Biometrics 000, 000–000 DOI: 000
000 0000

Marginal Bayesian Nonparametric Model for Time to Disease Arrival of Threatened Amphibian Populations

Haiming Zhou[1], Timothy Hanson[1,*], and Roland Knapp[2]

[1]Department of Statistics, University of South Carolina, Columbia, South Carolina, U.S.A.

[2]Sierra Nevada Aquatic Research Laboratory, University of California, Mammoth Lakes, California, U.S.A.

email: hansont@stat.sc.edu

Summary: The global emergence of *Batrachochytrium dendrobatidis* (Bd) has caused the extinction of hundreds of amphibian species worldwide. It has become increasingly important to be able to precisely predict time to Bd arrival in a population. The data analyzed herein present a unique challenge in terms of modeling because there is a strong spatial component to Bd arrival time and the traditional proportional hazards assumption is grossly violated. To address these concerns, we develop a novel marginal Bayesian nonparametric survival model for spatially correlated right-censored data. This class of models assumes that the logarithm of survival times marginally follow a mixture of normal densities with a linear dependent Dirichlet process prior as the random mixing measure, and their joint distribution is induced by a Gaussian copula model with a spatial correlation structure. To invert high-dimensional spatial correlation matrices, we adopt a full-scale approximation that can capture both large- and small-scale spatial dependence. An efficient Markov chain Monte Carlo algorithm with delayed rejection is proposed for posterior computation, and an R package `spBayesSurv` is provided to fit the model. This approach is first evaluated through simulations, then applied to threatened frog populations in Sequoia-Kings Canyon National Park.

Key words: Spatial survival data; Point-referenced; Bayesian nonparametric; Dependent Dirichlet process; Delayed rejection; Full-scale approximation; Copula; Proportional hazards.

This paper has been submitted for consideration for publication in *Biometrics*

1. Introduction

The Earth is currently experiencing the most severe mass extinction of species since the dinosaurs died off 65 million years ago. Scientists estimate that we are currently losing on the order of up to 50,000 species per year, 1,000 to 10,000 times greater than the fossil record (Chivian and Bernstein, 2008). The current mass extinction is almost entirely due to humankind in the form of destruction of natural habitats, but with disease being increasingly recognized as another important driver. The global emergence of *Batrachochytrium dendrobatidis* (Bd), a fungus that can kill frogs within a few weeks, has caused the extinction of hundreds of amphibian species worldwide (Wake and Vredenburg, 2008), including the recent rapid local extinction of many mountain yellow-legged frog populations in the Sierra Nevada mountains of California (Rachowicz et al., 2006; Vredenburg et al., 2010). These impacts of Bd have been described as "*...the most spectacular loss of vertebrate biodiversity due to disease in recorded history...*" (Skerratt et al., 2007). Once the most common amphibian in the region, mountain yellow-legged frogs now inhabit less than one-tenth of their range of one hundred years ago and continue to disappear at an alarming rate.

The mountain yellow-legged frog is a species complex of the southern mountain yellow-legged frog *Rana muscosa* and the Sierra Nevada yellow-legged frog *Rana sierrae* (Vredenburg et al., 2007). In part due to Bd-caused declines, these species were recently listed as "endangered" under the U.S. Endangered Species Act (Federal Register 2014). Bd often spreads in wavelike patterns (Cheng et al., 2011; Lips et al., 2008). One of the authors (Knapp) collected data over a 12-year period by hiking large areas of Sequoia and Kings Canyon National Parks, one of the few areas in the Sierra Nevada that still contained some Bd-negative frog populations at the beginning of this study (Vredenburg et al., 2010). Bd-negative frog populations were discovered during the first park-wide survey of all suitable

habitats (Knapp et al., 2003), and were revisited, typically every 1-2 years, over the study period to determine their Bd status over time.

It has become increasingly important to be able to predict when Bd is likely to arrive in a frog population, considering that interventions may be useful to prevent the extinction of frog populations following the Bd arrival (Vredenburg et al., 2010). We model the Bd arrival time of frog populations from discovery given two baseline measures of Bd proximity: (i) "bdwater," indicating whether Bd was present during the previous year (= 1) or not (= 0) in the watershed containing the frog population of interest, and (ii) "bddist," indicating the linear distance to the nearest Bd-positive location during the previous year. Once Bd arrived at a site, this site was assumed to always be Bd-positive in subsequent years (Vredenburg et al., 2010). The location was recorded as Universal Transverse Mercator (x, y)–coordinates in meters for each frog population. These data provide a unique challenge in terms of modeling: there is a strong spatial component to Bd arrival times across populations, and the proportional hazards assumption is rather severely violated.

Traditionally, most survival models incorporating spatial information have been semiparametric, conditional (so-called frailty) models; these include conditionally proportional hazards models (Banerjee et al., 2003; Hennerfeind et al., 2006), proportional odds models (Banerjee and Dey, 2005; Zhao et al., 2009), and accelerated failure time models (Zhao et al., 2009; Wang et al., 2012). Any of these models is preferable if they actually fit the data, as such semiparametric structure allows for easy conditional interpretation in terms of hazard ratios, odds ratios, or acceleration factors, respectively. However, these models impose constraints on survival densities. For example, both proportional hazards and accelerated failure time models induce stochastically ordered survival times for populations with different covariates or frailties. For this reason, if the semiparametric structure is not appropriate, the addition of frailties may not improve model fit, or can even make it worse by adding noise. Instead of

incorporating a spatial frailty term, Li and Lin (2006) proposed a marginal survival model, which they termed a semiparametric normal transformation model, where survival times are assumed to marginally follow the proportional hazards (PH) model and their joint distribution is specified by a Gaussian copula model with a spatial correlation structure. One advantage of their model over frailty models is that the regression coefficients have population-level interpretations. However, their model imposes constraints so that survival curves from different covariate levels are not allowed to cross, which is unrealistic in many practical applications (see De Iorio et al., 2009), including the mountain yellow-legged frog illustration used in the current study (e.g. see Kaplan-Meier estimates in Web Appendix F). For these data, a more flexible survival model is needed to quantify the risk factors associated with the Bd arrival time of frog populations while taking spatial correlation into account.

Flexible Bayesian nonparametric modeling techniques have been successfully developed for handling complex survival data that traditional semiparametric survival models fail to fit. One appealing feature about nonparametric approaches for estimating survival densities is their ability to avoid unrealistic constraints on how the variance, skewness, shape and even modality change with covariates. The linear dependent Dirichlet process mixture model (De Iorio et al., 2009; Jara et al., 2010, 2011), essentially a countable mixture of accelerated failure time models, provides a flexible way to capture crossing hazard and survival curves. However, it is unclear how to extend these Bayesian nonparametric models to a geostatistical setting for the analysis of spatially correlated survival data, and there has been virtually no related literature so far.

In this article, we develop a marginal Bayesian nonparametric spatial survival model for the analysis of the Bd arrival time of yellow-legged frog populations given the baseline covariates bdwater and bddist. This model assumes that the logarithm of survival times marginally follow a linear dependent Dirichlet process mixture (LDDPM) model, and specifies their

joint distribution via a Gaussian copula model. Two major features of the proposal are that the resulting survival curves are allowed to cross, or not, as dictated by the data, and that the inclusion of spatial information improves the prediction of Bd arrival time dramatically. To invert the high-dimensional spatial correlation matrices, we adopt a full scale approximation that can capture both large- and small-scale spatial dependence (Sang and Huang, 2012). To compare our proposal with the semiparametric normal transformation model proposed by Li and Lin (2006), we also present a Bayesian version of their model, denoted as a marginal PH spatial model. We develop efficient Markov chain Monte Carlo (MCMC) algorithms for both our model and the marginal PH spatial model. The analysis of the mountain yellow-legged frog data shows that our proposed spatial LDDPM model provides improvement over most traditional models, including the non-spatial LDDPM, PH, PH augmented with spatial frailties, and marginal PH spatial. For ease of implementation, we also developed an R package `spBayesSurv`, which is available for downloading at `http://cran.r-project.org/web/packages/spBayesSurv`, that can fit our proposed spatial LDDPM, the non-spatial LDDPM, PH, and the marginal PH spatial models.

