An Approach to Modelling a Highly Flexible Fully Bayesian Analysis of Survival Data

Juliana I. Consul, Bunakiye Richard Japheth & Joseph A. Erho
[1]Department of Mathematics, Niger Delta University, Bayelsa State, Nigeria.
E-mail: iwori2001@yahoo.com; bunakiye.japheth@ndu.edu.ng; joseph.erho@ndu.edu.ng
[2&3]Department of Computer Science, Niger Delta University, Bayelsa State, Nigeria.
Phone: +2348036724426; +2348061324564; +2348035079644

ABSTRACT

A popular method of analysis of survival data is the Cox proportional hazard method. The form of the dependence of the hazard function on the covariate values is usually specified through a linear predictor in the proportional hazard model. There could be a form of flexibility in the way the covariates are incorporated into the survival model. For this reason, in this paper we propose a continuous parameter space model using a Gaussian process prior which will allow the covariates to remain the way they are and each individual would have a hazard function which depends on the covariate vector of the individual. The model parameters were updated one at a time making the mixing of the samplers very poor in the posterior distribution. In this research, we introduce sampling the principal components of the correlated parameters to improve the mixing. The Breast cancer data was used for illustration of the proposed model. All computations are performed using R software. Furthermore, the trace plots were given to assess the performance of the model parameters. Finally, our research found that the proposed model performed well and could be beneficial in the analysis of various types of survival data.

Keywords: Gaussian process, Bayesian methods, McMC, posterior distribution, sampling, principal components

CISDI Journal Reference Format
Juliana I. Consul, Bunakiye Richard Japheth & Joseph A. Erho. (2022): An Approach to Modelling a Highly Flexible Fully Bayesian Analysis of Survival Data . Computing, Information Systems, Development Informatics & Allied Research Journal. Vol 13 No 4, Pp 1-.12
Available online at www.isteams.net/cisdijournal

1. INTRODUCTION

Survival analysis is generally known as the analysis of time-to-event data in which the outcome is often referred to as event time. It is very usual that there exists the possibility that a subject may not observe the event of interest. This is called censoring (Klein, J. P. and Moeschberger, M. L. (2003) and Kartsonaki, 2016). Both censored and uncensored observations should be correctly incorporated into the survival models when estimating the model parameters. In many real-life applications, survival times are affected by explanatory variables. One of the objectives of the analysis of survival data might be to examine the relationship of a set of explanatory variables with the survival time. The Cox proportional hazard model is very popular for examining this relationship since it provides an easy interpretation (Cox, 1972). The explanatory variables are related to the response variable through a hazard-based regression model which can be formulated in a number of ways. The baseline hazard function is usually combined with the hazard multipliers and it depends on the values of the covariates through a linear predictor. The form of the dependence on the covariate values is usually specified.

There is the need to model flexibility into the dependence of the hazard on the values of the covariates. This study uses the Bayesian approach to inference which allow the covariates to remain the way they are and each subject has a hazard which depends on the covariate vector. The research is aimed at incorporating a Gaussian prior and a continuous parameter space model to allow for flexibility in the model.

Bayesian technique being a powerful and flexible alternative to the classical inferential method have been applied to survival analysis (Ibrahim et al., 2001). The Bayesian survival analysis model using McMC algorithm is a very useful instrument for its computational rapid advancements. In our proposed model, the Gaussian process prior is used as a suitable prior distribution for a framework for highly flexible fully Bayesian analysis of survival data. Some authors have proposed using Gaussian process to model. For instance, Aguilar, (2016) proposed an McMC algorithm where they used a Gaussian process to perform inferences. while Ahmed, (2017) proposed a model which uses multi-task Gaussian process to capture complex non-linear interactions between the patients' covariates and cause-specific survival time.

Some flexible models have been proposed in recent times. Thomas et al., (2016) constructed a tractable semi parametric alternative to the piecewise exponential model that assumes a continuous hazard. Ibrahim et al, (2001). presented a popular choice of accommodating flexible Bayesian analysis to survival data by assuming a piecewise exponential model while Thomas et al., (2016) proposed a more realistic piecewise linear log-hazard formulation that facilitates time dependent and proportional hazard covariate effects. Sharef et al,(2010) presented a semi-parametric Bayesian method for modelling both the unknown baseline hazard using mixtures of B-spline and Cai et al., (2002) also used the B-spline for modelling the log-hazard additively.

