Partitioned and Non-Partitioned Regularized Additive Hazard Models with and Without Spatial Dependence

Peter Enesi Omaku[1*] & Benjamin Agboola Oyejola[2]

[1]*Department of Mathematics & Statistics, Federal Polytechnic Nasarawa – Nigeria*

[2]*Department of Statistics, University of Ilorin, Ilorin- Nigeria*

Corresponding author: email: omakupete@gmail.com

Abstract

Spatial effects are often simultaneously investigated with non-linear effects of continuous covariates and the usual linear effect. In this work the performance of models with and without spatial dependence in partitioned (PM) and non-partitioned models (NPM) for four (4) censoring percentages, three(3) levels of Weibull baseline variances (WBV), and sample sizes 100, 500 & 1000 were investigated. Hazard models were adapted to the generalized additive predictors and analyses were carried out via MCMC simulation technique. The performances of the models were again assessed when fitted to the diabetic data set. Results suggest that; partition models outperformed the non-partition ones. Models with spatial dependence perform better than models without spatial dependence in denser event times and when WBVs are low. The partition models perform better with spatial dependence than the Non-partitioned models. For the diabetic data set, it is seen that covariates Age and Blood Sugar level (BSL) violates the proportionality assumptions upon test. Further assessment from the graph of coefficient against time; suggest that Age be put to cut-points while BSL was estimated for models with and without Penalized splines for the sake of comparison, since the graph shows just a slight deviation from proportionality. Hazard rates for the time varying Age; indicate that as the time of study rolls by, the hazard of experiencing the event death from the disease increases steadily between intervals but constant within each time interval. A unit change in hazard rate for BSL indicates a decrease for PM implemented for with and without penalized splines. The model without penalized splines was however, seen to be better with smaller DIC (Deviance Information Criteria) value. Marriage is seen to be significant in the management of the disease in comparison to single patients. In addition patients are advised to visit their physicians on a regular basis to run a routine check to keep their BSL in good range. The study provides a means of moving out of non-linear ruts in survival data analysis. Intervals increase sample sizes (pseudo-observations), which in turn improves the modified Partitioned model when they are with or without spatial dependence.

Keywords: Generalized Additive Model (GAM), Modification, Non-partition, Partition, Spatial dependence.

1. Introduction

Analysis of survival time data has gained considerable attention, particularly in medicine, where the conventional denotation 'Survival analysis' arises from (Hennerfeind, 2006). In several other bio-statistical applications on censored follow-up time data, interest lies mainly on the prognostic role of clinical/biological covariates. To such end, non-parametric and semi-parametric methods have been preferred over parametric ones. The most widely adopted tool is the Cox model, which avoids any assumption of the functional form of the hazard function on time. However, such feature is not tenable if the interest lies on investigating the shape of the hazard or in predictive modeling (Kooperberg et al 1995) when the Cox-model is extended to time-varying covariates and time-dependent effects, which combine to give the most general version of the hazard. Functional forms of continuous covariates and spatial correlation as a form of frailty model in medical research may most often affect the overall estimate of the hazard function. Progress would require specifying the form of this function of time. In such scenario where time is seen to be truly continuous a flexible or semi-parametric strategy is required, where mild assumptions are made about the baseline hazard. Specifically, time may be subdivide into reasonably small intervals and assume that the baseline hazard is constant in each interval, leading to a piece-wise survival model (cited in Omaku & Ibinayin, 2020). What modifications are possible to fit models that are interpretable in the face of these complexities of non-linear ruts in survival modeling?

The Partitioned model (Piecewise models) is a special case of models that employ time-varying covariates (Rodriguez, 2007). This is because the Partition Model (PM) requires us to split up single-spell duration data in the same way that we had to when we wanted to incorporate time-varying covariates. According to (Fabio et al. 2010) the PM arises as a quite attractive alternative to parametric models for the analysis of time to event data. Although parametric in a strict sense, the PM can be thought of as a nonparametric model as far as it does not have a closed form for the hazard function, it is also a proportional hazard (PH) model as its basic hazard rate can be specified in the following way: $\lambda(t, X) = \lambda_0(t)e^{X\vartheta}$. The main difference is that the baseline hazard rate is allowed to vary in different time periods.

In the simplest case of proportional hazards, the relationship between the hazard function of the PCE model and the covariates is expressed as follows:

$$\lambda_{PE}(t/X_i) = \sum_{k=1}^{K}(I(t\epsilon T_k)\lambda_k)\exp(X_i^T\underline{\vartheta}) \quad (1)$$

where: $\lambda_1, \ldots \lambda_K$ are the unknown parameters of the baseline hazard function; $\underline{\vartheta}$ is the p × 1 vector of regression coefficients; and $I(t \in T_k)$ is an indicator function being equal to 1 if $\tau_k - 1 \leq t < \tau_k$ and 0 otherwise Cited in (Marano et al, 2016).

The model in (1) can be completely specified as

$$\lambda(t,\ X_t) = \begin{cases} \bar{\lambda}_0(t_1)e^{\ X_1\vartheta} & t\epsilon(0,\tau_1\ ) \\ \bar{\lambda}_0(t_2)e^{\ X_2\vartheta} & t\epsilon(\tau_1\ ,\tau_2) \\ \qquad\vdots & \\ \bar{\lambda}_0(t_k)e^{\ X_k\vartheta} & t\epsilon(\tau_{k-1}\ ,\tau_k) \end{cases} \tag{2}$$

$$\lambda(t,\ X_t) = \begin{cases} \exp[\log(\bar{\lambda}_0(t_1) + X_1\vartheta] & t\epsilon(0,\tau_1\ ) \\ \exp[\log(\bar{\lambda}_0(t_2) + X_2\vartheta] & t\epsilon(\tau_1\ ,\tau_2) \\ \qquad\vdots & \\ \exp[\log(\bar{\lambda}_0(t_k) + X_k\vartheta] & t\epsilon(\tau_{k-1}\ ,\tau_k) \end{cases} \tag{3}$$

It is seen in (Clark & Ryan, 2002, Bastos & Gamerman, 2006 & Kim et al. 2006) that Partitioning is done by increasing the size of the dataset, from the original one, to such end, the data of each subject is replicated for each interval in which it is followed. For each replicate, a row data indicating subject covariates, time interval and the status of the subject within the interval is created for intervals say k=1,...,K. The essence of following subjects within intervals is to treat the dataset of wriggles or non-proportionality which makes interpretation sometimes impossible and difficult.

