



A Bayesian semiparametric regression model for reliability data using effective age



Li Li, Timothy E. Hanson*

Department of Statistics, University of South Carolina, USA

ARTICLE INFO

Article history:

Received 8 January 2013

Received in revised form 28 October 2013

Accepted 22 November 2013

Available online 12 December 2013

Keywords:

Effective age

Repairable system

Tailfree process

Truncated data

ABSTRACT

A new regression model for recurrent events from repairable systems is proposed. The effectiveness of each repair in Kijima models I and II is regressed on repair-specific covariates. By modeling effective age in a flexible way, the model allows a spectrum of heterogeneous repairs besides “good as new” and “good as old” repairs. The density for the baseline hazard is modeled nonparametrically with a tailfree process prior which is centered at Weibull and yet allows substantial data-driven deviations from the centering family. Linearity in the predictors is relaxed using a B-spline transformation. The method is illustrated using simulations as well as two real data analyses.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Repairable systems have been widely studied in the literature. Systems fail, get repaired upon failure, and these recurrent events (failures, repairs) are observed. The event process generating the repeated events is closely related to the intensity function, denoted as $\lambda(t|H(t))$ and formally defined in Section 2.1, which describes the probability of an instantaneous new failure, given the history of maintenances and failures $H(t)$. In general, recurrent event modeling methods can be divided into categories based on the type of repairs a system receives. Renewal processes are used if all the repairs bring the system to the “good as new” state and Poisson processes are used if all the maintenances bring the system to a “good as old” state. Kijima (1989) introduced two classes of models using the notion of “effective age” (also known as “virtual age”) of the system to allow for a spectrum of repairs between “good as old” and “good as new”. Consider a system observed over $[0, \tau]$. Assume the repair times for the system are $0 < t_1 < t_2 < \dots < t_n$, and denote $\varepsilon(t)$ as the effective age of the system at time t . Suppose that the intensity $\lambda(t|H(t))$ is related to the unknown hazard, or failure rate, of a new system $r(t)$ through $\lambda(t|H(t)) = r(\varepsilon(t))$. Poisson models assume $\varepsilon(t) = t$ and renewal models assume $\varepsilon(t) = t - s_{N(t-)}$ where $s_{N(t-)}$ is the time at which the last repair occurred. Kijima models introduce an age reduction factor D_i for each repair, occurring at calendar time t_i . Define $\varepsilon(t_i) = \varepsilon(t_{i-1}) + [t_i - t_{i-1}]D_i$ for the Kijima type I model and $\varepsilon(t_i) = [\varepsilon(t_{i-1}) + t_i - t_{i-1}]D_i$ for the Kijima type II model. Assume $\varepsilon(t) = \varepsilon(t_{i-1}) + t - t_{i-1}$ for $t \in (t_{i-1}, t_i)$. Note that $D_i = 1$ implies a Poisson process in models I and II, and $D_i = 0$ implies a renewal process in type II.

Lindqvist (2006) provides a review of the modeling of effective age. Dorado et al. (1997) generalize Kijima's models that allow for repairs of varying degree by including known “life supplements” – numbers between zero and one indicating the degrees of the repairs. There is very limited literature dealing with unknown effective age processes. Doyen and Gaudoin (2004) studied a class of Kijima's models where the repairs reduce the effective ages by one overall effectiveness scalar q . Recently, Veber et al. (2008) propose an EM-algorithm to estimate q and use Weibull mixtures for the baseline failure time distribution. Using one scalar is inappropriate for systems where repairs of varying effectiveness occur. For example,

* Correspondence to: 219C LeConte College, University of South Carolina, Columbia, SC 29208, USA. Tel.: +1 803 777 3859; fax: +1 803 777 4048.
E-mail addresses: lil@email.sc.edu (L. Li), hanson@stat.sc.edu (T.E. Hanson).