The rest of this article is organized as follows. Section 2 describes the marginal LDDPM spatial survival model. Section 3 provides a recipe for efficient MCMC inference, discusses prediction, and proposes a cross-validated predictive model comparison criterion. Section 4 discusses simulation results, shows how ignoring spatial correlation grossly can bias inferences, and validates that our model selection criterion works well. Section 5 presents the analysis of the time to Bd arrival data, and shows how posterior inference can be interpreted and put to practical use. The paper concludes with concluding remarks in Section 6.

2. Marginal LDDPM Spatial Survival Model

2.1 *Model Specification*

Suppose right-censored spatial survival data $\{(t_i^o, \delta_i, \mathbf{x}_i, \mathbf{s}_i) : i = 1, \ldots, n\}$ are collected in a spatial region of interest $\mathcal{D}$, where t_i^o is a recorded event time, δ_i is a censoring indicator equalling 1 if t_i^o is the observed event time and equalling 0 if the event time is right-censored at t_i^o, $\mathbf{x}_i$ is a p-dimensional vector of covariates including the intercept, and $\mathbf{s}_i$ records the spatial location. We denote by t_i the latent true (unobserved if $\delta_i = 0$) event time corresponding to t_i^o, and then the relationship between δ_i and t_i can be described by $\delta_i = I(t_i = t_i^o)$, where $\delta_i = 0$ implies $t_i > t_i^o$. Since event times are all positive, it remains appealing to work on the logarithms of event times, i.e. $y_i = \log(t_i)$, $y_i^o = \log(t_i^o)$. In addition, we assume an independent censoring scheme; that is, the event and censoring times are independent given the observed covariates.

We assume that $y_i|\mathbf{x}_i$ marginally follows a linear dependent Dirichlet process mixture model of De Iorio et al. (2009),

$$F_{\mathbf{x}_i}(y|G) = \int \Phi\left(\frac{y - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) dG\{\boldsymbol{\beta}, \sigma^2\}, \tag{1}$$

where $\Phi(\cdot)$ is the cumulative distribution function (cdf) of the standard normal, and G follows the Dirichlet Process (DP) prior (Ferguson, 1973) with concentration parameter $\alpha > 0$ and base measure G_0 on $\mathbb{R}^p \times \mathbb{R}^+$, denoted by $G \sim DP(\alpha, G_0)$. This Bayesian nonparametric model treats the conditional distribution $F_{\mathbf{x}}$ as a function-valued parameter and allows its variance, skewness, modality and other features to flexibly vary with the $\mathbf{x}$ covariates. See Pati et al. (2013) for the sufficient conditions for posterior consistency. Note that there is no clear approach to incorporate simple spatial frailties into this model, as there is a countable number of linear predictors.

To incorporate the spatial correlation among the event times, we first define a residual

process $z(\mathbf{s}_i)$, after adjusting for covariate effects, as

$$z(\mathbf{s}_i) = \Phi^{-1}\left\{F_{\mathbf{x}_i}(y_i|G)\right\}, i = 1, \ldots, n. \tag{2}$$

Note that $z(\mathbf{s}_i)$ follows the standard normal distribution, providing a natural way of modeling spatial dependence. Specifically, we assume that $\mathbf{z} = (z(\mathbf{s}_1), \ldots, z(\mathbf{s}_n))'$ arises from a zero-mean Gaussian process (GP) $\{z(\mathbf{s}) : \mathbf{s} \in \mathcal{D}\}$ with a valid correlation function $\mathbf{C}_{\boldsymbol{\theta}}(\mathbf{s}, \mathbf{s}')$; that is, $\mathbf{z}$ follows a multivariate Gaussian distribution as $\mathbf{z} \sim N_n(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}})$, where $\mathbf{C}_{\boldsymbol{\theta}} = [\mathbf{C}_{\boldsymbol{\theta}}(\mathbf{s}_i, \mathbf{s}_j)]_{i,j=1}^n$ is the $n \times n$ correlation matrix depending on a parameter vector $\boldsymbol{\theta}$. To allow for a nugget effect, we consider $\mathbf{C}_{\boldsymbol{\theta}}(\mathbf{s}_i, \mathbf{s}_j) = \theta_1 \rho(\mathbf{s}_i, \mathbf{s}_j) + (1 - \theta_1) I(\mathbf{s}_i = \mathbf{s}_j)$, where $\rho(\cdot, \cdot)$ is a valid correlation function, and $\theta_1 \in [0, 1]$, also known as a "partial sill" in Waller and Gotway (2004), is a scale parameter measuring a local maximum correlation. The simplest parameterization for the correlation function $\rho(\cdot, \cdot)$ is an isotropic one, where the spatial correlation is assumed to be a function solely of the Euclidean distance d_{ij} between locations $\mathbf{s}_i$ and $\mathbf{s}_j$. In this paper, we consider the exponential correlation function $\rho(\mathbf{s}, \mathbf{s}') = \exp\{-\theta_2 ||\mathbf{s} - \mathbf{s}'||\}$, where θ_2 controls the spatial decay over distance. Other choices such as the spherical, Gaussian and Matérn correlation functions are also possible.

The above model specification is completely equivalent to Gaussian copula modeling (Song, 2000). In fact we have modeled the joint distribution of $\mathbf{y} = (y_1, \ldots, y_n)$ as a function of its marginal cdf, that is, $\mathbf{y} \sim \Phi_n\left(\Phi^{-1}\{F_{\mathbf{x}_1}(y_1|G)\}, \ldots, \Phi^{-1}\{F_{\mathbf{x}_n}(y_n|G)\}; \mathbf{C}_{\boldsymbol{\theta}}\right)$, where $\Phi_n(\cdot; \boldsymbol{\Sigma})$ is the cdf of an n-dimensional normal with mean zero and covariance matrix $\boldsymbol{\Sigma}$. Following Song (2000), the likelihood function based upon the complete data $\{(y_i, \mathbf{x}_i, \mathbf{s}_i), i = 1, \ldots, n\}$ is

$$\mathcal{L} = |\mathbf{C}_{\boldsymbol{\theta}}|^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{z}'(\mathbf{C}_{\boldsymbol{\theta}}^{-1} - \mathbf{I}_n)\mathbf{z}\right\} \prod_{i=1}^{n} f_{\mathbf{x}_i}(y_i|G), \tag{3}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $f_{\mathbf{x}_i}(y_i|G)$ is the density function corresponding to $F_{\mathbf{x}_i}(y_i|G)$.