Consider partitioning duration into J intervals with cut points $0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k = \infty$. Define the k[th] interval as $I_i = [\ \tau_{k-1}\ , \tau_k)$, extending from the $[k-1)$st boundary to the k[th] and including the former but not the latter, then assume that the baseline hazard is constant within each interval, so that

$$\lambda_0(t) = \lambda_i\ \ for\ \ t\ in\ \ [\tau_{k-1},\ \tau_k) \tag{4}$$

Model the baseline hazard $\lambda_0(t)$ using k parameters $\lambda_1, \lambda_2 \ldots \lambda_k$, each representing the risk for the reference group (or individual) in one particular interval. Since the risk is assumed to be piece-wise constant, the corresponding survival function is often called a piece-wise exponential (Marano et al, 2016).

Clearly, judicious choice of the cut points should approximate reasonably well almost any baseline hazard, using closely-spaced boundaries the hazard varies rapidly and wider intervals where the hazard changes more slowly.

1.1 Spatial Dependence

Spatial survival analysis has received a great deal of attention, due to the important role that geographical information can play in survival prediction, serving as a proxy for unmeasured and structured regional characteristics such as socioeconomic status, access to health care, pollution, etc. Literature on the spatial analysis of survival data has flourished over the last two decades, including the study of leukemia survival (Henderson et al, 2002), a geo-additive discrete time probit model by (Adebayo & Farmeir ,2003) to model regional and social economic variations of childhood mortality in Nigeria, childhood mortality (Kneib, 2006), asthma (Li & Lin, 2006), breast cancer (Banerjee and

Dey, 2005);(Zhou, Hanson, Jara, & Zhang, 2015a), political event processes (Darmofal, 2009), prostate cancer (Wang, Zhang, & Lawson 2012; Zhou, Hanson & Zhang 2017), pine trees (Li, Hong, Thapa, & Burkhart 2015a), threatened frogs (Zhou, Hanson, & Knapp 2015b), health and pharmaceutical firms (Arbia et al, 2016), emergency service response times (Taylor, 2017), and many others. Continous time geoadditive hazard rate model was proposed by (Hennefeind et al, 2006). Within a unified frame work, their approach extends the Cox model with respect to several aspects often needed in applications. They generalized the common linear predictor to an additive predictor, including unknown functional forms for the baseline hazard, time varying effects of metrical covariates and spatial component for spatial effects

1.2 Generalized Additive Models

Generalized additive models (GAMs) are statistical models in which the conventional multiple linear regression is generalized to permit a much broader class of time varying and nonlinear functional form of continuous covariates and their effects; but still with additive relationships between response and predictor variables. GAMs, derived from the work of (Hastie & Tibshirani, 1986, 1990), provide flexible and effective means of moving out of the "linear rut" (Jones & Almond, 1992) in which a considerable amount of bio-statistical modeling is still located. In this work GAM is used to estimate the effects of several functional forms of covariates incorporated to the Modified partitioned model.

In this study a modified Partitioned (PM) and Non-Partitioned (NPM) survival models with additive predictors to include the functional forms; of time varying covariates, nonlinear effects of metrical covariates and spatial terms for spatial dependence were considered. The performance of PM & NPM for with and without spatial dependence were assessed using the DIC (Deviance Information criterion).

2. Materials and Method

The risk data used for this paper was simulated from three levels of Weibull baseline hazard distribution, for sample sizes of 100, 500 & 1000 and four (4) censoring percentages of: no censoring "0%", low "about 25%", moderate "about 50%" and high "about 75%". Cut off points were observed from the graphs of coefficients against time.

Application to Diabetic Dataset

Diabetic dataset for 452 patients who were admitted at the Nigerian Air Force Hospital, Abuja. Time from diagnosis of the disease for in-patients to death defines the failure time while those whose records read "alive" were right-censored because such patients had not died as at the time of the study. Other variables considered were; Age, Gender, Marital status and Blood Sugar Level (BSL).

2.1 Model Specification

The Cox hazard model with Spatial effect as stated in (Zhou & Hanson, 2017) is as follows

$$S_{x_{ij}}(t) = S_0(t)^{e^{x_{ij}^T \vartheta + v_i}}, \quad (5)$$

$$f_{x_{ij}}(t) = e^{x_{ij}^T \vartheta + v_i} S_0(t)^{e^{x_{ij}^T \vartheta + v_i} - 1} f_0(t) \quad (6)$$

$$h_{x_{ij}}(t) = e^{x_{ij}^T \vartheta + v_i} . f_0(t) \quad (7)$$

where $S_{xij}(t)$ is the survival function of subjects i with x_j covariate, $f_{x_{ij}}(t)$ is the density function of subjects i with x_j covariate, $h_{x_{ij}}$ is the hazard function of subjects i with x_j covariate, $\vartheta = (\vartheta_1, \ldots, \vartheta_p)^T$ is a vector of regression coefficients, v_i is an unobserved spatial frailty associated with the survival function, and S_0 (t) is the baseline survival with density $f_0(t)$ corresponding to x_{ij} = 0 and v_i = 0.

2.2 Hennerfeind suggested the repraramетization of the Cox model (a Non Partition Model-NPM) as follows:

$$\eta(t; z, w, v, x) = f_0(t) + \sum_{j=1}^{p} f_j(t) z_j + \sum_{j=p+1}^{p+q} f_j(w_j) + f_{spat}(v_j) + x'\gamma \quad (8)$$

The function $f_0(t)$ is the baseline effect, and a function, $f_j(t)$represents a time–varying effect of the covariate z_j. The functions $f_j(w_1)$,..., $f_q(w_q)$ are possibly nonlinear effects of metrical covariates $w_1, \ldots, w_q$ and $f_{spat}(v_i)$ is a structured spatial effect, where v, v =1,. . .,V is either a spatial index, with $v = v_j$ if subject j is from area v or it is an exact spatial coordinate $v = (x_j, y_v)$, e.g. for centriods of regions or if exact locations of individuals are known. γ is the usual linear part of the predictor.

