{\epsilon_{22}}^2)$ prior, update $\vec{\epsilon}_{22}$ from a $\mathcal{N}(\mu, \Sigma)$ distribution, where

$$\Sigma = \left(\frac{1}{\sigma_{\epsilon_{22}}^2} I + (V_2^*)^T \Omega_2 V_2^*\right)^{-1}, \tag{10}$$

$$\mu = \Sigma\left((V_2^*)^T \Omega_2(z_2 - X\beta - V_1^*\vec{c} - V_2^*\vec{d} - V_1^*\vec{\epsilon}_{21})\right). \tag{11}$$

Step 12: Update l_{21}, σ_{21}

The update of l_{21} is similar to that of l_{11}. The Metropolis-Hasting algorithm is used to update l_{21} with acceptance probability

$$A = \min\left(1, \frac{\pi(l_{21}^*)}{\pi(l_{21})}\right) = \min\left(1, \frac{p(\vec{c} \mid K_{\sigma_{21},l_{21}^*}(h_1,\ldots,h_S))p(l_{21}^* \mid a_{l_1}, b_{l_1})}{p(\vec{c} \mid K_{\sigma_{21},l_{21}}(h_1,\ldots,h_S))p(l_{21} \mid a_{l_1}, b_{l_1})}\right).$$

Hyperparameter σ_{21} is drawn from the full conditional distribution

$$\sigma_{21}^2 \sim \mathcal{IG}\left(a_{\sigma_1} + \frac{S}{2}, b_{\sigma_1} + \frac{\vec{c}^T(K_{\sigma_{21},l_{21}}(h_1,\ldots,h_S))^{-1}\vec{c}}{2}\right),$$

where $K^{-1}_{\sigma_{21},l_{21}}$ is approximated using latent NNGP.

Step 13: Update l_{22}, σ_{22}^2

First, l_{22} is updated using the Metropolis-Hasting algorithm with acceptance probability

$$A = \min\left(1, \frac{\pi(l_{22}^*)}{\pi(l_{22})}\right) = \min\left(1, \frac{p(\vec{d} \mid K_{\sigma_{22},l_{22}^*}(w_1,\ldots,w_T))p(l_{22}^* \mid a_{l_2}, b_{l_2})}{p(\vec{d} \mid K_{\sigma_{22},l_{22}}(w_1,\ldots,w_T))p(l_{22} \mid b_{l_2}, b_{l_2})}\right).$$

Then, σ_{22} can be sampled from the full conditional as

$$\sigma_{22}^2 \sim \mathcal{IG}\left(a_{\sigma_2} + \frac{T}{2}, b_{\sigma_2} + \frac{\vec{d}^T(K_{\sigma_{22},l_{22}}(w_1,\ldots,w_T))^{-1}\vec{d}}{2}\right).$$

Step 14: Update $\sigma_{\epsilon_{21}}$

Assuming a $\mathcal{IG}(a_\epsilon, b_\epsilon)$ prior, draw $\sigma_{\epsilon_{21}}$ from the posterior distribution:

$$\sigma_{\epsilon_{21}}^2 \sim \mathcal{IG}\left(a_\epsilon + \frac{S}{2}, b_\epsilon + \frac{\sum_{s=1}^S \epsilon_{21s}^2}{2}\right).$$

Step 15: Update $\sigma_{\epsilon_{22}}$

Assuming a $\mathcal{IG}(a_\epsilon, b_\epsilon)$ prior, draw $\sigma_{\epsilon_{22}}$ from the posterior distribution:

$$\sigma_{\epsilon_{22}}^2 \sim \mathcal{IG}\left(a_\epsilon + \frac{T}{2}, b_\epsilon + \frac{\sum_{t=1}^T \epsilon_{22t}^2}{2}\right).$$

References

An, R., Xiang, X., 2015. Social vulnerability and obesity among US adults. Int. J. Health Sci. 3 (3), 7–21.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2003. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC.

Bruckhaus, A.A., Abedi, A., Salehi, S., Pickering, T.A., Zhang, Y., Martinez, A., Lai, M., Garner, R., Duncan, D., 2022. COVID-19 vaccination dynamics in the US: coverage velocity and carrying capacity based on socio-demographic vulnerability indices in California. J. Immigr. Minority Health 24 (1), 18–30.

Casella, G., George, E.I., 1992. Explaining the Gibbs sampler. Amer. Statist. 46 (3), 167–174.

Centers for Disease Control and Prevention, 2018. CDC/ATSDR Social Vulnerability Index 2018 Database Florida. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html. (Accessed 1 February 2022).

Chalupka, K., Williams, C.K., Murray, I., 2013. A framework for evaluating approximation methods for Gaussian process regression. J. Mach. Learn. Res. 14, 333–350.

Cheung, Y.B., 2002. Zero-inflated models for regression analysis of count data: a study of growth and development. Stat. Med. 21 (10), 1461–1469.

Cousin, A., Maatouk, H., Rullière, D., 2016. Kriging of financial term-structures. European J. Oper. Res. 255 (2), 631–648.