2.2 *Prior Specification and Hierarchical Modeling*

For the DP prior $G \sim DP(\alpha, G_0)$, we define the base measure G_0 through independent priors $N_p(\boldsymbol{\beta}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\text{Ga}(\sigma^{-2}|\nu_a, \nu_b)$. Here, $N_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\text{Ga}(\cdot|a, b)$ denote a p-dimensional normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and a gamma distribution with shape a and rate b, respectively. Note that it is critical to select an appropriate prior for the concentration parameter α, since its value controls the number of distinct components to which the data points are allocated. We assume a $\text{Ga}(\alpha|a_0, b_0)$ prior for α, which has been widely used by many researchers because of its tractability. As for the correlation parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$, they are typically not consistently estimable for a wide range of correlation functions, as demonstrated by Zhang (2004). This implies that increasing sample size does not necessarily obliterate the priors' impact, and thus weakly informative priors may be desirable to help identify the parameters $\boldsymbol{\theta}$, say, a $\text{Ga}(\theta_1|\theta_{1a}, \theta_{1b})$ for θ_1 and a $\text{Ga}(\theta_2|\theta_{2a}, \theta_{2b})$ for θ_2 with hyperparameters $\boldsymbol{\theta}_0 = (\theta_{1a}, \theta_{1b}, \theta_{2a}, \theta_{2b})$ being carefully choosen. Finally, we specify conjugate hyperpriors on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^{-1}$ using a normal $N_p(\boldsymbol{\mu}|\mathbf{m}_0, \mathbf{S}_0)$ and a Wishart $W_p\left((\kappa_0\boldsymbol{\Sigma}_0)^{-1}, \kappa_0\right)$, respectively, where the Wishart has mean $\boldsymbol{\Sigma}_0^{-1}$ and degrees of freedom κ_0.

Following de Carvalho et al. (2013), we suggest reasonable default hyperpriors as follows: $a_0 = b_0 = 2$, $\nu_a = 3$, $\nu_b = \hat{\sigma}^2$, $\boldsymbol{\theta}_0 = \mathbf{1}$, $\mathbf{m}_0 = \hat{\boldsymbol{\beta}}$, $\mathbf{S}_0 = \hat{\boldsymbol{\Sigma}}$, $\boldsymbol{\Sigma}_0 = 30\hat{\boldsymbol{\Sigma}}$, and $\kappa_0 = 7$, where $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are the maximum likelihood estimates of $\boldsymbol{\beta}$ and σ^2 from fitting the log-normal accelerated failure time model $\log(t_i) = \mathbf{x}_i'\boldsymbol{\beta} + \sigma\epsilon_i, \epsilon_i \sim N(0, 1)$, and $\hat{\boldsymbol{\Sigma}}$ is the asymptotic covariance estimate for $\hat{\boldsymbol{\beta}}$.

For ease of hierarchical modeling, we express the DP prior G in the stick-breaking form (Sethuraman, 1994) as

$$G = \sum_{k=1}^{\infty} w_k \delta_{(\boldsymbol{\beta}_k, \sigma_k^2)}, \quad w_k = V_k \prod_{j<k}(1 - V_j), \tag{4}$$

where δ_a is a Dirac probability measure concentrated at a, $V_k \overset{iid}{\sim} \text{Beta}(1, \alpha)$ and $(\boldsymbol{\beta}_k, \sigma_k^2) \overset{iid}{\sim} G_0$ are mutually independent for $k = 1, \ldots, \infty$. In practical implementations, either fixed

(Ishwaran and James, 2001) or random stopping (Papaspiliopoulos and Roberts, 2008) approximation procedures of the infinite sum representation (4) can be considered. In this paper, we use the truncation approximation, replacing G with $G^N = \sum_{k=1}^{N} w_k \delta_{(\boldsymbol{\beta}_k, \sigma_k^2)}$, with N being pre-specified, where w_ks result from a truncated version of the stick-breaking construction: $w_1 = V_1,\ w_k = V_k \prod_{j=1}^{k-1}(1 - V_j), k = 2, \ldots, N,\ V_N = 1$. The truncation level N can be determined by considering the properties of the higher-order w_k values in the infinite sum representation (4), i.e. $U_N = \sum_{k=N+1}^{\infty} w_k$. Ishwaran and Zarepour (2000) demonstrated that $E(U_N|\alpha) = \alpha^N/(1+\alpha)^N$ and $\mathrm{Var}(U_N|\alpha) = \alpha^N/(2+\alpha)^N - \alpha^{2N}/(1+\alpha)^{2N}$. Then for any given truncation level N, we can approximate these expressions by averaging over the gamma prior for α. For example, setting $N = 10$ and placing $\mathrm{Ga}(2,2)$ on α in our simulation study will result in $E(U_N) \approx 0.0055$ and $\mathrm{Var}(U_N) \approx 0.0002$, which is more than adequate for data analyses.

In order to determine which component the ith data point is allocated, we introduce configuration variables K_i. Then the hierarchical model for the data, together with the augmented latent true event-times, can be written as follows:

$$
\begin{aligned}
y_i|\mathbf{B}, \boldsymbol{\sigma}^2, \mathbf{K} &\sim N(\mathbf{x}_i'\boldsymbol{\beta}_{K_i}, \sigma_{K_i}^2) \\
z(\mathbf{s}_i) &= \Phi^{-1}\left\{\sum_{k=1}^{N} w_k \Phi\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}_k}{\sigma_k}\right)\right\} \\
(z(\mathbf{s}_1), \ldots, z(\mathbf{s}_n))'|\boldsymbol{\theta} &\sim N_n(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}}) \\
P(K_i = k|\mathbf{V}) &= w_k,\ k = 1, \ldots, N \\
(V_k|\alpha) &\overset{iid}{\sim} \mathrm{Beta}(1, \alpha),\ k = 1, \ldots, N-1 \\
(\boldsymbol{\beta}_k, \sigma_k^{-2})|\boldsymbol{\mu}, \boldsymbol{\Sigma} &\overset{iid}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \times \mathrm{Ga}(\nu_a, \nu_b) \\
\alpha &\sim \Gamma(a_0, b_0) \\
(\theta_1, \theta_2) &\sim \mathrm{Beta}(\theta_{1a}, \theta_{1b}) \times \mathrm{Ga}(\theta_{2a}, \theta_{2b}) \\
(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}) &\sim N_p(\mathbf{m}_0, \mathbf{S}_0) \times W_p\left((\kappa_0 \boldsymbol{\Sigma}_0)^{-1}, \kappa_0\right)
\end{aligned}
\tag{5}
$$

where $\mathbf{B} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_N')$, $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_N^2)$, $\mathbf{K} = (K_1, \ldots, K_n)$, and $\mathbf{V} = (V_1, \ldots, V_N)$.

3. Posterior Inference

3.1 *MCMC Sampling*

We develop an efficient MCMC algorithm for posterior sampling from the hierarchical model representation (5). The full conditionals are straightforward to derive, but most of them are not recognizable due to the incorporation of spatial dependence. A complete description and derivation of the updating steps is available in Web Appendix A . The posterior samples for the model parameters are used for all inferences of interest.