2.3 Expression (4) may be modified, when partitioning is done, as:

$$\lambda_{PM}(t_\omega; z, w, v) = \{I(t \epsilon T_k(f_k(t_\omega))\} + \sum_{j=1}^{p} f_j(t_\omega) z_{jk} + \sum_{j=p+1}^{p+q} f_{jk}(w_{jk}) + f_{spat}(v_{jk}) \quad (9)$$

With its various terms defined as t_ω is the time variable varied for three (3) levels of Weibull distribution, the function $f_k = log\lambda_k$ is the baseline effect varied for three (3) variance levels of Weibull distribution for the kth interval of PM, $f_0(t)$ the baseline effect for the NPM, functions $f_j(t)z_{jk}$,..., $f_p(t)z_{pk}$ are functional forms of time varying covariates $z_{1k}, \ldots, z_{pk}$ in the k^{th} interval , The functions $f_j(w_{jk})$,..., $f_q(w_{qk})$ are possibly nonlinear effects of metrical covariates $w_{1k}, \ldots, w_{qk}$ and $f_{spat}(v_{jk})$ is a structured spatial effect, where v, v =1, . . . ,V is either a spatial index, with $v = v_{jk}$ if subject j in the k^{th} interval is from area v or it is an exact spatial coordinate $v = (x_i, y_v)$, e.g. for centriods of regions or if exact locations of individuals are known.

2.4 Model Likelihood for the modified partitioned model

$$L_{PM}\left(\underline{\lambda},\underline{\gamma},\underline{\beta};D,\Delta,z,w,v\right)=\prod_{j=1}^{n}\prod_{k=1}^{K_j}(\lambda_k\exp(Z_k\gamma_{jk}+\beta_k w_{jk}+v_{jk})^{d_{ik}}.\exp(\lambda_k\exp\left(Z_k\gamma_{jk}+\beta_k w_{jk}+v_{jk}\right)\Delta_{jk}). \quad (10)$$

where for each subject j there is a product of Kj terms, K_j being the number of intervals in which the subject is followed. In the expression above, d_{jk} is the status of the j^{th} subject within the interval T_k (0 = alive or censored, 1 = failed); Δ_{jk} is the time spent in T_k by the subject. From expression (10) it may be seen that L_{PM} is proportional to the product of Poisson likelihoods for Dj_k with mean parameters: $\mu_{jk}=\lambda_k\exp\left(Z_k\gamma_{jk}+\beta_k w_{jk}+v_{jk}\right)\Delta_{jk}$. As a consequence, the expression of the Poisson regression model is:

$$D_{jk}\sim POISSON(\mu_{jk});\log(\mu_{jk})=\underline{\alpha_k}+Z_k\gamma_{jk}+\beta_k w_{jk}+v_{jk}+\log(\Delta_{jk})\ . \quad (11)$$

where $h(j)$ indicate the interval where t_j falls, i.e. the interval where individual j died or was censored, $\underline{\alpha_k}$ = $\log(\lambda_k)$ are log-hazard parameters, and the term $\log(\Delta_{jk})$ is an offset .

2.5 Partition model with regularized effects:

$$\begin{cases} d_{jk}\sim POISSON(\mu_{jk}) \\ \log(\mu_{jk})=\underline{B_0^T\alpha}+\sum_{j=1}^{p}Z_{jk,}\underline{B_0^T\gamma_j}+w_{k,j}\underline{B_0^T\beta_k}+v_{jk}+\log(\Delta_{jk}) \\ (\underline{\alpha}|\tau^2)\sim RW(\tau^2,P_d);\ \tau^2\sim\pi_{\tau^2} \\ \left(\underline{\gamma_j}\middle|\tau_j^2\right)\sim RW\left(\tau_j^2,P_d^{(j)}\right);\ \tau_j^2\sim\pi_{\tau^2 j};j=1,\dots,p \\ \left(\underline{\beta_j}\middle|\tau_j^2\right)\sim RW\left(\tau_{p+l}^2,P_d^{(p+l)}\right);\ \tau_{p+l}^2\sim\pi_{\tau^2 j};l=1,\dots,q \\ V_i/\{v_j\}_{j\neq i}\sim N\left(-\sum_{\{j:j\neq i\}}P_{ijV_{ij}}/P_{ii}\ ,\tau^2/P_{ii}\right) \end{cases} \quad (12)$$

d_{jh} is the status of the subject j in the k^{th}, interval, μ_{jk} is the mean likelihood of the partitioned model of subject j in the k^{th} interval, The time-dependent effects for each covariate are: $Z_j\underline{B_0^T\gamma_{1j}}$; j = 1, …,p. Thus, for each Z_j, its values multiplied for a piecewise constant function: $\underline{B_0^T\gamma_j}$; in the parameters. $\gamma_{1j}=(\gamma_{1,j,1},\dots,\gamma_{1,j,K})^T$. This enables the effect of each Z_j to vary in each interval T_k of the original partition of the follow-up: $\underline{B_0^T\alpha}+Z_j\underline{B_0^T}=\alpha_h+z_j\gamma_{j,h}$ for $t\in T_h$.

τ^2 is the smoothing parameter, P_d is the penalized term of order d for the random walk process of non-linear ruts (the log baseline hazard, the functional form of time varying covariate & the functional form of continuous covariate). π_β and π_{τ^2} are generic prior densities for the regression coefficients (Omaku & Oyejola , 2020).

2.5.1 Bayesian P-Splines

The Bayesian P-splines method is based on a hierarchical model with non-informative priors for the regression coefficients (β) and a Gaussian Random Walk (RW) prior of order d for the coefficients of the hazard function (B-spline), conditional to a smoothing parameter τ^2 The general expression of the RW prior as suggested by (Lang & Brezger, 2004) and (Kooperberg & Intrator, 1995) is the following:

$$\beta_j/\tau_j^2 \propto \exp(-\frac{1}{2\tau_j^2}\ \lambda_j\ \beta'_j K_j \beta_j) \quad (13)$$

The penalty matrix K_j is of the form $K_j = D'D$, where D is a first or second order difference matrix. For an independent and identical random effect, the penalty matrix is the identity matrix, i.e. K_j= I. The variance parameter τ_j^2 controls the trade off between flexibility and smoothing and an inverse gamma prior (the conjugate prior) is assumed. i.e. $\tau_j^2 \sim IG(a, b)$.