different maintenance types, or levels of experience in those carrying out the repairs, can have drastic differences in repair effectiveness. Very recently, [Yuan and Uday \(2012\)](#) extend the single scalar parameter q to a time-dependent function, e.g. $q(t) = \exp(-et)$ where e is estimated, and assume the baseline distribution to be a parametric power-law distribution. In this work, we regress the effectiveness of each repair on covariates, e.g. materials used or the technician, and relax the parametric assumption of the baseline distribution using nonparametric priors in a Bayesian framework. Time trends in the effectiveness of repairs, characteristics of each repair, and association among repairs within each system can be flexibly coded into the covariate process. Specifically, the effectiveness measure D_i is regressed on a vector of covariates \mathbf{w}_i ; let $D_i = \exp\{\beta'\mathbf{w}_i\}/\{1 + \exp(\beta'\mathbf{w}_i)\}$ or $D_i = \exp(\beta'\mathbf{w}_i)$. The associations between the covariates and the effective age reduction are characterized by β . When the hazard of the system is monotone increasing, a repair with covariates resulting in a smaller age reduction factor D_i tends to be more effective than other repairs performed at the same effective age of the system.

Other generalizations of renewal and Poisson processes allowing for covariates also assume the effective age process $\varepsilon(t)$ is known, including for example, modulated renewal processes ([Cox, 1972](#)), point-process models incorporating renewals and time trends ([Lawless and Thiagarajah, 1996](#)), and a general class of semiparametric models ([Peña et al., 2007](#)) which simultaneously accommodates the effects of increasing numbers of events, covariates, interventions (repairs), and association among the interevent times within a system. This literature encompasses a rich and widely used family of reliability models. However, it is difficult to assume that the effective age process is known. There might even be interplay among the effective age process and history-dependent covariates and the baseline hazard function, as noted in [Peña et al. \(2007\)](#). Moreover, understanding the performance of repairs is often crucial to decision-making and even predictions.

A parametric analysis of our proposed model can be performed by choosing an appropriate distribution family, e.g. Weibull, for $r(t)$. In this work, we seek a more flexible approach where the entire density, the cumulative hazard function, or the hazard is assigned a nonparametric prior distribution. Bayesian nonparametric priors have achieved prominent success due to their flexibility in modeling unknown distributions; examples include the Dirichlet process ([Ferguson, 1973](#)), Polya tree priors ([Lavine, 1992](#)), Dirichlet process mixtures ([Escobar and West, 1995](#)), etc. However, the use of these nonparametric priors in recurrent event models has been quite limited. Very recently, [Taddy and Kottas \(2012\)](#) used Dirichlet process mixtures for the interfailure density in Poisson process models. Priors on the cumulative hazard $R(t) = \int_0^t r(s)ds$ include the beta and gamma processes ([Lo, 1992](#); [Kuo and Ghosh, 1997](#)) which are discrete and not readily used in our context. The weighted gamma process ([Ishwaran and James, 2004](#)) is centered at one unique baseline intensity and is also not appropriate for a model that involves a factor in the argument of the intensity. Our proposed framework uses tailfree priors ([Freedman, 1963](#); [Ferguson, 1974](#); [Jara and Hanson, 2011](#)), on the space of densities, centered at the Weibull family, but allows for substantial data-driven deviations from the centering families. A special case of the tailfree prior, the Polya tree prior, has been widely used for models that warp the baseline r ; see [Hanson \(2006\)](#), [Walker and Mallick \(1999\)](#), and [Hanson and Yang \(2007\)](#) for applications involving the accelerated failure time model and the proportional odds model. Like the Dirichlet process, tailfree priors also have desirable consistency and large support properties ([Jara and Hanson, 2011](#)). The general framework proposed herein allows model comparisons using the goodness-of-fit measures LPML and DIC so that comparisons among renewal processes, Poisson processes and Kijima models are readily made. We develop a full, automated MCMC sampling scheme to fit our proposed model and illustrate our method using simulations as well as on real data.

This paper is organized as follows: Section 2 presents a description of our model and an introduction to tailfree priors. Section 3 provides the MCMC algorithm and an approach to relax linearity in the linear predictor, and Section 4 presents simulation results. Section 5 summarizes the results for two real dataset analyses and in Section 6 we provide some concluding remarks.