Let $\boldsymbol{\Omega} = (\mathbf{K}, \mathbf{y}, \mathbf{B}, \boldsymbol{\sigma}^2, \mathbf{V}, \alpha, \boldsymbol{\theta})$ denote collectively the model parameters to be updated. Note that the likelihood function involves the inversion and determinant calculation of a very large global correlation matrix $\mathbf{C}_{\boldsymbol{\theta}}$ and these matrix operations have to be repeated for every MCMC iteration. For large values of the sample size n, e.g., $n \geq 500$, we suggest replacing $\mathbf{C}_{\boldsymbol{\theta}}$ with $\mathbf{C}_{\boldsymbol{\theta}}^{\dagger}$ based on the full scale approximation (FSA) approach as described in Web Appendix B. Conditional on all other parameters, K_i is sampled from a multinomial distribution. For updating y_i, $\boldsymbol{\beta}_k$, σ_k^2 and V_k, we notice that the full conditional for each is proportional to a recognizable density multiplied by a common part $\exp\left\{-\frac{1}{2}\mathbf{z}'(\mathbf{C}_{\boldsymbol{\theta}}^{-1} - \mathbf{I}_n)\mathbf{z}\right\}$. Thus we use Metropolis-Hastings (M-H) with independent proposals, where each proposal is based on the recognizable density. However, we observed for this inititial MCMC scheme that often some $\boldsymbol{\beta}_k$ got "stuck" for a long period, leading to poor MCMC mixing. As a remedy, we found that delayed rejection (Tierney and Mira, 1999) works very well. Upon a rejection in the M-H, instead of retaining the same position, a second-stage proposal corresponding to a random walk is proposed. The precision parameter α, and hyper-parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1})$ are updated from their conjugate full conditionals. Finally, to update the correlation parameters

$\boldsymbol{\theta}$, we first take transformations $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)'$ with $\vartheta_1 = \log\left(\frac{\theta_1}{1-\theta_1}\right)$ and $\vartheta_2 = \log(\theta_2)$, and then update $\boldsymbol{\vartheta}$ using adaptive Metropolis-Hastings algorithms (Haario et al., 2001).

Given a set of posterior samples $\{\boldsymbol{\Omega}^{(l)}, l = 1, \ldots, L\}$, the marginal conditional density of log event time y given the covariates $\mathbf{x}$ is estimated by

$$\hat{f}_{\mathbf{x}}(y) = \frac{1}{L}\sum_{l=1}^{L}\sum_{k=1}^{N} w_k^{(l)} \frac{1}{\sigma_k^{(l)}}\phi\left(\frac{y - \mathbf{x}'\boldsymbol{\beta}_k^{(l)}}{\sigma_k^{(l)}}\right), \tag{6}$$

where $\phi(\cdot)$ is the density of the standard normal. The marginal conditional survival and hazard functions can be estimated similarly. Then all the marginal density, survival and hazard functions of an event time $t = \exp\{y\}$ given $\mathbf{x}$ can be easily obtained. An R package `spBayesSurv` accompanying this paper is provided to implement the MCMC algorithm and plot the estimated curves; see Web Appendix G for sample R code.

3.2 *Spatial Prediction*

In geostatistics, one major interest is predicting the survival t_0 at a new location $\mathbf{s}_0$ with associated covariate values $\mathbf{x}_0$. Given the parameters $\boldsymbol{\Omega}$, by noting that $\mathbf{z} \sim N_n(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}})$, we can easily obtain that $z(\mathbf{s}_0) \sim N(\mu(\mathbf{s}_0), \tau^2(\mathbf{s}_0))$ with $\mu(\mathbf{s}_0) = \mathbf{h}(\mathbf{s}_0)'\mathbf{C}_{\boldsymbol{\theta}}^{-1}\mathbf{z}$ and $\tau^2(\mathbf{s}_0) = 1 - \mathbf{h}(\mathbf{s}_0)'\mathbf{C}_{\boldsymbol{\theta}}^{-1}\mathbf{h}(\mathbf{s}_0)$, where $\mathbf{h}(\mathbf{s}_0) = [\theta_1\rho(\mathbf{s}_0, \mathbf{s}_i)]_{i=1}^n$ is an $n \times 1$ vector. Note that both $\mathbf{C}_{\boldsymbol{\theta}}$ and ρ would be replaced by $\mathbf{C}_{\boldsymbol{\theta}}^{\dagger}$ and $\rho^{\dagger}$ respectively if the FSA is used. Based on the definition of $z(\cdot)$ in equation (2) and the N-level truncation of G, we have

$$z(\mathbf{s}_0) = \Phi^{-1}\left\{\sum_{k=1}^{N} w_k \Phi\left(\frac{\log t_0 - \mathbf{x}_0'\boldsymbol{\beta}_k}{\sigma_k}\right)\right\}. \tag{7}$$

It follows that the predictive density of t_0 is given by

$$\begin{aligned} f(t|\boldsymbol{\Omega}, \mathbf{x}_0, \mathbf{s}_0) =& \frac{1}{t \cdot \tau(\mathbf{s}_0)}\phi\left(\frac{z(\mathbf{s}_0) - \mu(\mathbf{s}_0)}{\tau(\mathbf{s}_0)}\right) \\ & \times \frac{\sum_{k=1}^{N} w_k \frac{1}{\sigma_k}\phi\left(\frac{\log t - \mathbf{x}_0'\boldsymbol{\beta}_k}{\sigma_k}\right)}{\phi(z(\mathbf{s}_0))}. \end{aligned} \tag{8}$$

Given a set of posterior samples $\{\boldsymbol{\Omega}^{(l)}, l = 1, \ldots, L\}$ obtained from Section 3.1, we first draw $z^{(l)}(\mathbf{s}_0)$ from $N(\mu(\mathbf{s}_0), \tau^2(\mathbf{s}_0))$ for $l = 1, \ldots, L$, and then make the transformation according to (7) to obtain a sample of predictive event-times $\{t^{(l)}(\mathbf{s}_0), l = 1, \ldots, L\}$ at the new location

$\mathbf{s}_0$. The final predicted value of t_0 can be either the mean or median of $t^{(l)}(\mathbf{s}_0)$s. In practice, it is difficult to observe all the covariates at the whole study region $\mathcal{D}$, so it is not practically feasible to create a map for predictive event times. Alternatively, we may show a spatial map for the residual process, which can be interpreted in a manner similar to the spatial frailties in conditional survival models; that is, the higher the $z(\mathbf{s})$ is, the larger the event time $y(\mathbf{s})$ would be on average. Note that the predictive density of $t(\mathbf{s}_0)$ is simply (8) averaged over the MCMC iterates. It is worth highlighting here that (8) is different from the truncated version of the marginal density in (1). This is due to the fact that the event-times are spatially correlated so that the prediction at a new location will borrow information from locations where the data have been collected. On the other hand, if we assume there is no spatial correlation, i.e. $\theta_1 = 0$, it is easy to see that the predictive density in (8) reduces to the N-level truncation version of (1).

3.3 *Model comparison*

To compare the predictive ability of competing models, we consider the conditional predictive ordinate (CPO) statistic as suggested by Geisser and Eddy (1979). Let $\mathcal{D}_{-i}$ denote the observed data excluding the ith data point. For a given model, the CPO statistic for the ith observation is defined as $\text{CPO}_i = f\left(t_i^o|\mathcal{D}_{-i}\right)^{\delta_i} S\left(t_i^o|\mathcal{D}_{-i}\right)^{1-\delta_i}$, where f and S denote the marginal posterior predictive density and survival functions of t_i^o given $\mathcal{D}_{-i}$ respectively. A higher value of CPO_i under one model implies a better fit of that model to the ith observation. Let $\mathbf{t} = (t_1, \ldots, t_n)$ be the vector of latent true survival times and $\mathbf{t}_{-i}$ be the corresponding vector with the ith element removed. According to the hierarchical model in (5), given all the model parameters $\boldsymbol{\Theta} = (\mathbf{B}, \boldsymbol{\sigma}^2, \mathbf{V}, \boldsymbol{\theta})$, we show that the CPO of Gelfand and Dey (1994) is generalized to

$$\text{CPO}_i = \left(E_{(\mathbf{t},\Theta|\mathcal{D})} \left[\frac{1}{f(t_i^o|\mathbf{t}_{-i},\Theta)^{\delta_i} S(t_i^o|\mathbf{t}_{-i},\Theta)^{1-\delta_i}} \right] \right)^{-1}. \tag{9}$$

See Web Appendix C for its derivation and the expressions of $f(t_i^o|\mathbf{t}_{-i}, \Theta)$ and $S(t_i^o|\mathbf{t}_{-i}, \Theta)$. To give an aggregate summary measure of a model's predictive ability, we define the log pseudo marginal likelihood (LPML) as LPML $= \sum_{i=1}^{n} \log(\text{CPO}_i)$. The LPML is a cross-validated predictive measure: the larger a model's LPML is, the better predictive ability the model has. From (9), one can easily compute LPML from the MCMC output.