2.5.2 Gaussian Random Field (GRF) priors

For spatial data, it is assumed in (Zhou & Hanson, 2017) that $v_i = v(s_i)$ comes from a Gaussian random field (GRF) $\{v(s), s\epsilon S\}$ such that $v = (v_1, \dots, v_m)$ follows a multivariate Gaussian distribution as $v \sim N_m(0, \tau^2 R)$, where τ^2 measures the amount of geographical variation across locations and the (i,j) element of R is modeled as $R[i,j] = \rho(s_i, s_j)$. Here $\rho(.,.)$ is a correlation function controlling the spatial dependence of v(s). In "survregbayes" package in R, the powered exponential correlation function $\rho(s_i, s_j)$. = $\rho(s_i, s_j, \varphi)$.= $\exp\{-(\varphi\|s - s'\|)^v\}$ is used, where φ> 0 is a range parameter controlling the spatial decay over distance, $v\epsilon$ (0,2] is a prespecified shape parameter, and $\|s - s'\|$ refers to the distance (e.g., Euclidean, great-circle) between s and s' Therefore, the prior GRF($\tau^2, \emptyset$) is defined as $V_i/\{v_j\}_{j\neq i} \sim N\left(-\sum_{\{j:j\neq i\}} P_{ijV_{ij}}/P_{ii}, \tau^2/P_{ii}\right) i = 1, \dots, m$ where P_{ij} is the (i,j) element of R^{-1} .

2.6 Variable Transformation for Diabetic dataset

For the sake of analysis the covariate "age" is considered as a continuous variable, Gender was coded "1" for "male" patients and "0" for "female" patients. Marital status was coded "0" for "single" patients and "1" for patients that are "married". Blood Sugar Level (BSL) was considered to be continuous in nature. For the outcome variable patients that experienced the event death were coded "1" while those that were alive as at the time of study or those that were discharged were considered to be right censored and coded "0".

2.6.1 Models for Diabetic Dataset

$$\eta_{NPM} = f_0(t) + AGE(t)\delta + \beta.BSL + \gamma_1 Gender + \gamma_2 Marital\ status \quad (14)$$

$$\log(\mu_{jh}, \alpha, \beta, \gamma_1, \gamma_2) = \underline{B_0^T \alpha} + AGE(t)\delta + \beta.BSL + \gamma_1 Gender + \gamma_2 Marital\ status \quad (15)$$

$$\log(\mu_{jh}, \alpha, \beta_i, \gamma_1, \gamma_2) = \underline{B_0^T \alpha} + AGE(t)\delta + \beta_{\mathrm{i(pslines)}} BSL + \gamma_1 Gender + \gamma_2 Marital\ status \quad (16)$$

Equation 15 differs from 16 as the latter estimates the effect of BSL (β_i) via Penalty splines through some numbers of selected knots. Simulations and analysis were carried out in R using the coda package for spBayesSurv, version 3.6.2. Comparisons were done using Deviance Information Criterion (DIC) (smaller is better).

2.7 Test for Non-Proportionality

To test the hypothesis that the proportional hazard assumption is valid, the following statement of hypothesis is made.

H_0: $\delta_1 = \delta_2 = \cdots = \delta_p = 0$ (Assumption is valid)

H_1: $at\ least\ one\ of\ the\ \delta_i's\ is\ not\ equal\ to\ zero$ (Assumption violated)

Decision rule: Reject H_0 if $p - value \leq \alpha$ (level of significance)

Residual measures are used to investigate the departure from the proportional hazard assumption. Schoenfeld residuals are usually calculated at every failure of time under the proportional hazard assumption, and usually not defined for censored observations. The overall significance test is called the global test of the model (cited in Adeniyi & Akinrefon, 2018).

2.8 Simulation Study

The survival function is given by

$$S(t) = exp(-H_0(t)e^{\beta(t)z_j + w_j + v_j}) \quad (17)$$

Then, the cumulative distribution function of the non-proportional hazards model is

$$F(t) = 1 - exp(-H_0(t)e^{\beta(t)z_j + w_j + v_j}) \quad (18)$$

The distribution function follows a uniform distribution on the interval from 0 to 1, denoted by U if $U \sim U(0,1)$ then $1 - U \sim U(0,1)$. Therefore, the survival function follows a uniform distribution on the interval from 0 to 1. That is

$$U = exp\left(-H_0(t)e^{\beta(t)z_j + w_j + v_j}\right) \sim U(0,1) \quad (19)$$

The failure time t can be solved by inverting H_0

$$t = H_0^{-1}\left(\frac{-logU}{e^{\beta(t)z_j + w_j + v_j}}\right) \quad (20)$$

For Weibull baseline distribution, then $H_0^{-1}(\quad) = \alpha t^{1/\eta}$ and the failure time is

$$t = \alpha\left(\frac{-logU}{e^{\beta(t)z_j + w_j + v_j}}\right)^{1/\eta} \quad (21)$$

The simulations apply the functional form of time varying covariate by (Bender et al , 2005) given as

$$f(z) = 0.5\sqrt{t} * y. \qquad y \sim binom(N, 1, 0.5) \tag{22}$$

The functional form of the continuous covariates as in (Brezger, 2004) is given as:

$$f(w_j) = 1.0/1.80 w_j \tag{23}$$

where $w_i \sim U(-3,3)$.

Co-ordinates for spatial correlations follow the uniform distribution. $v_1 = runif(N, 0, 40)$ and $v_2 = runif(N, 0, 100)$. Ulviya (2013), obtained the shape and scale parameters of the Weibull distribution from the formulas below

$$\eta = \frac{1}{\Gamma(1+\frac{1}{\alpha})} \tag{24}$$

and

$$\left(\frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{(\Gamma(1+\frac{1}{\alpha}))^2} - 1\right) = 0.5 \tag{25}$$

for a convenience choice of mean 1 and variance 0.5. Using the uniroot function in R. parameters were given to be approximately $\alpha = 1.435523$ and $\eta = 1.101321$. We considered studying the impact of increasing and decreasing the variance of the Weibull distribution while keeping the mean at 1. The result is displayed in table 1 below

Table 1: Shape and scale parameters of the Weibull distributions

E(T)	Var (T)	α	η
1	0.25	2.101377	1.129063
1	0.5	1.435523	1.101321
1	0.75	1.157975	1.052847

3. Results and Discussion

Table 2: DIC values for levels of censoring, three (3) levels of Weibull baseline variance, sample sizes for with and without spatial dependence in Non-Partition Model