2. Model development

2.1. Likelihood construction

Consider a system starting from new. Suppose the system gets repaired at times t_i , $i = 1, \dots, n$ and $0 < t_1 < t_2 < \dots < t_n < \tau$ where τ is the time when data collection stops. We assume τ is independent of the failure process. If a repair is performed without an accompanying failure, the observation of event time is right censored. Let the indicator δ_i take the value 1 if the system fails at time t_i and 0 otherwise. Further we assume a d -dimensional covariate vector for each repair, independent of the failure process, i.e. $\mathbf{w}_i = (w_{i0}, w_{i1}, \dots, w_{i,d-1})$ for the repair at time t_i . This vector may incorporate information concerning technician skills, repair type, materials used, time trend, etc. Let the counting process $\{N(t), t \geq 0\}$ record the cumulative number of failures over time and $H(t) = \{N(s) : 0 \leq s < t\}$ be the history of the process at time t . The intensity function for an event process is defined as

$$\lambda(t|H(t)) = \lim_{\Delta \rightarrow 0} \frac{P\{N(t + \Delta) - N(t) = 1|H(t)\}}{\Delta^+}. \quad (1)$$

The Kijima models for the event data assume $\lambda(t|H(t)) = r\{\varepsilon(t)\}$ where $\varepsilon(t)$ is the effective age. A Kijima type I model has $\varepsilon(t_i) = \varepsilon(t_{i-1}) + [t_i - t_{i-1}]D_i$ and the type II model has $\varepsilon(t_i) = [\varepsilon(t_{i-1}) + t_i - t_{i-1}]D_i$ where $t_i - t_{i-1}$ is the time since last repair. Denote the effective age right before t_i as $\varepsilon(t_{i-})$. The i th repair at t_i reduces the effective age right before t_i by

a fraction of the time since last repair, that is, $(t_i - t_{i-1})(1 - D_i)$ in the type I model and a proportion of the effective age, i.e. $\varepsilon(t_{i-1})(1 - D_i)$ in the type II model. Note that $D_i = 0$ sets the clock back to the status right after last repair in the type I model and to a new status in the type II model.

We propose to model D_i as a function of \mathbf{w}_i through the regression coefficient β . Let $D_i = \exp(\beta' \mathbf{w}_i)$ or $D_i = \text{logit}(\beta' \mathbf{w}_i)$. Let $w_{i0} = 1$ be an intercept. When the link function is the CDF of a logistic distribution, $D_i \in (0, 1)$ for all repairs, i.e. all repairs are between “good as new” and “bad as old”. If the link is exponential; then $D_i \in (0, +\infty)$. An interesting case would be $D_i > 1$ where the system actually gets worse than “bad as old” after the repair. When the baseline hazard is monotone nondecreasing or nonincreasing, β can be interpreted directly with respect to effectiveness of repairs, i.e. if the baseline hazard is nondecreasing and β_j is positive, one may conclude that an increase in w_j results in less effective repairs overall.

We refer to Lindqvist (2006) in deriving the likelihood of observing a system with failures at $0 < t_1 < t_2 < \dots < t_n < \tau$ ($\delta_i = 1, i = 1, \dots, n$):

$$L = \prod_{i=1}^n r(\varepsilon(t_{i-1}) + x_i) \exp \left\{ - \sum_{i=1}^n \int_0^{x_i} r(\varepsilon(t_{i-1} + u)) du - \int_0^{\tau - t_n} r(\varepsilon(t_n + u)) du \right\}, \quad (2)$$

where $x_i = t_i - t_{i-1}$. The likelihood is equivalent to

$$L = \prod_{i=1}^n \frac{f(\varepsilon(t_{i-1}) + x_i)}{S(\varepsilon(t_{i-1}))} \cdot \frac{S(\varepsilon(t_n) + \tau - t_n)}{S(\varepsilon(t_n))},$$