The application of Bayesian technique and its usage to analyze cancer data is common in recent years. Breast cancer is the commonest cause of mortality among women worldwide and the most common cancer among Nigerian women. The data used in this research is the breast cancer data from the University of Illorin teaching hospital, Illorin, Nigeria More research are required towards creating awareness on how the prevalence of breast cancer can be reduced. The Bayesian approach to inference produces precise estimates for modelling breast cancer but experience has shown that the mixing of the trace plots of the sampling of the model parameters might be poor due to correlated parameters and hence sampling the principal components of the correlated parameters will improve mixing in the sampling of parameters.

Principal component analysis can be applied to reduce the number of variables included in the model as well as eliminating possible correlation between covariates. It is formed from two stages, namely, forming the eigenvalues and eigen vectors of the samples covariance matrix which produces the principal components. Some authors have used principal components in survival analysis. For instance, Lin et al., (2006) applied principal components to the proportional hazard regression model in condition-based maintenance (CBM) optimization. Ma, (2021) used the principal component regression (PCR) technique as an approach to model reduction based on weight least squared estimate as it is very insensitive to the number of covariates. Yao, (2007) also proposed an approach to jointly model longitudinal and survival data using flexible basis function such as B-splines and the model dimension was reduced by the functional principal component analysis. Junttila & Laine, (2017) used a linear regression model that accounts for multicollinearity in the covariates by principal components and Bayesian regularization. In this research, the principal components will be used for improving the mixing of samples in Bayesian survival analysis.

2. METHODOLOGY

Two important functions in the analysis of Survival data are the survival and hazard functions (Lule & Fatmir, 2019). These are the key concepts for describing the distribution of survival times. The survival function $S_i(t)$ of a subject i be defined as the probability that the subject lives than some specified time t, for $0 < t < \infty$. The hazard function gives the potential that the event will occur per time unit, given that the subject has survived up to the specified time. Some of the most common ways of relating the covariates of the model to the survival distribution are the proportional hazard model (PHM) (Cox, 1972 and Fauziah, 2021) and accelerated life models (Abdul-Fatawu, 2020). Suppose that there are G covariates for $g = 1, 2, \dots . G$ and n individuals for $i = 1, 2, \dots \dots n$. The covariate vector of an individual is therefore denoted by $\underline{X_i} = \left(1, x_{i,1}, x_{i,2}, \dots . x_{i,G}\right)$.

The proportional hazard model assumes that the hazard function of an individual i given as $h_i(t)$ is written as $h_i(t) = \lambda_i \times h_0(t)$

where $h_0(t)$ is the baseline hazard function and the quantity λ_i is the hazard multiplier which depends on the covariates of the i^{th} individual. The linear predictor (also called the prognostic index) of the i^{th} individual η_i is written as a logarithmic link function given as follows:

$$\log \lambda_i = \eta_i = \beta_0 + \sum_{g=1}^{G} \beta_g x_{i,g} \qquad (1)$$

where β_0 is the baseline parameter and $x_{i,g}$ is the value of covariate g for individual i.

2.1 Continuous Parameter Space Model

In this research, we discuss a new approach of Bayesian modelling of flexibility into the dependence of the hazard on the covariate values but allowing the covariates to remain the way they are and giving the log – hazards a Gaussian process prior. We suppose that the form of the baseline hazard is a Weibull distribution and that each individual has a hazard which depends on the covariate vector (Li, 2012). It is possible that two individuals share the same covariate vector. This implies that the number of distinct covariate vectors or different hazard is n' where $n' < n$ given that the number of individuals in the study is n. And so, the hazard multipliers are $\lambda_1, \lambda_2, \dots \dots \lambda_{n'}$ where $\eta_p = \log(\lambda_p)$. We will also suppose that we have a linear model for $\eta_1, \eta_2, \dots \dots \eta_{n'}$ with $G + 1$ coefficients$\beta_0, \beta_1, \dots \dots \beta_G$.

Let $\underline{\beta} = (\beta_0, \beta_1, \dots \dots \beta_G)^T$ then,

$$\underline{\eta} = X \, \underline{\beta} \qquad (2)$$

where X is the design matrix (Puntanen, 2011) with n'rows and $G + 1$ columns where row i is the covariate vector for the i^{th} individual.

There is therefore exactly one parameter model for each covariate vector. These parameters are $\eta_1, \eta_2, \dots \dots \eta_{n'}$ are the logarithm of hazard (log-hazard).