Censoring %	Variance	N=100		N=500		N=1000	
		WSD	**WOSD**	**WSD**	**WOSD**	**WSD**	**WOSD**
0	**0.25**	853.6004	854.1665	4068.979	4069.006	8779.792	8780.939
	0.50	962.164	962.253	4707.551	4709.29	10137.67	10143.15
	0.75	1032.945	1033.307	5138.187	5136.902	11028.65	11035.66
25	**0.25**	653.4635	653.4906	3210.955	3246.832	7007.161	7001.295
	0.50	716.3681	716.0612	3623.79	3578.655	7820.168	7828.612
	0.75	756.2604	755.0077	3848.511	3813.489	8272.722	8341.513
50	**0.25**	512.0976	498.3278	2407.989	2445.837	5000.534	5034.049
	0.50	520.4511	519.2795	2620.238	2606.079	5512.023	5520.44
	0.75	543.8143	543.4234	2607.02	2601.271	5698.33	5692.22
75	**0.25**	255.7324	255.4689	1280.065	1263.73	2695.128	2725.757
	0.50	312.231	488.3278	1338.623	1298.471	2850.098	2833.02
	0.75	269.8102	269.1235	1350.597	1340.108	2831.845	2807.207

From table 2: models for 0% censoring, were better when represented With Spatial Dependence (WSD) than models Without Spatial Dependence at low and intermediate Weibull Baseline Variances (WBV) for all sample sizes. At high level of WBV, the Non-Partition Model without Spatial Dependence (NPMWOSD) was only seen to be better when sample size is 500.

For 25% censoring, it is seen that models WSD are better for sample sizes 100 and 500, at low Weibull Baseline Variance. Model without Spatial Dependence (WOSD) outperformed models with spatial dependence at intermediate & high levels of Weibull Baseline Variances. Again, when sample size is increased to 1000, it is observed that models WSD at intermediate and high WBV are better.

For 50% censoring, it is seen observed that when sample size is 100, models Without Spatial Dependence (WOSD) outperformed on comparison to model incorporated WSD for all levels of WBV.

Increment in sample sizes impact the models; at low WBV, when n=500 & 1000, models WSD were observed to be better on comparison to models WOSD. Model at intermediate level of WBV is seen to be better With Spatial Dependence, when sample size is 1000, while models for all sample sizes at high level of WBV were better when modeled WOSD

For 75% censoring, models WOSD outperformed in most combination of sample size and levels of WBV. At low level of WBV, NPMWOSD had better DIC values to NPMWSD for sample sizes 100 & 500 but not the case, when sample size is increased to 1000. At intermediate level of WBV it is seen

that NPMWOSD performed better than NPMWSD for sample sizes 500 and 1000. While at high level of Weibull Baseline Variance, NPMWOSD were better for all sample sizes.

Table 3: DIC values for levels of censoring, three (3) levels of Weibull baseline variance, sample sizes for with and without spatial dependence in "Partitioned Model" (PM)

Censoring %	Variance	N=100		N=500		N=1000	
		WSD	**WOSD**	**WSD**	**WOSD**	**WSD**	**WOSD**
0	**0.25**	832.1385	837.187	3999.343	4945.71	8597.94	8878.76
	0.50	943.7934	**936.7658**	4686.069	**4143.179**	10085.62	10119.98
	0.75	1005.326	1014.516	5076.758	**5072.229**	10970.23	**10961**
25	**0.25**	644.8024	648.6781	3181.119	3233.434	6939.78	6997.888
	0.50	702.3201	708.4976	3613.921	**3538.044**	7813.243	**7808.675**
	0.75	746.4749	753.0131	3829.991	**3797.836**	8263.318	8286.271
50	**0.25**	488.0027	**468.4429**	2356.165	2423.315	4942.05	5022.247
	0.50	515.8235	517.3664	2604.272	**2556.05**	5514.923	**5508.376**
	0.75	540.5304	**527.3858**	2552.504	**2548.174**	5707.325	**5655.104**
75	**0.25**	287.7841	**266.2221**	1275.756	**1237.466**	2677.361	2706.389
	0.50	299.871	461.421	1350.044	**1289.735**	2848.829	**2706.389**
	0.75	267.6131	**261.0132**	1353.066	**1336.481**	2841.659	**2773.892**

For 0% censoring, models WSD are better than WOSD for all sample sizes when level of Weibull Baseline Variance (WBV) is low, increment in WBV to intermediate presents a better DIC values for models WOSD when sample sizes are 100 and 500 but better for PMWSD when the sample size is increased to 1000. At high level of WBV, PMWSD was better for sample sizes 100 and 500 but was outperformed by PPMWOSD when sample size is 1000.

When censoring is 25%, it is noticed that models WSD are better for all sample sizes at low level of WBV and for all levels of WBV for sample size 100. Models WOSD perform better for sample sizes 500 & 1000 at intermediate WBV and for sample size 500 for high WBV.

When censoring is 50%, it is observed that, models WSD are better at low level of WBV for sample sizes 500 and 1000, models WOSD were observed to be better than Models WSD for sample sizes 500 and 1000 at both intermediate and high WBVs.

When censoring is 75%, at low level of WBV, the models WOSD out performed models WSD when sample sizes are 100 and 500, but better for model WSD when sample size is increased to 1000. At intermediate level of WBV for sample sizes 500 and 1000 PMWOSD are better in comparison to PMWSD. At high level of WBV, PMWOSD outperform PMWSD for all sample sizes.

Table 4: DIC values for levels of censoring, three (3) levels of Weibull baseline variance, sample sizes for Non-Partitioned Model (PM) and Partitioned Model with spatial dependence

Censoring %	Variance	N=100		N=500		N=1000	
		NPM	**PM**	**NPM**	**PM**	**NPM**	**PM**
0	**0.25**	853.6004	**832.1385**	4068.979	**3999.343**	8779.792	**8597.94**
	0.50	962.164	**943.7934**	4707.551	**4686.069**	10137.67	**10085.62**
	0.75	1032.945	**1005.326**	5138.187	**5076.758**	11028.65	**10970.23**
25	**0.25**	653.4635	**644.8024**	3210.955	**3181.119**	7007.161	**6939.78**
	0.50	716.3681	**702.3201**	3623.79	**3613.921**	7820.168	**7813.243**
	0.75	756.2604	**746.4749**	3848.511	**3829.991**	8272.722	**8263.318**
50	**0.25**	512.0976	**488.0027**	2407.989	**2356.165**	5000.534	**4942.05**
	0.50	520.4511	**515.8235**	2620.238	**2604.272**	5512.023	5514.923
	0.75	543.8143	**540.5304**	2607.02	**2552.504**	5698.33	5707.325
75	**0.25**	255.7324	**287.7841**	1280.065	**1275.756**	2695.128	**2677.361**
	0.50	312.231	**299.871**	1338.623	1350.044	2850.098	**2848.829**
	0.75	269.8102	**267.6131**	1350.597	1353.066	2831.845	2841.659

From table 4, it is observed that the Partitioned models outperformed the Non-partition models in most occasion of censoring percentages levels of Weibull baseline variances and sample sizes, the Non-partition models were better. For 50% censoring at intermediate & high levels of WBVs, when $n = 1000$. For high censoring percentage (75%) at low levels of WBVs, when $n = 100$; at intermediate and high levels of WBV, when $n = 500$ and at high level of WBV, when $n = 1000$.