where S and f are the unique survival and density functions corresponding to r . Denote F as the cumulative distribution function for r . Now suppose we observe m identical systems. Denote t_{ij} as the event time for the i th repair of system j , \mathbf{w}_{ij} as the covariate vector and δ_{ij} as the censoring indicator. Let τ_j be the termination time for observing system j . Conditional on the collection of observables $\text{data} = \{t_{ij}, \tau_j, \mathbf{w}_{ij}, \delta_{ij}, i = 1, 2, \dots, n_j, j = 1, 2, \dots, m\}$, the likelihood of observing m independent event processes is then

$$L = \prod_{j=1}^m \prod_{i=1}^{n_j} \frac{\{f(\varepsilon(t_{i-1,j}) + x_{ij})\}^{\delta_{ij}} \{S(\varepsilon(t_{i-1,j}) + x_{ij})\}^{1-\delta_{ij}}}{S(\varepsilon(t_{i-1,j}))} \cdot \prod_{j=1}^m \frac{S(\varepsilon(t_{n_j,j}) + \tau_j - t_{n_j,j})}{S(\varepsilon(t_{n_j,j}))}, \quad (3)$$

where $x_{ij} = t_{ij} - t_{i-1,j}$.

2.2. Prior specifications

2.2.1. Tailfree process prior on F

We place a tailfree process prior on F , centered at the Weibull family. Denote G_θ as the cumulative distribution function for Weibull, $G_\theta(t) = 1 - \exp(-(t/\eta)^\alpha)$ for $t \geq 0$ and $\theta = (\log(\alpha), \log(\eta))'$. Let $\Pi_j = \{B_{\epsilon_1 \dots \epsilon_j} : \epsilon_i \in \{0, 1\}\}$ be a partition of the positive reals \mathbb{R}^+ and each set in Π_j be split into two sets in Π_{j+1} , e.g. $\{B_0, B_1\}$ at the first level; $\{B_{00}, B_{01}, B_{10}, B_{11}\}$ at the second level, and so on. Following Lavine (1992), the sets are given by quantiles of the centering family; if m is the base-10 representation of the binary number $\epsilon_1 \dots \epsilon_j$, then $B_{\epsilon_1 \dots \epsilon_j}$ is the interval $(G_\theta^{-1}(m/2^j), G_\theta^{-1}((m+1)/2^j)]$. Let $\Pi = \{\Pi_j, j = 1, 2, \dots\}$ be the sequence of partitions. We also refer to Π as the partition tree and $j = 1, 2, \dots$ as the tree levels.

Define $F(A)$ to be the probability of any set A for distribution F ; note that $F(A)$ is a random variable. The tailfree prior for F is constructed from the sequence of partitions Π and their associated pairwise conditional probabilities $(Y_{\epsilon_1 \dots \epsilon_{j-1}0}, Y_{\epsilon_1 \dots \epsilon_{j-1}1})$, assuming $Y_{\epsilon_1 \dots \epsilon_{j-1}0} = 1 - Y_{\epsilon_1 \dots \epsilon_{j-1}1} = F\{B_{\epsilon_1 \dots \epsilon_{j-1}0} | B_{\epsilon_1 \dots \epsilon_{j-1}}\}$. Let $\mathcal{Y} = \{Y_{\epsilon_1 \dots \epsilon_{j-1}0}, j = 1, 2, \dots\}$. Further, the tailfree prior assumes the random probabilities in \mathcal{Y} are mutually independent, and the random measure F is related to the probabilities through the relation: $F\{B_{\epsilon_1 \dots \epsilon_j}\} = \prod_{i=1}^j Y_{\epsilon_1 \dots \epsilon_{i-1}\epsilon_i}$. Let $\lambda_{\epsilon_1 \dots \epsilon_{j-1}0}$ be the logit transformation of $Y_{\epsilon_1 \dots \epsilon_{j-1}0}$. By assuming $\lambda_{\epsilon_1 \dots \epsilon_{j-1}0}$ has the normal prior $N(0, 2/[c\rho(j)])$, $Y_{\epsilon_1 \dots \epsilon_{j-1}0}$ approximately follows the beta($c\rho(j)$, $c\rho(j)$) distribution (Jara and Hanson, 2011). That is,

$$\text{logit}\{Y_{\epsilon_1 \dots \epsilon_{j-1}0}\} = \lambda_{\epsilon_1 \dots \epsilon_{j-1}0}, \quad \lambda_{\epsilon_1 \dots \epsilon_{j-1}0} \sim N\left(0, \frac{2}{c\rho(j)}\right). \quad (4)$$

The sequence of partitions Π forms a generator of the Borel σ -field of \mathbb{R}^+ and hence for any measurable set $A \in \mathbb{R}^+$, $F(A)$ is defined.