We think of a Gaussian process prior as a suitable prior distribution for $\underline{\eta}$ and suppose that the lifetime random variable has a Weibull distribution with parameters $(\boldsymbol{\lambda}, \boldsymbol{\alpha})$ where $\boldsymbol{\lambda}$ is the scale parameter and $\boldsymbol{\alpha}$ is the shape

parameter which describes the shape of the hazard function (Muse, 2022). We will let the number of cases with covariate profile c be η_c, for $c = 1, 2, \ldots \ldots . n'$ and the number of these cases where the event of interest was observed be $\eta_{d,c}$. So, the number of all the cases that observed the event of interest will be
$\eta_d = \sum_{c=1}^{n'} \eta_{d,c}$

Let the hazard multiplier for covariate profile c be $\lambda_c = \exp\{\eta_c\}$

The likelihood contribution from covariate profile c, written as $L(\lambda_c, \alpha|D)$ from the data set, D is

$$L(\lambda_c, \alpha|D) = \left\{\prod_{k=1}^{n_c}\left[\alpha\lambda_c t_{c,k}{}^{\alpha-1}\right]^{d_{c,k}}\right\} \exp\left\{-\lambda_c \sum_{k=1}^{n_c} t_{c,k}{}^{\alpha}\right\} \qquad (3)$$

where $t_{c,k}$ is the event or censoring time for individual k in profile c and $d_{c,k} = 0$ if it is a censoring time and $d_{c,k} = 1$ if it is an event time.

We will suppose that the vector of the logarithm of hazards $\underline{\eta}$, has a multivariate normal prior distribution with vector of means $E(\underline{\eta})$ and covariance matrix $Var(\underline{\eta})$. We would think of a systematic way of obtaining $E(\underline{\eta})$ and $Var(\underline{\eta})$. In an attempt to make the prior means and variances different, we set up a preliminary linear model where we give all coefficients of the covariate's prior means and variances and hence the prior information or elicitation depends on the effects of the covariates on the log – hazard of the individuals. The prior means of the log – hazards $E(\underline{\eta})$ is then given as:

$$E\left(\underline{\eta}\right) = X\,E\left(\underline{\beta}\right) \qquad (4)$$

where $E(\underline{\beta})$is the vector of means of the effects of the covariates $\underline{\beta}$.

The covariance matrix of $\underline{\eta}$, $Var(\underline{\eta})$ would be constructed by first thinking of an implied covariance matrix

$$Var^*\left(\underline{\eta}\right) = X\,Var\,\left(\underline{\beta}\right)X^T \qquad (5)$$

where $Var\ \left(\underline{\beta}\right)$ is the covariance matrix of the effects of the covariates $\underline{\beta}$.

We will think of constructing a correlation matrix R based on the "distances" between the covariate vectors since $Var^*\left(\underline{\eta}\right)$ will be singular. The diagonal elements of $Var^*\left(\underline{\eta}\right)$ will be the marginal variances $\underline{V}^* = (V_1{}^*, V_2{}^*, \ \ldots \ldots, V_c{}^*)^T$ and $\underline{s}^* = (s_1{}^*, s_2{}^*, \ \ldots \ldots, s_c{}^*)^T$ be the vector of standard deviation where $s_k{}^* = \sqrt{V_k{}^*}$ and $S^* = diag\ (\underline{s}^*)$, which is a diagonal matrix with diagonal $\underline{s}^*$. Then,

$$Var\left(\underline{\eta}\right) = S^* R\, S^*$$

where R is a correlation matrix that we will want to construct.

The elements of the correlation matrix R are

$$r_{i,j} = \Lambda_{i,j} \left\{a + (1-a)\, exp\{-d_{i,j}\}\right\} \text{ for } i \neq j \qquad (6)$$

where the factor Λ for $0 \leq \Lambda_{i,j} \leq 1$ gives a provision for the inclusion of frailty (Guure et al., 2020). The value of $\Lambda_{i,j} = 1$ if $i = j$ and $\Lambda_{i,j} < 1$ otherwise. We also think of a value less than 1 for a.

The form of the correlation matrix will be based on the "distances" between the covariate vectors of the covariate profiles. The values of $d_{i,j}$ is the distance measure between the i^{th} and j^{th} covariate profiles and is given by

$$d_{i,j} = \sqrt{(\underline{x}_i - \underline{x}_j)^T D^* (\underline{x}_i - \underline{x}_j)}$$

where $\underline{x}_i$ and $\underline{x}_j$ are the covariate vectors for the i^{th} and j^{th} covariate profiles. The matrix D^* is a symmetric positive definite matrix which rescales the covariates.

2.2 Sampling from the posterior distribution of the logarithm of the hazards
In Bayesian Inference, it is often not feasible to draw independent samples from the posterior distribution since it might not have a standard form. The Markov chain Monte Carlo (McMC) which is a generalized and flexible way of simulating samples from the joint posterior distribution of the model parameters can be used in this case (Gilks et al., 1996, Ibrahim et al., 2001). The samples follow a Markov chain where each sample may depend on the previous one.