4. Simulations

We conduct a simulation study to illustrate the proposed model (LDDPM-spatial) and assess its performance. We also compare it with the Bayesian version of Li and Lin (2006) (PH-spatial, see Web Appendix D) and the model by De Iorio et al. (2009) (LDDPM-ind). All analyses can be run in `R` using the package `spBayesSurv`.

We randomly select 400 locations over a spatial region $[0, 40] \times [0, 100]$ (mirroring from the frog data) and hold out 100 of them for assessing the predictive performance, yielding a total sample of $n = 300$ subjects for estimation. The log event times $y(\mathbf{s}) = \log t(\mathbf{s})$ at these 400 locations are simulated from a mixture model $f(y|x) = 0.4N(3.5 + 0.5x, 1^2) + 0.6N(2.5 - x, 0.5^2)$ with the spatial dependence described in Section 2.1, where x is generated independently from a uniform distribution over $(-1.5, 1.5)$ and $\mathbf{C}_{\boldsymbol{\theta}}$ is specified with $\theta_1 = 0.98$ and $\theta_2 = 0.1$. The choice of this mixture model is based on a modification of the simulation study in De Iorio et al. (2009). The log censoring times are simulated from a uniform distribution on $(3, 4)$ so that the censoring rate is about $20\% \sim 50\%$. This simulation study is referred to as Scenario I, for which 100 Monte Carlo replicate datasets are generated.

First, we fit the LDDPM-spatial model using truncation level $N = 10$ under the default prior specifications. We also fit the LDDPM-spatial model with $\mathbf{C}_{\boldsymbol{\theta}}$ approximated using the FSA approach (denoted as LDDPM-spatial-FSA), where we experiment with $m = 10$ regularly spaced knots and $B = 10$ blocks taken as equally sized squares. Second, we fit the PH-spatial model with default priors. Finally, we fit the LDDPM-ind model using the

same priors as the LDDPM-spatial model. For each MCMC, we retain $10,000$ scans thinned from $50,000$ after a burn-in period of $10,000$ iterations. To assess the prediction ability and accuracy, for each above model, we calculate the LPML and mean squared prediction error (MSPE), where MSPE $= \sum_{i=1}^{100}(y_i - \tilde{y}_i)^2/100$ with y_is being the held-out true log survival times and $\tilde{y}_i$s being the corresponding predicted values based on posterior means. The models are also compared by computing the integrated squared error (ISE) for estimated survival curves, given by ISE $= \int_0^\infty \left\{\hat{S}(y|x) - S(y|x)\right\}^2 dy$, where $\hat{S}(y|x)$ and $S(y|x)$ are estimated and true survival functions given x respectively.

[Table 1 about here.]

The posterior inferences for spatial correlation parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$ under each approach are presented in Table 1, where the bias of corresponding point estimates (i.e. posterior means), the Monte Carlo mean of posterior standard deviation estimates (MEAN-SD), the Monte Carlo standard deviation of point estimates (SD-MEAN), and the Monte Carlo coverage probability of 95% credible intervals (CP) are reported. The results suggest that the point estimates of $\boldsymbol{\theta}$ are almost unbiased, and that the observed biases under the LDDPM-spatial model are much smaller than the corresponding values under the PH-spatial model. The MEAN-SD and SD-MEAN values are fairly close indicating that the posterior standard deviation is an appropriate estimator of the frequentist standard error. The CPs are around the nominal 95% level. In contrast, the PH-spatial model provides substantially lower coverage probabilities. Furthermore, the posterior estimates with FSA are very close to those using the exact model, suggesting that FSA is a good approximation of the correlation matrix $\mathbf{C}_{\boldsymbol{\theta}}$. Table 1 also presents the Monte Carlo mean of computing times under each approach, where we see that FSA does speed up the computation as expected.

[Figure 1 about here.]

Figure 1 shows boxplots of the ISEs for estimated survival curves, LPMLs, and MSPEs

under the considered models. The LDDPM-spatial models (with and without FSA) provide much smaller biases of the fitted survival functions on average, compared with LDDPM-ind and PH-spatial, indicating that either violation of the PH assumption or ignorance of the spatial dependence could bias the inference. As for prediction ability and accuracy, the proposed models (with and without FSA) yield the best prediction performance as measured by both LPML and MSPE, compared with the PH-spatial and LDDPM-ind models. These simulations, i.e., when the truth is known, show that the LPML is consistent with the MSPE and hence validates its use for complex spatial models exhibiting dependence.

In Web Appendix E, we also tested the performance of LDDPM-spatial model when the PH assumption is satisfied and compared it with the PH-spatial model. The results show that two models provide almost the same boxplots of LPMLs and MSPEs, indicating that the LDDP-spatial model is quite competitive even when the PH assumption is satisfied.

5. Application to Frog Data

From 1997 to 2002, all mapped lentic water bodies in Sequoia and Kings Canyon National Parks were surveyed for mountain yellow-legged frogs. Starting in 2002 and continuing through 2011, nearly all the frog populations that were discovered during the initial surveys were visited every 1-2 years. A primary objective of these resurveys was to determine the Bd status of each frog population over time. Bd attacks the keratinized tissues of tadpole mouthparts and disrupts the normal black pigmentation of these structures. Therefore, early efforts to detect Bd in live amphibians often relied on visual inspections of tadpole mouthparts (see Table 4 in Knapp and Morgan, 2006). This method was replaced by a real-time quantitative PCR assay (qPCR) in the mid-2000s (Boyle et al., 2004), and remains the most reliable method for detecting Bd. In our data set, Bd status was determined using the inspections of tadpole mouthparts during resurveys conducted from 2002 to 2004 and using the qPCR after 2004. The data consist of $n = 309$ frog populations that were initially

discovered during park-wide surveys conducted from 1997 to 2002, and then resurveyed regularly through 2011. The observed event time is calculated as the number of years from the initial survey to either Bd arrival or the last resurvey. By the end of the study, about 11% of the frog populations remained Bd-negative (censored). Table 2 presents a summary of the data.

[Table 2 about here.]

We fit the LDDPM-spatial and PH-spatial models to the data using the same prior specifications as previous simulations. We also fit the LDDPM-ind model, the standard PH model $\lambda_i(t) = \lambda_0(t)\exp\{\mathbf{x}_i'\boldsymbol{\beta}\}$ as well as the PH model with a point-referenced frailty term $\lambda_i(t) = \lambda_0(t)\exp\{\mathbf{x}_i'\boldsymbol{\beta} + w(\mathbf{s}_i)\}$ (PH-frailty), where $w(\mathbf{s})$ follows a Gaussian process (GP) with the exponential covariance function, and the baseline hazard $\lambda_0(t)$ is modeled in the same way as the PH-spatial model. Based upon examintion of trace plots for model parameters in both the simulations and real data analyses, for each MCMC run we retain $20,000$ scans thinned from $200,000$ after a burn-in period of $200,000$ iterations. These are grossly conservative numbers; burnin and thinning requirements may change considerably depending on the amount of spatial correlation and data. The Markov chains mix reasonably well for all fitted models. For the LDDPM models, the number of components with non-negligible mass ranged from two to five, indicating at most five components in the mixture. In Web Appendix F, we present the posterior trace plots for $\boldsymbol{\theta}$ under the LDDPM-spatial model.