Cressie, N., Wikle, C.K., 2015. Statistics for Spatio-Temporal Data. John Wiley & Sons.

Dadaneh, S.Z., Zhou, M., Qian, X., 2018. Bayesian negative binomial regression for differential expression with confounding factors. Bioinformatics 34 (19), 3349–3356.

Dasgupta, S., Bowen, V.B., Leidner, A., Fletcher, K., Musial, T., Rose, C., Cha, A., Kang, G., Dirlikov, E., Pevzner, E., et al., 2020. Association between social vulnerability and a county's risk for becoming a COVID-19 hotspot—United States, June 1–July 25, 2020. MMWR Morb. Mortal Wkly. Rep. 69 (42), 1535.

Datta, A., Banerjee, S., Finley, A.O., Gelfand, A.E., 2016a. Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. J. Amer. Statist. Assoc. 111 (514), 800–812.

Datta, A., Banerjee, S., Finley, A.O., Hamm, N.A., Schaap, M., 2016b. Nonseparable dynamic nearest neighbor Gaussian process models for large spatio-temporal data with an application to particulate matter analysis. Ann. Appl. Stat. 10 (3), 1286.

Diana, A., Dennis, E., Matechou, E., Morgan, B., 2021. Fast Bayesian inference for large occupancy data sets, using the polya-Gamma scheme. arXiv preprint arXiv:2107.14656.