One possible way of sampling from the posterior distribution is to sample each of the log – hazards one at a time. In this case, the conditional prior distribution of each log – hazard given the other log – hazards will be a normal distribution. For instance, the conditional distribution of the c^{th} log – hazard given the other log – hazards $\eta_c | \eta_{c'}$ has a normal distribution with the conditional mean $M_{c|c'}$ and variance $V_{c|c'}$.

The conditional mean $M_{c|c'}$ is given as

$$M_{c|c'} = m_c + C_{cc'} V_{c'}^{-1} (\underline{\eta}_{c'} - \underline{m}_{c'})$$

and the conditional covariance matrix $V_{c|c'}$ is given as

$$V_{c|c'} = V_{cc} - C_{cc'} V_{c'}^{-1} C_{c'c}$$

where $\underline{\eta}_{c'}$ is a $(n'-1) \times 1$ column matrix without $\underline{\eta}_c$,Type equation here.
$\underline{m}_{c'}$ is the vector of means of the log – hazards without the mean of $\underline{\eta}_c$,
V_{cc} and $C_{cc'}$ can be got by partitioning the covariance matrix of the log – hazards as follows:

$$[|matrix *********]$$

where $C_{c'c}$ is the transpose of C_{cc} . All other conditional log – hazards are computed in a similar way.

Therefore, the conditional prior distribution of the c^{th} log – hazard $\pi(\eta_c)$ has conditional mean $M_{c|c'}$ and variance $V_{c|c'}$.
We use the Metropolis within Gibbs algorithm and we fix the Weibull shape parameter α and then we sample each of the log – hazard from the full conditional distribution which is given as

$$\pi(\eta_c, \alpha |\, D) \propto \pi(\eta_c)\, L(\eta_c, \alpha |\, D)$$

where $L(\eta_c, \alpha |\, D)$ is the likelihood contribution from the c^{th} covariate profile.
A value ${\eta_c}^*$ for η_c is proposed from a normal distribution.

The proposed log – hazard ${\eta_c}^*$ is accepted with probability

$$A_{Prob} = min \left\{ 1, \qquad \frac{\pi({\eta_c}^* D)}{\pi\left(\eta_{c_p} |D\right)} \right\}$$

using the Metropolis – Hastings algorithm and symmetric proposal where the proposal densities cancel out. The accepted values of the log – hazards $\underline{\eta}$ are fixed and the Weibull shape parameter α using a gamma prior distribution with parameters a, b which is given as

$$\pi(\alpha|a, b) \;\propto\; \alpha^{a-1}\, exp\{-\, b\, \alpha\}.$$

The posterior density is then given as

$$\pi(\alpha|\, \underline{\eta}\,, D) \propto \pi(\alpha|a, b)\, L(\alpha|\, \underline{\eta}\,, D)$$

A value α^* for α is proposed from a Gamma distribution with some specified a^* , b^*. The proposal density of α^* given the value of α has a Gamma distribution and is denoted as $q(\alpha^*|\alpha)$ and the proposal density of α given the value of α^* is similarly given as $q(\alpha|\alpha^*)$.

The proposed value of α^* is accepted with probability

$$A_{Prob} = min \left\{ 1, \qquad \frac{\pi(\alpha^*|\, \underline{\eta}\,, D)}{\pi(\alpha|\, \underline{\eta}\,, D)}\; \frac{q(\alpha|\alpha^*)}{q(\alpha^*|\alpha)} \right\}$$

Experience has shown that the disadvantage of updating the model parameters one at a time is that the mixing (Yao, 2007) can be poor if the model parameters are highly dependent in the posterior. Sometimes, “blocking” of the groups of correlated parameters can be used to improve mixing. In this research, instead of sampling the correlated parameters, we introduce sampling the principal components of the correlated parameters which are not correlated in the prior. This will improve the poor mixing in the sampling of parameters.

2.3 Sampling the Principal Components of the logarithm of the hazards

The logarithm of the hazards (log-hazards) are correlated when sampling one at a time and we transform the log-hazards to a new set of parameters which are not correlated in the prior. This is done using the principal components. The main reason for sampling the principal components of the log-hazards is that the principal components given the prior covariance matrix are linear functions of the log-hazards. Again, the principal components are independent and uncorrelated in the prior.