The infinite number of levels in the partition tree Π is usually capped off by some fixed level J , typically $4 \leq J \leq 8$, which yields partitions up to level J , say Π^J . Furthermore, on partition sets $B_{\epsilon_1 \dots \epsilon_j} \in \Pi^J$ at level J we assume F follows the base measure G_θ , i.e. for all measurable $A \subset B_{\epsilon_1 \dots \epsilon_j}$,

$$F\{A | B_{\epsilon_1 \dots \epsilon_j}\} = G_\theta(A) / G_\theta\{B_{\epsilon_1 \dots \epsilon_j}\}. \quad (5)$$

We use $TF^J(c, \rho(\cdot), G_\theta)$ to denote this finite tailfree prior on F with level J . For $F \sim TF^J(c, \rho(\cdot), G_\theta)$, the survival function $S(t) = 1 - F(t)$ is given by

$$S(t) = p\{s(t)\} \{s(t) - 2^J G_\theta(t)\} + \sum_{j=s(t)+1}^{2^J} p(j), \quad (6)$$

where $s(t) = \lceil 2^J G_\theta(t) \rceil$, $\lceil \cdot \rceil$ is the ceiling function. Here $p(j)$, $j = 1, \dots, 2^J$ is defined as

$$p(j+1) = F\{B_{\epsilon_1 \dots \epsilon_j}\} = \prod_{i=1}^j Y_{\epsilon_1 \dots \epsilon_i}, \quad (7)$$

where $\epsilon_1 \dots \epsilon_j$ is the base-2 representation of j . Formula (6) can be obtained from (5) and (7) and

$$F(A) = F\{B_{\epsilon_1 \dots \epsilon_j}\} G_\theta(A) / G_\theta\{B_{\epsilon_1 \dots \epsilon_j}\}$$

for $A \subset B(\epsilon_1 \dots \epsilon_j)$. By differentiating (6), the density with respect to F is given by

$$f(t) = 2^J p\{s(t)\} g_\theta(t), \quad (8)$$

where $g_\theta(\cdot)$ is the density corresponding to G_θ .

A common choice for $\rho(j)$ is j^2 . The parameter c is a precision parameter; lower values of c allow mass of F to move easily from the centering distribution G_θ . As $c \rightarrow 0^+$, $E\{F(\cdot)\}$ tends to the empirical CDF of the data (Hanson and Johnson, 2002); as $c \rightarrow \infty$, all conditional probabilities $\pi(\epsilon)$ go to 0.5 and hence $F(A) \rightarrow G_\theta(A)$ a.s. for all measurable sets. We assign c a gamma prior $c \sim \Gamma(a_c, b_c)$; typically $a = 5$ or 10 and $b = 1$. Alternatively, some authors simply set c as small values, e.g. $c = 1$.

It is well known that fixing θ results in “jumpy” densities as f defined in (8) has discontinuities at each partition interval endpoint. Placing a continuous prior on θ smooths out the posterior density and hazard curves, yielding a mixture of tailfree processes for F (Jara and Hanson, 2011). For the Kijima models, we suggest an empirical approach: an easily-fit special case of the model, e.g. a renewal process or the Poisson process, coupled with the underlying parametric Weibull family G_θ is fitted to obtain the maximum likelihood estimate μ_θ and the inverse information matrix \mathbf{V}_θ associated with μ_θ . A Gaussian prior $N_2(\mu_\theta, \mathbf{V}_\theta)$ is placed on θ . For example, in the first data analysis in Section 5, on the reliability of valve seats, many authors have fit Poisson processes; a Poisson process could be used to center θ . Without such prior knowledge, the first failures of all systems (i.i.d. samples) can be used for a parametric inference on θ .