DuPre, N.C., Karimi, S., Zhang, C.H., Blair, L., Gupta, A., Alharbi, L.M.A., Alluhibi, M., Mitra, R., McKinney, W.P., Little, B., 2021. County-level demographic, social, economic, and lifestyle correlates of COVID-19 infection and death trajectories during the first wave of the pandemic in the United States. Sci. Total Environ. 786, 147495.
Foo, O., Hiu, S., Teare, D., Syed, A.A., Razvi, S., 2021. A global country-level analysis of the relationship between obesity and COVID-19 cases and mortality. Diabetes Obes. Metab. 23 (12), 2697–2706.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall/CRC.
Ghosh, S.K., Mukhopadhyay, P., Lu, J.-C.J., 2006. Bayesian analysis of zero-inflated regression models. J. Statist. Plann. Inference 136 (4), 1360–1375.
Gramacy, R.B., 2016. laGP: large-scale spatial modeling via local approximate Gaussian processes in R. J. Stat. Softw. 72, 1–46.
Greene, W.H., 1994. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. NYU Working Paper No. EC-94-10.
Gu, X., Yan, X., Ma, L., Liu, X., 2020. Modeling the service-route-based crash frequency by a spatiotemporal-random-effect zero-inflated negative binomial model: An empirical analysis for bus-involved crashes. Accid. Anal. Prev. 144, 105674.
Hughes, M.M., Wang, A., Grossman, M.K., Pun, E., Whiteman, A., Deng, L., Hallisey, E., Sharpe, J.D., Ussery, E.N., Stokley, S., et al., 2021. County-level COVID-19 vaccination coverage and social vulnerability—United States, December 14, 2020–March 1, 2021. MMWR Morb. Mortal Wkly. Rep. 70 (12), 431.
Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., Meers, Z., Jackman, M.S., Imports, M., 2015. Package 'pscl'. Political Sci. Comput. Lab. 18 (04.2017).
Kang, E.L., Cressie, N., 2011. Bayesian inference for the spatial random effects model. J. Amer. Statist. Assoc. 106 (495), 972–983.
Karmakar, M., Lantz, P.M., Tipirneni, R., 2021. Association of social and demographic factors with COVID-19 incidence and death rates in the US. JAMA Netw. Open 4 (1), e2036462.
Khedhiri, S., 2021. Statistical modeling of COVID-19 deaths with excess zero counts. Epidemiol. Methods 10 (s1).
Kim, S., Lu, Z., Cohen, A.S., 2018. An improved estimation using polya-gamma augmentation for Bayesian structural equation models with dichotomous variables. Measur. Interdiscip. Res. Perspect. 16 (2), 81–91.
Kimeldorf, G.S., Wahba, G., 1970. A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. Ann. Math. Stat. 41 (2), 495–502.
Lázaro-Gredilla, M., Quinonero-Candela, J., Rasmussen, C.E., Figueiras-Vidal, A.R., 2010. Sparse spectrum Gaussian process regression. J. Mach. Learn. Res. 11, 1865–1881.
Lewsey, J.D., Thomson, W.M., 2004. The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. Community Dent. Oral Epidemiol. 32 (3), 183–189.
Liu, H., Ong, Y.-S., Shen, X., Cai, J., 2020. When Gaussian process meets big data: A review of scalable GPs. IEEE Trans. Neural Netw. Learn. Syst. 31 (11), 4405–4423.
Lunn, D., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D., 2013. The BUGS book. In: A Practical Introduction to Bayesian Analysis. Chapman Hall, London.
McCullagh, P., Nelder, J.A., 2019. Generalized Linear Models. Routledge.
Mountantonakis, S.E., Epstein, L.M., Coleman, K., Martinez, J., Saleh, M., Kvasnovsky, C., Brown, R.-M., McCulloch, E., Kuvin, J., Richardson, S., et al., 2021. The association of structural inequities and race with out-of-hospital sudden death during the COVID-19 pandemic. Circ. Arrhythm. Electrophysiol. 14 (5), e009646.
Neelon, B., 2019. Bayesian zero-inflated negative binomial regression based on pólya-gamma mixtures. Bayesian Anal. 14 (3), 829–855.
Neelon, B., Mutiso, F., Mueller, N.T., Pearce, J.L., Benjamin-Neelon, S.E., 2021. Spatial and temporal trends in social vulnerability and COVID-19 incidence and death rates in the United States. PLoS One 16 (3), e0248702.
Neelon, B.H., O'Malley, A.J., Normand, S.-L.T., 2010. A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. Stat. Model. 10 (4), 421–439.
Neelon, B., Wen, C.-C., Benjamin-Neelon, S.E., 2022. A multivariate spatiotemporal model for tracking COVID-19 incidence and death rates in socially vulnerable populations. J. Appl. Stat. 1–24.
Pillow, J., Scott, J., 2012. Fully Bayesian inference for neural models with negative-binomial spiking. Adv. Neural Inf. Process. Syst. 25, 1898–1906.
Polson, N.G., Scott, J.G., Windle, J., 2013. Bayesian inference for logistic models using Pólya–Gamma latent variables. J. Amer. Statist. Assoc. 108 (504), 1339–1349.
Quinonero-Candela, J., Rasmussen, C.E., 2005. A unifying view of sparse approximate Gaussian process regression. J. Mach. Learn. Res. 6, 1939–1959.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/.
Rottoli, M., Bernante, P., Belvedere, A., Balsamo, F., Garelli, S., Giannella, M., Cascavilla, A., Tedeschi, S., Ianniruberto, S., Del Turco, E.R., et al., 2020. How important is obesity as a risk factor for respiratory failure, intensive care admission and death in hospitalised COVID-19 patients? Results from a single Italian centre. Eur. J. Endocrinol. 183 (4), 389–397.
Smola, A.J., Bartlett, P.L., 2001. Sparse greedy Gaussian process regression. In: Adv. Neural Inf. Process. Syst.. pp. 619–625.
Snelson, E., Ghahramani, Z., 2007. Local and global sparse Gaussian process approximations. In: Artificial Intelligence and Statistics. PMLR, pp. 524–531.
Stan Development Team, 2021. RStan: the R interface to Stan. URL https://mc-stan.org/. R package version 2.21.3.
Tartof, S.Y., Qian, L., Hong, V., Wei, R., Nadjafi, R.F., Fischer, H., Li, Z., Shaw, S.F., Caparosa, S.L., Nau, C.L., et al., 2020. Obesity and mortality among patients diagnosed with COVID-19: results from an integrated health care organization. Ann. Intern. Med. 173 (10), 773–781.
Titsias, M., 2009. Variational learning of inducing variables in sparse Gaussian processes. In: Artificial Intelligence and Statistics. PMLR, pp. 567–574.
Vasudevan, S., Ramos, F., Nettleton, E., Durrant-Whyte, H., 2009. Gaussian process modeling of large-scale terrain. J. Field Robotics 26 (10), 812–840.
Wahba, G., 1990. Spline Models for Observational Data. SIAM.
Wang, S.-H., Bai, R., Huang, H.-H., 2023. Mixed-type multivariate Bayesian sparse variable selection with shrinkage priors. arXiv:2201.12839.
Wang, X., Chen, M.-H., Kuo, R.C., Dey, D.K., 2016. Dynamic spatial pattern recognition in count data. In: New Developments in Statistical Modeling, Inference and Application. Springer, pp. 185–202.
Yau, K.K., Wang, K., Lee, A.H., 2003. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. Biom. J. 45 (4), 437–452.
Zhang, L., Banerjee, S., Finley, A.O., 2021a. High-dimensional multivariate geostatistics: A Bayesian matrix-normal approach. Environmetrics e2675.
Zhang, Y., Chang, H.H., Iuliano, A.D., Reed, C., 2022. Application of Bayesian spatial-temporal models for estimating unrecognized COVID-19 deaths in the United States. Spatial Stat. 100584.
Zhang, H.-Y., Zhang, A.-R., Lu, Q.-B., Zhang, X.-A., Zhang, Z.-J., Guan, X.-G., Che, T.-L., Yang, Y., Li, H., Liu, W., et al., 2021b. Association between fatality rate of COVID-19 and selenium deficiency in China. BMC Infect. Dis. 21 (1), 1–8.
Zhou, M., Carin, L., 2013. Negative binomial process count and mixture modeling. IEEE Trans. Pattern Anal. Mach. Intell. 37 (2), 307–320.
Zuur, A.F., et al., 2012. Zero inflated models and generalized linear mixed models with R. Number 574.50182 Z8.