When finding the principal components, the eigen values and eigen vectors are usually involved. We will suppose that γ is a matrix where the columns** are the eigen vectors of the log-hazards $\underline{(\underline{\eta})}$. The vector of the principal components $(\underline{P})$ is given by

$$\underline{P} = \gamma \, \underline{\eta} \tag{7}$$

The covariance matrix of the principal components of the log – hazards $Var\ (\underline{P})$ which is a diagonal matrix of the eigen values is be given as

$$Var\ (\underline{P}) = \gamma \, Var\ \left(\underline{\eta}\right) \gamma^T \tag{8}$$

And the vector of means of the principal components of the log - hazards $E(\underline{P})$ is also given as

$$E(\underline{P}) = \gamma \, E\left(\underline{\eta}\right) \tag{9}$$

Bayesian modelling of the principal components of the log - hazard

The prior distribution of the p^{th} principal component of the log – hazard has a normal distribution with mean given as $E(P_p)$, variance as $Var(P_p)$ and probability density $\pi(P_p)$, given as follows:

$$\pi(P_p) \propto \exp\left\{-\frac{1}{2}\frac{\left(P_p - \ E(P_p)\right)^2}{Var(P_p)}\right\}$$

The likelihood contribution from the log – hazards is given the data $L\left(\underline{\eta}\,|D\right)$.

The Metropolis within Gibbs sampling is used since we can not sample directly from the full conditional distribution. The full conditional distribution of the p^{th} principal component of the log – hazard P_i is given as

$$\pi(P_p|D) \ \propto \ \pi(P_p) L\left(\underline{\eta}\,|D\right)$$

A value $P_p{}^*$ for P_p is proposed from a normal distribution. Since the principal component of the log – hazards are a function of the log – hazards, the proposed value of the log - hazards $\underline{\eta}^*$ are computed as follows:

$$\underline{\eta}^* = \gamma^{-1}\ \underline{P}^* \tag{10}$$

The proposed log – hazard $\eta_i{}^*$ is accepted with probability

$$A_{Prob} = min\left\{1, \qquad \frac{\pi\left(P_p{}^*|D\right)}{\pi\left(P_p|D\right)}\right\}$$

using the Metropolis – Hastings algorithm and symmetric proposal where the proposal densities cancel out. The value of the log – hazard will be transformed from the proposed principal component of the log – hazard if the proposed value of the principal component of the p^{th} log – hazard is accepted using Equation (10).

3. APPLICATION

We applied the discussion in Section 2.1 using a breast cancer data set from the University of Illorin teaching hospital, Illorin, Nigeria for a period of five years (Oguntunde & Okagbue, 2017). The breast cancer data set consists of the length of stay and the status (dead or alive) after treatment from year 2011 to 2016. The study time are recorded in months. The other four covariates are as follows:

- Age: This is the age (in years) of the patient.
- Sex: This is the gender of the patient. Female was indicated as "1" while male as "2".
- Mode of diagnosis (mode): This is the mode of diagnosis of the cancer. Cytological was indicated as "1" while Histological was indicated as "2".
- Location of breast cancer (location): This indicates the location of the breast cancer on the survivability of the breast cancer patients. Left breast was indicated as "1", right breast was indicated as "2" and both breast was indicated as "3". We will follow the discussions in Section 2.1.

3.1 Improving Mixing by sampling the principal components of the log – hazards

The parameters of the continuous parameter space model seem to be strongly correlated and hence, the mixing was poor. We think of sampling the log – hazards. We use the Metropolis within Gibbs algorithm and follow discussions in Section 2.3 to sample the principal components of the log – hazards from the full conditional distribution while we fix the Weibull shape parameter α. The value of the Weibull shape parameter α is then samples while the sampled values of the log – hazards are fixed.

Following a burn-in of 3000 iterations, 7000 iterations were taken. The convergence was assessed by visual inspection of the trace plots of the log – hazards. Figure 2 shows that the mixing was good. The numerical summaries of some of the values of the log – hazards using the breast cancer data set are given in Table 1

Table 1: Numerical summaries of the posterior means and standard deviations of some of the parameters using the Continuous parameter space model for the breast cancer data set.