[Table 3 about here.]

We first obtain the LPML values for all models under consideration as follows: -276.7 for LDDPM-spatial, -304.4 for PH-spatial, -631.5 for LDDPM-ind, -705.3 for PH and -703.4 for PH-frailty. The LDDPM-spatial model provides significantly better prediction as measured by LPML, with differences ranging from 27 to 428. Interestingly, the PH model

augmented with a GP frailty surface hardly improves inference over the standard PH model. In what follows, we only focus on interpreting results from the proposed model. Table 3 shows posterior estimates of the spatial dependence parameters θ_1 and θ_2. The partial sill parameter θ_1 measures the maximum correlation between frog populations if they were located in the same area location. Our analysis shows that such a correlation is very strong, almost equal to 1. The parameter θ_2 controls the decay of spatial dependence over distance measured by kilometers. For instance, the posterior mean of $\theta_2 = 0.133$ indicates that the correlation decays by $1 - e^{-0.133\times 5} \approx 48.6\%$ for every 5-km increase in distance. This tells us that the spatial dependence still does not disappear even when the two frog populations are located 5-km apart.

[Figure 2 about here.]

Figure 2(b) shows, for example, for frogs living in a watershed that was Bd-positive during the previous year, the population-averaged median Bd arrival time is cut from around 9.8 years to 4.5 years, more than half, when the distance to nearest Bd-positive location goes from 7.731-km to 0.443-km. Interestingly, the *shapes* of densities also change, going from unimodal to bimodal in Figures 2(a) and 2(d). This figure clearly shows the invalidation of the standard AFT model for these data. Similarly, Figure 2(c) shows that the non-proportional change in hazards clearly invalidates the PH assumption. Comparing Bd-positive to Bd-negative in the watershed yields strikingly different outcomes. Populations in basins in which Bd is present, holding the distance to the nearest Bd-positive location constant, have a hazard spike near zero reaching up to around 0.7 (Figure 2(f)). This implies that the population-averaged probability of Bd arrival within one-year is about 70%. The corresponding survival curves cross (Figure 2(e)), invalidating most common semiparametric models. These results are for any frog population randomly found in the study region. The simulations in Section 4 show

that if the accompanying spatial information $\mathbf{s}_0$ is also used, prediction will be accurately refined.

[Figure 3 about here.]

We are also interested in predicting which areas have overall lower survival rates. Since the marginal distribution at each location cannot be predicted unless the associated baseline covariates are available, we instead show a spatial map (Figure 3) by smoothing the predicted residual process $z(\mathbf{s})$ at $10,000$ randomly simulated new locations over the national park. One may interpret this map in a manner similar to the frailty map in the GP frailty PH model in the absence of covariate information; lower value of $z(\mathbf{s})$ indicates lower survival rate at location $\mathbf{s}$. Overall, the frog populations living in darker regions became infected by Bd more earlier than lighter areas.

6. Concluding Remarks

This paper presents a unified approach to the nonparametric modeling of point-referenced survival data. The methodology is broadly illustrated on an analysis of Bd arrival time of yellow legged frog populations throughout Sequoia-Kings Canyon National Park. The frog data presents a unique challenge and opportunity, as most common semiparametric survival models are grossly violated and there is a strong spatial component to the data. In the analysis, we considered two important measures of Bd proximity: bdwater and bddist; additional covariates could be easily incorporated into the vector $\mathbf{x}$ in our model (De Iorio et al., 2009; Jara et al., 2010). As expected, the closer a frog population is to a Bd-positive location, the less time that population has until Bd infects them. This analysis shows how modern, cutting edge statistical techniques can be used to understand a real ecological problem, here predicting the arrival time by Bd in Bd-negative frog populations.

One surprising result is just how poorly the PH model can do in terms of prediction when

it is wrong, even when a traditional GP spatial frailty surface is added to the linear predictor. Zhao et al. (2009) note that in a conditional spatial survival setting, the survival model itself is the most important factor affecting prediction, in their case choosing among PO, PH, or AFT model. The survival model itself is more important than whether frailties are included, or whether the frailties are spatially varying or exchangeable. This observation also holds for the frog data: the assumptions that most semiparametric models imply are not met. In particular, the traditional PH model is inadequate relative to a more flexible nonparametric model. Once an appropriate nonparametric model has been chosen, the copula model can drastically improve prediction. It may be that, for many data sets, the marginal copula-based geostatistical modeling approach provides superior prediction over more traditional GP random-effects models, e.g., the types that can be fit in SAS `glimmix` and `mixed` procedures, the free `BayesX` program (Brezger et al., 2005). For this reason we have provided an efficient R package `spBayesSurv` for others to explore the use of these geostatistical copula models.

One drawback of the proposed model is that covariate effects are interpreted by examining the plots, which becomes challenging for high-dimensional covariates. Due to the inverse of high-dimensional correlation matrices during MCMC, the LDDPM model can also suffer longer computational times. In addition, the proposed model cannot handle time-dependent covariates. The traditional PH model can handle time-dependent covariates, and can have time-varying regression effects, which are not considered here nor were considered by Li and Lin (2006). However, once the PH model is augmented with time-varying effects, inference is also reduced to the examination of plots (e.g. survival curves and densities) and one might as well start with a purely nonparametric model. Future research will examine the fit of so-called additive spatial PH models, i.e. with time-varying regression effects.

An alternative way to incorporate spatial dependence into model (1) is to assume a multivariate latent stick-breaking process prior (Rodríguez et al., 2010) for the random

mixing measure G. This approach uses a copula representation only on the random mixing measure G but requires constraints on the atoms in the mixing measure. The model is broadly developed for the univariate case, but multivariate cases are problematic in that the order constraint becomes complex even for two dimensions. The dimensionality of $(\boldsymbol{\beta}_k, \boldsymbol{\sigma}_k^2)$ in our approach is four, but likely to be much higher in general. By taking the full LDDPM model as the marginal in the copula representation, analysis is greatly simplified.

Supplementary Material

Web Appendices referenced in Sections 1, 3, 4, and 5, and sample R code fitting the proposed model are available with this paper at the Biometrics website on Wiley Online Library.

References

Banerjee, S. and Dey, D. K. (2005). Semiparametric proportional odds models for spatially correlated survival data. *Lifetime Data Analysis* **11,** 175–191.

Banerjee, S., Wall, M. M., and Carlin, B. P. (2003). Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics* **4,** 123–142.

Boyle, D., Boyle, D., Olsen, V., Morgan, J., and Hyatt, A. (2004). Rapid quantitative detection of chytridiomycosis (Batrachochytrium dendrobatidis) in amphibian samples using real-time Taqman PCR assay. *Diseases of Aquatic Organisms* **60,** 141–148.

Brezger, A., Kneib, T., and Lang, S. (2005). BayesX: Analyzing Bayesian structured additive regression models. *Journal of Statistical Software* **14,** 1–22.

Cheng, T. L., Rovito, S. M., Wake, D. B., and Vredenburg, V. T. (2011). Coincident mass extirpation of neotropical amphibians with the emergence of the infectious fungal pathogen Batrachochytrium dendrobatidis. *Proceedings of the National Academy of Sciences* **108,** 9502–9507.

Chivian, E. and Bernstein, A. (2008). *Sustaining life: how human health depends on biodiversity.* Oxford University Press.

de Carvalho, V. I., Jara, A., Hanson, T., and de Carvalho, M. (2013). Bayesian nonparametric ROC regression modeling. *Bayesian Analysis* **8,** 623–646.