Again, from table 5, the Partitioned models outperformed the Non-partition models in most cases of censoring percentages, levels of Weibull baseline variances and sample sizes in terms of DIC. However, the Non-partition models were seen to be better for 0% censoring at intermediate level of WBV when sample sizes are 500 & 1000. For high censoring percentages (75%) at low level of WBV, when $n = 100$.

In this study, all models portray the same trend of DIC values with respect to increase in censoring percentages, levels of Weibull baseline variance and increase in sample sizes. It is observed that, a reduction in the event times (increased censoring) present better DIC values, increase in levels of baseline variance - increases the spread of the data which consistently reduce precision. Increase in sample sizes inflates the DIC values.

Table 5: DIC values for levels of censoring, three (3) levels of Weibull baseline variance, sample sizes for Non-Partitioned Model (PM) and Partitioned Model without spatial dependence

Censoring %	Variance	N=100		N=500		N=1000	
		NPM	**PM**	**NPM**	**PM**	**NPM**	**PM**
0	**0.25**	854.1665	**837.187**	4069.006	4945.71	8780.939	8878.76
	0.50	962.253	**936.7658**	4709.29	**4143.179**	10143.15	**10119.98**
	0.75	1033.307	**1014.516**	5136.902	**5072.229**	11035.66	**10961**
25	**0.25**	653.4906	**648.6781**	3246.832	**3233.434**	7001.295	**6997.888**
	0.50	716.0612	**708.4976**	3578.655	**3538.044**	7828.612	**7808.675**
	0.75	755.0077	**753.0131**	3813.489	**3797.836**	8341.513	**8286.271**
50	**0.25**	498.3278	**468.4429**	2445.837	**2423.315**	5034.049	**5022.247**
	0.50	519.2795	**517.3664**	2606.079	**2556.05**	5520.44	**5508.376**
	0.75	543.4234	**527.3858**	2601.271	**2548.174**	5692.22	**5655.104**
75	**0.25**	255.4689	266.2221	1263.73	**1237.466**	2725.757	**2599.078**
	0.50	488.3278	**461.421**	1298.471	**1289.735**	2833.02	**2706.389**
	0.75	269.1235	**261.0132**	1340.108	**1336.481**	2807.207	**2773.892**

3.1 Results for Diabetes Data

This section presents the results of analysis of diabetic patients' dataset who were admitted at the Nigerian air force hospital Abuja.

Table 6: Test for Proportional hazard assumption for the diabetic data set

Covariates	Chisq	p-value
Age	29.43	**5.8e-08**
BSL	11.88	**0.00057**
Gender	0.41	0.52180
M.Status	4.25	0.04915
Global Test	31.95	2.0e-06

It is observed from table 6 that covariate Age and Blood Sugar Level (BSL) are not proportional with p-values 5.8e-08 and 0.00057, less than the significance level. The global test again indicates a violation of the proportional hazard assumption a double check from the plot of beta against time for Age & BLS were then examined for non-proportionality.

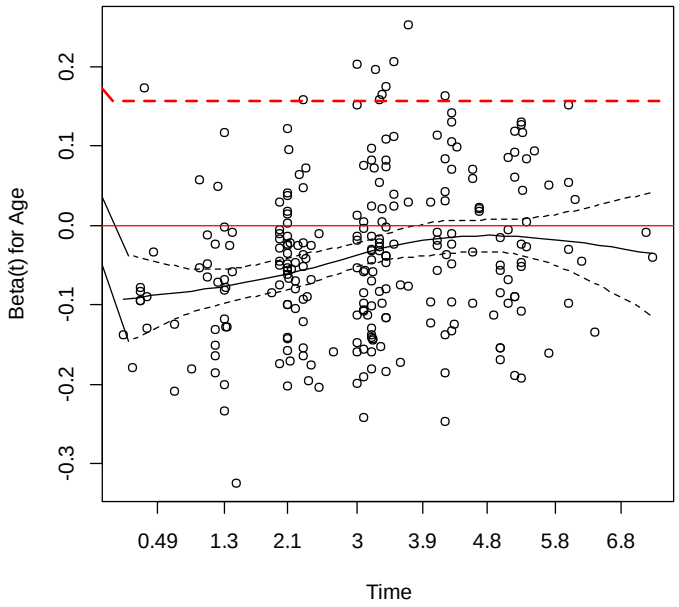

Figure 1: Graph of varying Coefficient of Age against time for diabetic patients

The solid smooth line is the estimated curve of time dependent departure of Age from proportionality, and the dotted lines are the estimated confidence bands.

From fig 1; it is observed from the graph of coefficient against time, that the effect of age on patient with the Diabetic disease is not constant through time, due to the non-parallel nature, which makes interpretation of hazard ratio difficult noticeable at time points 1.3, 4.2 and 5.8, the lower limits for the intervals. Non-overlapping intervals were then established to include the lower limits but not the upper limits, [0, 1.3), [1.3, 4.2), [4.2, 5.8), [5.8,∞).

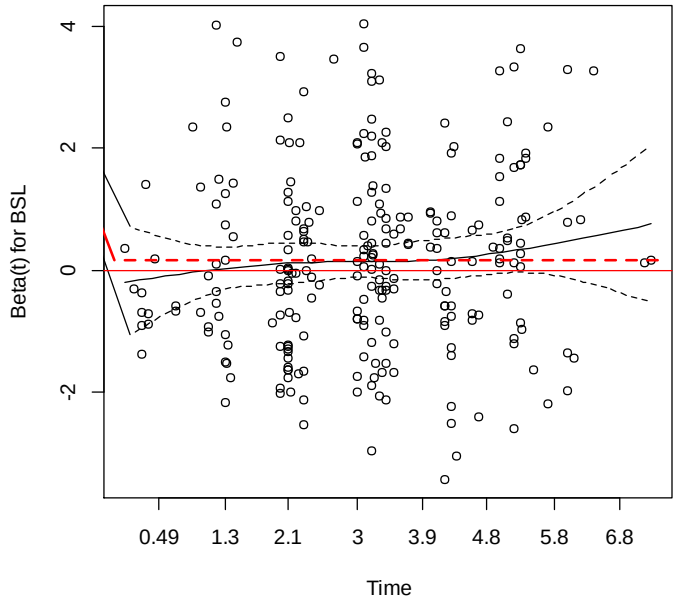

Figure 2: Graph of Varying Coefficient of Blood Sugar Level against time for diabetic patients

The solid smooth line is the estimated curve of time dependent departure of BSL from proportionality, and the dotted lines are the estimated confidence bands.