Parameters	Posterior means	Posterior standard deviations
$\underline{\eta}_1$	-2.4612256	0.57280387
$\underline{\eta}_2$	-2.1933383	0.61172503
$\underline{\eta}_3$	-3.4784913	0.86017873
$\underline{\eta}_4$	-1.7643344	0.54795379
$\underline{\eta}_5$	-2.3751723	0.69799227
$\underline{\eta}_6$	-3.7921777	0.88400490
$\underline{\eta}_7$	-2.8174202	0.70769464
$\underline{\eta}_8$	-1.7624426	0.71721082
$\underline{\eta}_9$	-1.4605110	0.78574080
$\underline{\eta}_{10}$	-3.1151225	0.82175697
α	0.5051676	0.03678672

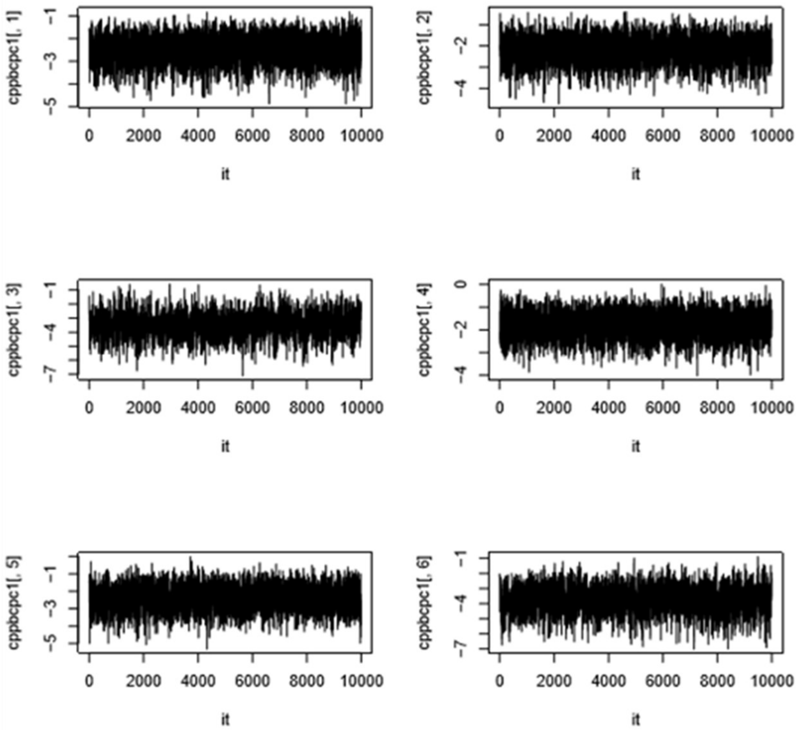

Figure 1 and 2 show the visual summary of some of the parameters of the breast cancer data set were updated one at a time and sampling the principal components respectively. Figure 2 shows that the McMC sampling process converges to the posterior distribution and thus the chains converged.

4. CONCLUSION

In this study we explored how to incorporate flexibility into the dependence of the hazard function on the covariates using a continuous parameter space and McMC techniques as an alternative for estimating the parameters of the proposed model that is more flexible than using the usual traditional method. The proposed method was illustrated using a real-world data involving a right-censored breast cancer data set. The Bayesian inference was performed with a Gaussian process prior and suppose that the lifetime random variable has a Weibull distribution. The convergence pattern was investigated using trace plots. Furthermore, the trace plots were given to assess the performance of the model parameters. The Bayesian approach in the analysis of breast cancer data is aimed at making an individual life affordable and comfortable Finally, our research found that the proposed model performed well and could be beneficial in the analysis of various types of survival data. We hope that this study encourages researchers to employ and conduct their analysis using this approach with the help of R software. In terms of future work, we intend to improve on this method by including frailties and also apply this model to including other types of censoring such as left, interval and double censoring.

REFERENCES

1. Abdul-Fatawu Majeed. Accelerated Failure Time Models: An Application in Insurance Attrition. The Journal of Risk Management and Insurance, 2020. ffhal-02953269
2. Aguilar, T & Rivera, Nicolas & Teh, YW. (2016). Gaussian processes for survival analysis, Conference: Advances in Neural Information Processing Systems.
3. Ahmed M. Alaa (2017), Deep Multi-task Gaussian Processes for Survival Analysis with Competing Risks , 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
4. Alvares, Danilo & Lázaro, Elena & Gómez Rubio, Virgilio & Armero, Carmen. (2020). Bayesian survival analysis with BUGS.
5. Bartoš, František & Aust, Frederik & Haaf, Julia. (2021). Informed Bayesian survival analysis.,Journal of Theoretical and Applied Information Technology, 100(19)
6. Cai, T., Hyndman, R., and Wand, M. (2002). “Mixed model-based hazard estimation.” Journal of Computational and Graphical Statistics, 11(4): 784–798. MR1944263. doi: http://dx.doi.org/10.1198/106186002321018786. 382
7. Cox D., (1972) Regression models and life-tables. Journal of the Royal Statistical Society, Series B 34: 187-220.
8. Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (1996) Markov chain Monte Carlo in practice. Chapman & Hall.
9. Guure, Chris & Alotaibi, Refah & Rezk Hoda (2020). Bayesian frailty modeling of correlated survival data with application to under five mortality. BMC Public Health, 20, 1-24. 10.1186/s12889-020-09328-7.
10. Hennerfeind, A., Brezger, A., and Fahrmeir, L. (2006). “Geoadditive Survival Models.” Journal of the American Statistical Association, 101(475): 1065–1075. MR2324146. doi: http://dx.doi.org/10.1198/016214506000000348. 382
11. Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). Bayesian Survival Analysis. New York: Springer. MR1876598. doi: http://dx.doi.org/10.1007/978-1-4757- 3447-8. 382, 384
12. Junttila, Virpi & Laine, Marko. (2017). Bayesian Principal Component Regression model with spatial effects for forest inventory under small field sample size. Remote Sensing of Environment. 192. 45-57. 10.1016/j.rse.2017.01.035.