Note that there is little difference between the standard Polya tree prior and the tailfree process prior for the distribution function of the baseline hazard. Since we use adaptive updating of the logit-transformed conditional probabilities (Section 3.1), it is slightly easier to fit the tailfree version rather than the Polya tree version.

2.2.2. Priors on β

We recommend Zellner’s g -prior (Zellner, 1983) on β , a “reference informative prior”. g -Prior can be used to take into account the correlation among the predictor covariates and has many advantages, as commonly seen in variable selection and linear or nonlinear regressions (Bové and Held, 2011; Marin and Robert, 2007; Fouskakis et al., 2009). Let $\mathbf{W}_j = (\mathbf{w}_{j1}, \dots, \mathbf{w}_{jn_j})$ and $\mathbf{W}_{m^* \times d} = (\mathbf{W}'_1, \dots, \mathbf{W}'_m)'$ where $m^* = \sum_{j=1}^m n_j$. g -Prior for β is then

$$\pi(\beta) \sim N_d(0, g\mathbf{m}^*(\mathbf{W}'\mathbf{W})^{-1}).$$

To avoid choosing g , one can assign g^{-1} a gamma prior $\Gamma(a_g, b_g)$. When $a_g = b_g = 1/2$, the prior on β is a multivariate Cauchy distribution (Zellner and Siow, 1980).

In our simulations, we use a g -prior for the logistic link and obtained excellent performance. We found that the g -prior improves the overall mixing of MCMC chain for both links, but an uninformative prior ($\pi(\beta) \propto 1$) for the logistic link leads to extremely poor MCMC mixing in many data sets. For the exponential link function, we found that uninformative prior shows pretty good and stable performance.

3. Posterior inferences

3.1. MCMC computing

MCMC is used to obtain posterior inferences. The likelihood L is defined in (3) and the prior on β is discussed in Section 2.2.2. Recall that we propose a mixture of tailfree processes prior on F with partitions capped off by J . The prior on θ is defined at the end of Section 2.2.1; the prior on c is $\Gamma(a_c, b_c)$, and the prior on g^{-1} is $\Gamma(a_g, b_g)$. Let $E_0 = \{\epsilon = \epsilon_1 \dots \epsilon_{j-1} 0, j = 1, \dots, J\}$. Each λ_ϵ is assigned a normal prior as detailed in (4). The posterior is then proportional to

$$\pi(\beta, \lambda_\epsilon, c, g, \theta | data) \propto L \cdot \pi(\beta) \Gamma(c | a_c, b_c) \Gamma(g^{-1} | a_g, b_g) \pi(\theta) \prod_{\epsilon \in E_0} N\left(\lambda_\epsilon | 0, \frac{2}{c^2}\right). \quad (9)$$

Parameters $\{\beta, \lambda(\epsilon), \epsilon \in E_0, \theta\}$ are updated using random-walk Metropolis–Hastings updates (Tierney, 1994). We build two blocks to update these parameters. Let \mathbf{b}_1 be a vector of all $\{\lambda_\epsilon, \epsilon \in E_0\}$ with dimension $2^J - 1$ and $\mathbf{b}_2 = (\beta, \theta)$. Gaussian random-walk proposals are used for the two blocks

$$\mathbf{b}'_1 \sim N(\mathbf{b}_1^*, \mathbf{V}_1) \quad \text{and} \quad \mathbf{b}'_2 \sim N(\mathbf{b}_2^*, \mathbf{V}_2),$$