Fig 2 shows that the effect of Blood Sugar Level on patient with the Diabetic disease is slightly not constant through time, due to the continuous nature of such measures, which again makes interpretation of hazard ratio difficult, parameter estimation of this variable will be carried out via Penalized splines of degree 1 (linear) since it has a better DIC than those of quadratic (degree 2) and higher polynomials.

Table 7: Posterior coefficients for Non-Partition Model and Partition Model (PM) for Diabetic patient data

NPM				Model Selection Criteria	
Covariates		$\hat{\beta}$	Hazard ratios	DIC	LPML
Age		-0.03962	0.961155	1099.58	-551.3091
BSL		0.13826	1.148274		
Gender		0.31786	1.374184		
M.Status		-0.12299	0.884273		
PM					
Covariates		$\hat{\beta}$	Hazard ratios	DIC	LPML
Age	$\hat{\beta}_1$	-0.11674	0.889817	1081.807	-544.0822
	$\hat{\beta}_2$	0.07586	1.078812		
	$\hat{\beta}_3$	0.10176	1.107118		
	$\hat{\beta}_4$	0.11198	1.11849		
BLS		0.13685	1.146656		
Gender		0.33816	1.402365		
M.Status		-0.07655	0.926307		

From Table 7, it is seen that PM is a better model choice as it has the smaller DIC value and high predictive power with the LPML.

The bit of time interval for Age suggest that for every additional year of age the baseline hazard is associated with approximately 11% decrease in hazard for the first interval of time. 8% increase, 11% increase, and about 12% increases in hazard for the second, third and fourth bits of time respectively. This means that as the time of study rolls by, the hazard for experiencing the event death from the disease increases steadily between intervals but constant within each time interval.

BLS is continuous type variable, which suggest that for every unit increase in Blood Sugar Level (BSL) the baseline is associated with 15% increase in the hazard rate.

Male patients are 1.4024 times at risk than their female counterparts. Those that are married are 0.9263 times at risk than patients that are single.

Table 8: Posterior coefficients for Non-Partition and Partitioned model with P-splines for continuous covariate "Blood Sugar Level (BSL)

NPM (PI)		n=452		Model Selection Criteria	
Covariates		$\hat{\beta}$	Hazard ratios	DIC	LPML
Age		-0.03885	0.961895	1100.81	-553.025
pspline(BSL)					
1	$\hat{\beta}_1$	-13.36687	1.57E-06		
2	$\hat{\beta}_2$	-9.25078	9.6E-05		
3	$\hat{\beta}_3$	-11.26618	1.28E-05		
4	$\hat{\beta}_4$	-10.50300	2.75E-05		
5	$\hat{\beta}_5$	-10.19127	3.75E-05		
6	$\hat{\beta}_6$	-11.27552	1.27E-05		
7	$\hat{\beta}_7$	-9.70266	6.11E-05		
Gender		0.34914	1.417848		
M.Status		-0.14009	0.86928		
PM (PD)		n=964			
Covariates		$\hat{\beta}$	Hazard ratios	DIC	LPML
Age	$\hat{\beta}_1$	-0.11393	0.89232	1087.397	-548.6908
	$\hat{\beta}_2$	0.07522	1.078121		
	$\hat{\beta}_3$	0.09795	1.102908		
	$\hat{\beta}_4$	0.10517	1.110899		
pspline(BSL)					
1	$\hat{\beta}_1$	-10.5240	2.69E-05		
2	$\hat{\beta}_2$	-7.18950	0.000754		
3	$\hat{\beta}_3$	-8.77347	0.000155		
4	$\hat{\beta}_4$	-8.37850	0.00023		
5	$\hat{\beta}_5$	-7.71938	0.000444		
6	$\hat{\beta}_6$	-9.17565	0.000104		
7	$\hat{\beta}_7$	-6.70699	0.001222		
Gender		0.36060	1.43419		
M.Status		-0.09846	0.906232		

From table 8, it is noticed that the P-splines for the explanatory variable "BSL" execute in seven knots, to cushion in the smoothen effect that enables interpretation of hazard rates for BSL over time for Non-partitioned and Partitioned Models. We notice that PM outperforms the NPM with better values of DIC and LPML. Again, it is seen that the PM model in table 7 presents a better model when compared

with the PM on table 8 with respect to the DIC and LPML values of 1081.807 & -544.0822 respectively, also a model with the least number of parameters.

The first value of the estimate β= -10.5240 represent the effect of the linear term as the hazard rates for BSL were estimated in six knots over the period of study for the PM.

For every unit change in BSL, the baseline hazards across the six knots are associated with 99.92%, 99.98%, 99.98%, 99.95%, 99.99%, 99.88% decrease in the hazard rates respectively. This represents a slight change in the hazard rate estimated through the study period for smooth effects from the linear combination of basic splines

4. Conclusion

Non-Partition models and Partition models with or without spatial dependence work better when we have more events times (ie low censoring percentage) often at low level of WBV. Increase in censoring and levels of WBVs adversely affect both models but better with increase in sample size. It is seen that the models without spatial dependence adapts well when censoring are high and when survival times gain more spread. Partition Models incorporates spatial effects better than those of Non-Partition Models, as Partition models were most of the time observed to out-perform the Non-partition models.

It is observed that Age and Blood Sugar Level (BSL) covariates for the diabetic data set, violates proportionality assumptions upon test. Further assessment made from the graph of coefficient against time; suggest that Age be put to cut-points. This made the models interpretable via constant hazards for several mean posterior coefficients, this also enables a good assessment of changes due to Age differences for every interval for which such individual(s) were being followed and treated, this by implication could enhance; proper medical diagnoses/test, effective drug prescription, diet advise and overall management of patients with different peculiarities with regards to these variables. The Time changing variable Age, suggest that as the time of study rolls by, the hazard of experiencing the event death from the disease increases steadily between intervals but constant within each time interval. For every unit change in hazard rates the baseline hazard is associated with a decrease throughout the period of study for partition models fitted for with and without Penalized splines. Marriage is seen to be significant in the management of the disease in comparison to single patients. In addition, patients are advised to visit their physicians on a regular basis to run a routine check to keep their BSL in good range.