13. Kartsonaki, Christiana. (2016). Survival analysis. Diagnostic Histopathology. 22(7). 263-270, 10.1016/j.mpdhp.2016.06.005.
14. Klein, J. P. and Moeschberger, M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data. New York, NY: Springer. 383
15. Li, Erling & Lim, Johan & Kim, Kyunga & Lee, Shin-Jae. (2012). Distribution-free Tests of Mean Vectors and Covariance Matrices for Multivariate Paired Data. Metrika. 75. 833-854. 10.1007/s00184-011-0355-7.
16. Lin, Daming & Banjevic, D & Jardine, Andrew. (2006). Using principal components in a proportional hazards model with applications in condition-based maintenance. Journal of The Operational Research Society - J OPER RES SOC. 57. 910-919. 10.1057/palgrave.jors.2602058.
17. Lule Basha & Fatmir Hoxha (2019). Kernel Estimation of the Baseline Function in the Cox Model, European Scientific Journal, 15(6): 105-118. Doi: 10.19044/esj.2019.v15n6p105
18. Ma, Steven. (2021). Principal Component Analysis in Linear Regression Survival Model with Microarray Data. Journal of Data Science. 5. 10.6339/JDS.2007.05(2).326.
19. Michele Campolieti (2000), Bayesian Estimation and Smoothing of the Baseline Hazard in Discrete Time Duration Models, The Review of Economics and Statistics, 82 (4), 685-694.
20. Muse, A. Hassan, Muse, O. N, Samuel. M, Huda M. A & Abdal-Aziz H. (2022) A flexible Bayesian Parametric Proportional Hazard Model: Simulation and Applications to Right- censored Healthcare Data. Journal of Healthcare Engineering, https://doi.org/10.1155/2022/2051642
21. Oguntunde, P. E. and Okagbue, H. I. (2017). Breast cancer patients in Nigeria: Data Exploration approach. Data in Brief, 15, 47-57.
22. Perperoglou, A. et al. (2006). Reduced – rank hazard regression for modelling non – proportional hazards. Statistics in Medicine, 25(16), 2831-2845. Doi:10.1002/sim.2360.
23. Puntanen, Simo & Styan, George & Isotalo, Jarkko. (2011). Matrix tricks for linear statistical models. Our personal top twenty. 10.1007/978-3-642-10473-2.
24. Royston, P., & Parmar, M. K.B. (2002). Flexible parametric proportional hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatments effects. Statistics in Medicine, 21(15), 2175-2197. Doi: 10.1002/sim.1203.
25. Sharef, E., Strawderman, R. L., Ruppert, D., Cowen, M., and Halasyamani, L. (2010). “Bayesian adaptive B-spline estimation in proportional hazards frailty models.” Electronic Journal of Statistics, 4: 606–642. MR2660535. doi: http://dx.doi.org/ 10.1214/10-EJS566. 382, 386
26. Shen Y & Huang S (2006) Improve survival Prediction using Principal Components of Gene Expression Data, Genomics Proteomics, 4(2): 110-119. doi: 10.1016/S1672-0229(06)60022-3
27. Thomas A. Murray, Brian P. Hobbs, Daniel J. Sargent and Bradley P. Carlin (2016), Flexible Bayesian Survival Modeling with Semiparametric Time-Dependent and Shape-Restricted Covariate Effects, International Society for Bayesian Analysis, 11(2), 381-402. DOI: 10.1214/15-BA954
28. Yao Fang (2007), Functional Principal component analysis for longitudinal and survival data, Statistica Sinica 17: 965-983