De Iorio, M., Johnson, W. O., Müller, P., and Rosner, G. L. (2009). Bayesian nonparametric nonproportional hazards survival modeling. *Biometrics* **65,** 762–771.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1,** 209–230.

Geisser, S. and Eddy, W. F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association* **74,** 153–160.

Gelfand, A. E. and Dey, D. K. (1994). Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society, Series B* **56,** 501–514.

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli* **7,** 223–242.

Hennerfeind, A., Brezger, A., and Fahrmeir, L. (2006). Geoadditive survival models. *Journal of the American Statistical Association* **101,** 1065–1075.

Ishwaran, H. and James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association* **96,** 161–173.

Ishwaran, H. and Zarepour, M. (2000). Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika* **87,** 371–390.

Jara, A., Hanson, T., Quintana, F., Müller, P., and Rosner, G. (2011). DPpackage: Bayesian semi- and nonparametric modeling in R. *Journal of Statistical Software* **40,** 1.

Jara, A., Lesaffre, E., De Iorio, M., and Quitana, F. (2010). Bayesian semiparametric inference for multivariate doubly-interval-censored data. *The Annals of Applied Statistics* **4,** 2126–2149.

Knapp, R. A., Matthews, K. R., Preisler, H. K., and Jellison, R. (2003). Developing probabilistic models to predict amphibian site occupancy in a patchy landscape. *Ecological Applications* **13,** 1069–1082.

Knapp, R. A. and Morgan, J. A. (2006). Tadpole mouthpart depigmentation as an accurate indicator of chytridiomycosis, an emerging disease of amphibians. *Copeia* **2006,** 188–197.

Li, Y. and Lin, X. (2006). Semiparametric normal transformation models for spatially correlated survival data. *Journal of the American Statistical Association* **101,** 591–603.

Lips, K. R., Diffendorfer, J., Mendelson, J. R., and Sears, M. W. (2008). Riding the wave: reconciling the roles of disease and climate change in amphibian declines. *PLoS Biology* **6,** e72.

Papaspiliopoulos, O. and Roberts, G. O. (2008). Retrospective Markov chain Monte Carlo methods for Dirichlet process hierarchical models. *Biometrika* **95,** 169–186.

Pati, D., Dunson, D. B., and Tokdar, S. T. (2013). Posterior consistency in conditional distribution estimation. *Journal of Multivariate Analysis* **116,** 456–472.

Rachowicz, L. J., Knapp, R. A., Morgan, J. A., Stice, M. J., Vredenburg, V. T., Parker, J. M., and Briggs, C. J. (2006). Emerging infectious disease as a proximate cause of amphibian mass mortality. *Ecology* **87,** 1671–1683.

Rodríguez, A., Dunson, D. B., and Gelfand, A. E. (2010). Latent stick-breaking processes. *Journal of the American Statistical Association* **105,** 647–659.

Sang, H. and Huang, J. Z. (2012). A full scale approximation of covariance functions for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **74,** 111–132.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica* **4,** 639–650.

Skerratt, L. F., Berger, L., Speare, R., Cashins, S., McDonald, K. R., Phillott, A. D., Hines,

H. B., and Kenyon, N. (2007). Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth* **4,** 125–134.

Song, P. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics* **27,** 305–320.

Tierney, L. and Mira, A. (1999). Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine* **18,** 2507–2515.

Vredenburg, V., Bingham, R., Knapp, R., Morgan, J., Moritz, C., and Wake, D. (2007). Concordant molecular and phenotypic data delineate new taxonomy and conservation priorities for the endangered mountain yellow-legged frog. *Journal of Zoology* **271,** 361–374.

Vredenburg, V. T., Knapp, R. A., Tunstall, T. S., and Briggs, C. J. (2010). Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proceedings of the National Academy of Sciences* **107,** 9689–9694.

Wake, D. B. and Vredenburg, V. T. (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences* **105,** 11466–11473.

Waller, L. A. and Gotway, C. A. (2004). *Applied spatial statistics for public health data.* John Wiley & Sons.

Wang, S., Zhang, J., and Lawson, A. B. (2012). A Bayesian normal mixture accelerated failure time spatial model and its application to prostate cancer. *Statistical Methods in Medical Research* .

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association* **99,** 250–261.

Zhao, L., Hanson, T. E., and Carlin, B. P. (2009). Mixtures of Polya trees for flexible spatial frailty survival modelling. *Biometrika* **96,** 263–276.