The study adds to knowledge in the following ways; provides a means of moving out of non-linear ruts in survival data analysis. Intervals increase sample sizes (pseudo-observations), which in turn improves the modified Partitioned model when they are with or without spatial dependence.

5. Acknowledgements

Thanks to God Almighty for the privileged to contribute to knowledge. My appreciation goes to Prof. B.A Oyejola for his meticulous assessment of the work and to the Editorial team of JOSMA for their conscious effort in reviewing the manuscript.

6. References

Adebayo, S.B. & Fahrmeir, L. (2005): Analyzing Child Mortality in Nigeria with Geoadditive Discrete-time Survival Models. *Statistics in Medicine*, 24:709-728.

Adeniyi, O.I & Akinrefon, A.A. (2018): First Birth Interval: Cox Regression Model with Time Varying Covariates. *Covenant Journal of Physical and Life Sciences*, 6(1): 1-7.

Arbia, G., Espa, G., Giuliani, D., Micciolo, A. (2016). A Spatial Analysis of Health and Pharmaceutical Firm Survival. *Journal of Applied Statistics*, 44(9): 1560-1575.

Banerjee, S., Carlin, B.P & Gelfand, A.E (2014). *Hierarchical Modeling and Analysis for Spatial Data*, Second Edition. Chapman and Hall/CRC Press.

Banerjee, S., & Dey, D.K (2005). Semiparametric Proportional Odds Models for Spatially Correlated Survival Data, *Lifetime Data Analysis*, 11(2), 175–191.

Bender, R., Augustin, T. & Blettner, M. (2005). Generating Survival Times to Simulate Cox Proportional Hazards Models. *Statistics in Medicine*. 24(11): 1713-1723.

Bastos, L. S. & Gamerman, D. (2006). Dynamic Survival Models with Spatial Frailty, *Lifetime Data Anal*, 12: 441–460.

Clark, D. E. & Ryan, L. M. (2002). Concurrent Prediction of Hospital Mortality and Length of Stay from Risk Factors On Admission, *Health Services Res*, 37: 631–645.

Darmofal, D. (2009). Bayesian Spatial Survival Models for Political Event Processes. *American Journal of Political Science*, 53(1), 241–257. ISSN 1540-5907.

Fabio, N. D, Rosangela, H. L, Enrico A, & Dipak, K. (2010), Extensions of the Piecewise Exponential Model. Corpus ID: 53392910, *Life Data Anal*, 14: 333-356

Hastie, T. & Tibshirani, R. (1990). *Generalized Additive Models*. Chapman and Hall London.

Hastie, T., and Tibshirani, R (1986). Generalized Additive Models. *Journal of Statistical Science*, 1(3): 297-310.

Henderson, R., Shimakura, S. & Gorst, D. (2002). Modeling Spatial Variation in Leukemia Survival Data. *Journal of the American Statistical Association*, 97(460): 965–972.

Hennerfeind, A., Brezger, A. & Fahrmeir, L. (2006). Geoadditive Survival Models. *Journal of the American Statistical Association*, 101(475): 1065–1075.

Jones, K. & Almond, S. (1992). Moving Out of the Linear Rut the Possibilities of Generalized Additive Models. *Trans Inst Br Geogr*, 17: 434-47.

Kim, S., Chen, M. H., Dey, D. K. & Gamerman, D. (2006). Bayesian Dynamic Models for Survival Data with a Cure Fraction, *Lifetime Data Anal*, 13: 17–35.

Kneib, T. (2006). Mixed Model-Based Inference in Geoadditive Hazard Regression for Interval Censored Survival Times. *Computational Statistics & Data Analysis*, 51(2), 777 – 792.

Kooperberg , C. & Intrator, N. (1995).Trees and Splines in Survival Analysis. *Statistical Methods in Medical Research*, 4: 237–261.

Lang, S. & Brezger, A. (2004): Bayesian P-splines. *Journal of Computational and Graphical Statistics*, 13:183-212.

Li, J., Hong, Y., Thapa, R. & Burkhart, H.E (2015a). Survival Analysis of Loblolly Pine Trees with Spatially Correlated Random Effects. *Journal of the American Statistical Association*, 110(510), 486–502.

Li, Y. & Lin, X. (2006). Semiparametric Normal Transformation Models for Spatially Correlated Survival Data. *Journal of the American Statistical Association*, 101(474), 591–603.

Marano, G., Boracchi, P. & Biganzoli, E.M. (2016) Estimation of the Piecewise Exponential Model by Bayesian P-Splines via Gibbs Sampling: Robustness and Reliability of Posterior Estimates. *Open Journal of Statistics*, 6: 451-468.

Omaku, P.E, Braimah J.O. Salisu, S.U & Abdulazeez, A.S (2021). A Piece Wise Modelling Function of Survival Analysis Data. *Kwara State University journal of Mathematical Science*, 2(1): 34-55 ISSN 2735-962x.

Omaku, P.E & Oyejola B.A (2020). A Piece-Wise Additive Model of Survival Data with Non - Linear Rut. *Australian Journal of Science and Technology*, 4(4): 386-391. ISSN Number (2208-6404).

Taylor, B.M. (2017). Spatial Modelling of Emergency Service Response Times. *Journal of the Royal Statistical Society A*, 180(2), 433–453.

Ulviya, .A. (2013). Frailty Models for Modelling Heterogeneity, McMasters University. Masters Dissertation.

Wang, S., Zhang, J. & Lawson, A.B. (2012). A Bayesian Normal Mixture Accelerated Failure Time Spatial Model and Its Application to Prostate Cancer. *Statistical Methods in Medical Research*, 25(2): 793-806. http://dx.doi.org/10.1177/0962280212466189.

Zhou, H. & Hanson, T. (2015). Bayesian Spatial Survival Models. *In Nonparametric Bayesian Inference in Biostatistics*, pp. 215–246. Springer-Verlag.

Zhou, H. & Hanson, T. (2017). A Unified Framework for Fitting Bayesian Semiparametric Models to Arbitrarily Censored Survival Data, Including Spatially-Referenced Data. *Journal of the American Statistical Association*, 113(522): 571-522.