APPENDIX

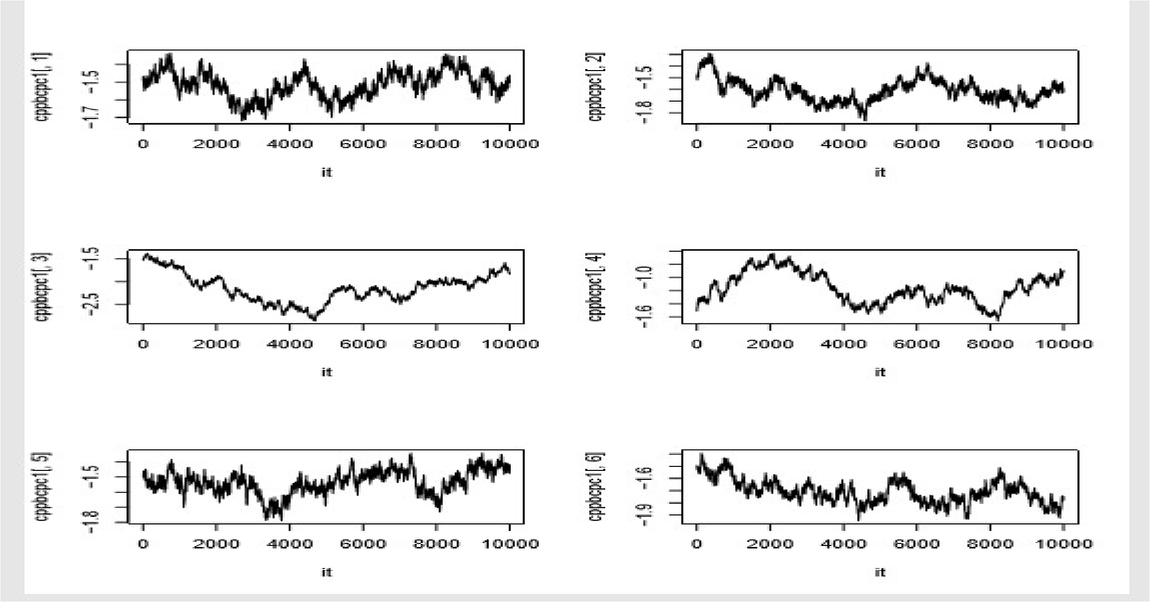

Figure 1: Visual summary of some of the parameters of the breast cancer data when sampling one at a time

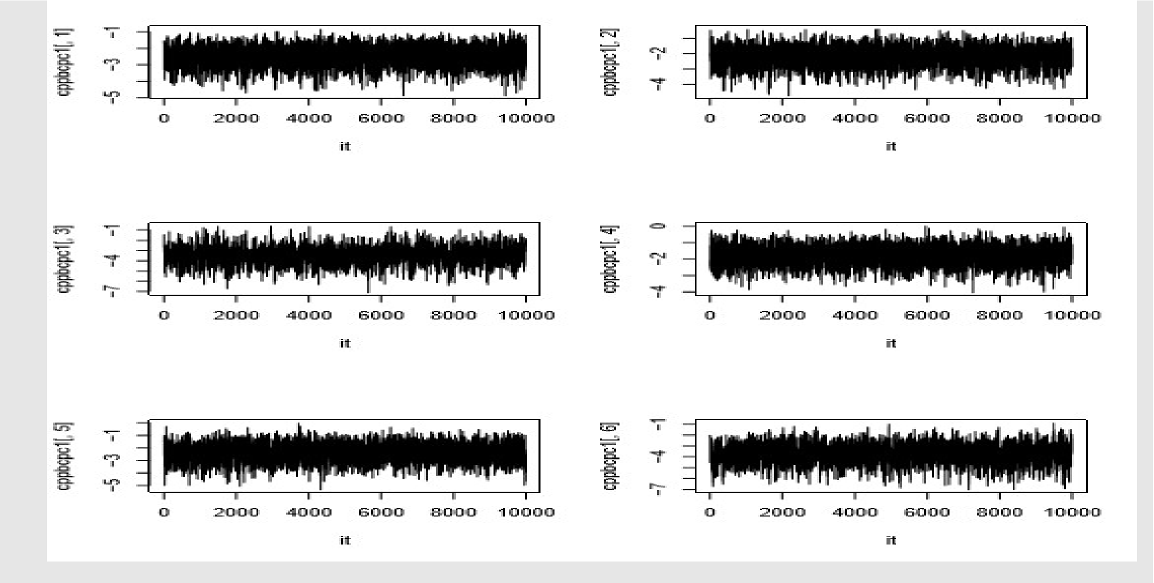

Figure 2: Visual summary of some of the parameters of the breast cancer data when the principal components were sampled