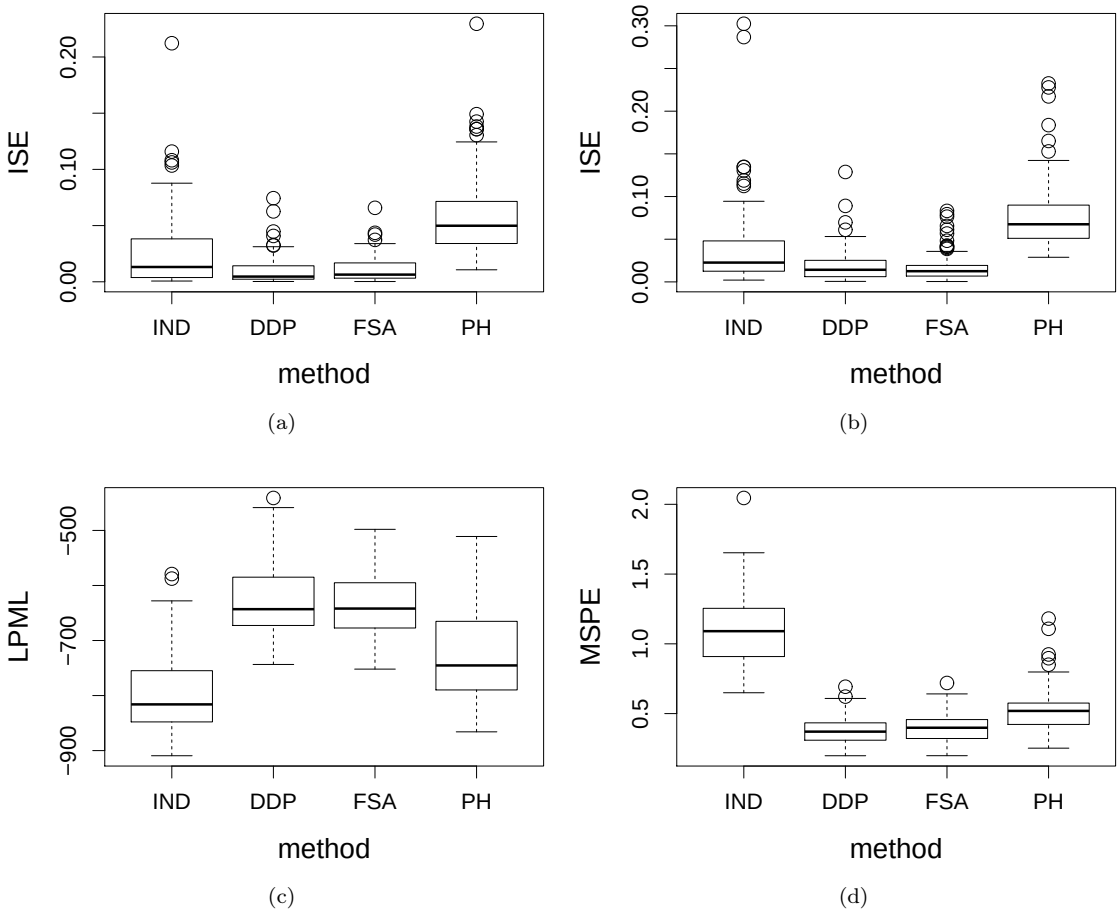
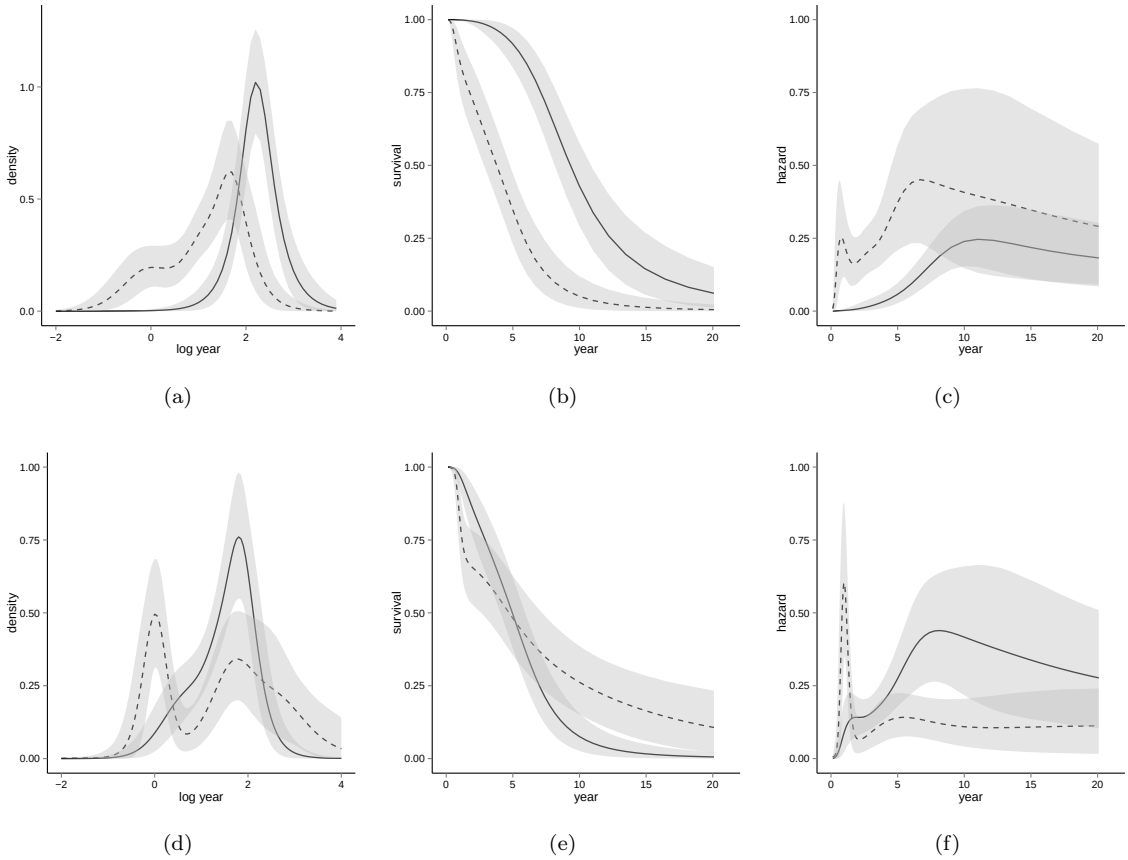


Figure 1. Simulated data – Scenario I. Panel (a) and (b): boxplots of ISEs for fitted survival curves when $x = -1$ and $x = 1$, respectively. Panel (c): boxplots of LPMLs. Panel (d): boxplots of MSPE. In each panel, the four models from left to right are LDDPM-ind, LDDPM-spatial, LDDPM-spatial-FSA, and PH-spatial, respectively.



(a) (b) (c)

(d) (e) (f)

Figure 2. Frog data. Fitted marginal densities (panels (a) and (d)), survival curves (panels (b) and (e)) and hazard curves (panels (c) and (f)) with 90% point-wise credible intervals for high versus low value of bddist when bdwater is equal to 0 (panels (a), (b) and (c)) and for bdwater=0 versus bdwater=1 when bddist is equal to its population mean of 2.717-km (panels (d), (e) and (f)). In panels (a), (b) and (c), the results for bddist=95% and bddist=5% quantiles are displayed as solid and dashed lines, respectively. In panels (d), (e) and (f), the results for bdwater=0 and bdwater=1 are displayed as solid and dashed lines, respectively.

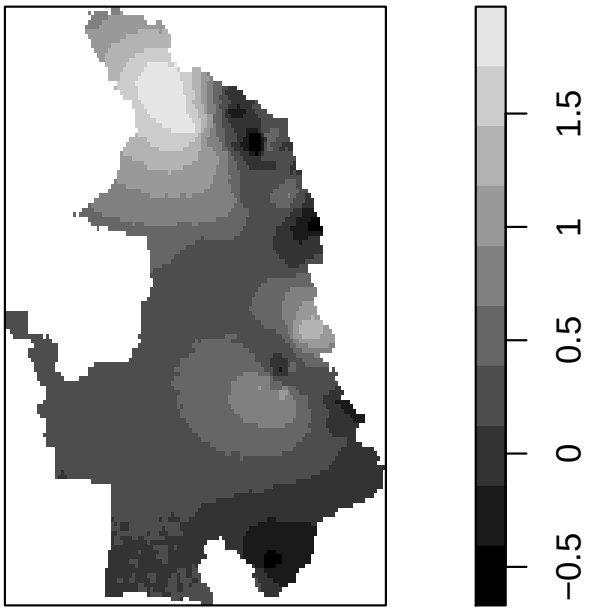


Figure 3. Frog data. Spatial map for the predicted residual process $z(\mathbf{s})$ at $10,000$ randomly simulated new locations. Higher value implies better survival overall.

Table 1

Simulated data – Scenario I. True value, bias of the point estimator (posterior mean), mean (across Monte Carlo replicates) of the posterior standard deviations (MEAN-SD), standard deviation (across Monte Carlo replicates) of the point estimator (SD-MEAN), and Monte Carlo coverage probability for the 95% *credible interval (CP) for the spatial correlation parameter* $\boldsymbol{\theta}$*. The averaged computing time is also presented.*

Model	Parameters	True	BIAS	MEAN-SD	SD-MEAN	CP
LDDPM-spatial	θ_1	0.98	-0.026	0.028	0.028	0.91
(32 minutes)	θ_2	0.10	0.017	0.026	0.025	0.95
LDDPM-spatial-FSA	θ_1	0.98	-0.017	0.025	0.024	0.96
(24 minutes)	θ_2	0.10	0.016	0.025	0.023	0.93
PH-spatial	θ_1	0.98	-0.065	0.037	0.047	0.61
(24 minutes)	θ_2	0.10	-0.009	0.023	0.030	0.78

Table 2
Frog data. Summary for event times, censoring status, bdwater, and bddist.

Time to Bd (yrs)	Bd status	bdwater	bddist (km)
5 (median)	33 (censored)	57 (bdwater=1)	1.811 (median)
1-11 (range)	276 (event)	252 (bdwater=0)	0.092-9.189 (range)

Table 3
Frog data. Posterior statistics for θ_1 and θ_2 under the LDDPM-spatial model assuming the exponential correlation function. The computing time is also presented.

Model	Parameters	Mean	Median	Std. dev.	95% CI
LDDPM-spatial	θ_1	0.991	0.992	0.004	(0.982, 0.998)
(3.2 hours)	θ_2	0.133	0.130	0.040	(0.060, 0.216)