where \mathbf{b}_1^* and \mathbf{b}_2^* are the latest accepted values for \mathbf{b}_1 and \mathbf{b}_2 . We have found automatic tuning of \mathbf{V}_1 and \mathbf{V}_2 to work very well in practice (Haario et al., 2005) leading to proposal acceptance rates in the 20%–50% range as typically desired. Specifically, let the sequence $\mathbf{b}_1^{(1)}, \mathbf{b}_1^{(2)}, \dots$ be the states of the Markov chain for \mathbf{b}_1 . When deciding the t -th state \mathbf{b}_1 , we sample $\mathbf{b}_1^* \sim N(\mathbf{b}_1^{(t-1)}, \mathbf{V}_1^{(t)})$ with

$$\mathbf{V}_1^{(t)} = \begin{cases} \mathbf{V}_1^{(0)}, & t < t_0 \\ s \text{Var} \{ \mathbf{b}_1^{(1)}, \dots, \mathbf{b}_1^{(t-1)} \} + s_0 \mathbf{I}_p, & t > t_0 \end{cases}$$

where p is the dimension of \mathbf{b}_1 , s is recommended to be $2.4^2/p$, s_0 is a small constant, $\mathbf{V}_1^{(0)}$ is the initial covariance of the proposal distribution and \mathbf{I}_p is an identity matrix. A similar automatic tuning procedure applies to \mathbf{b}_2 . The parameter c is updated through the full conditional distribution

$$p(c|\lambda) \sim \Gamma \left\{ (a_c + 2^{J-1} - 1/2), b_c + \sum_{\epsilon_1 \in E_2 \dots \epsilon_j \in E_0} \lambda_{\epsilon_1 \dots \epsilon_j}^2 j^2 / 4 \right\}.$$

The full conditional distribution for g^{-1} given the remaining parameters is $\Gamma(a_g + 1, b_g + \beta' \mathbf{W}' \mathbf{W} \beta / 2m^* + b_g)$. FORTRAN 90 codes for fitting the models in this paper are available from the first author, upon request.

3.2. Model comparison

We compare models using log pseudo-marginal likelihood (LPML) (Geisser and Eddy, 1979), a measure of a model's predictive ability and the deviance information criterion (DIC) (Spiegelhalter et al., 2002), a model selection criterion related to AIC but for use with Bayesian models. Both are easy to compute based on the MCMC output.

Let $\Theta = (\lambda, \theta, \beta)$ and $t_{nj+1,j} = \tau_j$. By definition,

$$\text{LPML} = \sum_{j=1}^m \sum_{i=1}^{n_j+1} \log \{ p(t_{ij} | \mathbf{t}_{-ij}) \},$$

where $p(t_{ij} | \mathbf{t}_{-ij})$ is the predictive density ($\delta_{ij} = 1$) or survival probability ($\delta_{ij} = 0$) for t_{ij} based on the remaining data, $p(\cdot | \mathbf{t}_{-ij})$, evaluated at t_{ij} . This is called the ij -th conditional predictive ordinate (CPO) statistic, and measures how well t_{ij} is predicted from the remaining \mathbf{t}_{-ij} through the model. For system j that has events at $0 < t_{1j} < t_{2j} < \dots < t_{n_j+1,j}$, we compute the predictive density or survival at t_{ij} based on failure and maintenance history for this system during time periods $(0, t_{i-1,j}]$ and $[t_{i+1,j}, \tau_j]$, plus partial information during $(t_{i-1,j}, t_{i+1,j})$ that a certain repair was performed at t_{ij} , plus the information from other systems. That is, to predict $p(t_{ij} | \mathbf{t}_{-ij})$, a repair is still assumed to be done at t_{ij} . The LPML simply aggregates the log of these. For this type of prediction, we are able to share the same form of computing as recommended by Gelfand and Dey (1994). As stated in Section 3.1, the likelihood contribution of failure at t_{ij} depends on repair times and their effectivenesses before t_{ij} for system i , i.e. $t_{ik}, D_{ik}, k = 1, \dots, j-1$. Conditional on Θ , the joint likelihood is $\prod_{j=1}^m \prod_{i=1}^{n_j+1} p(t_{ij} | \mathbf{t}_{1:i-1,j}, \Theta)$.

Following Gelfand and Dey (1994), we have

$$\begin{aligned} p(t_{ij} | \mathbf{t}_{-ij}) &= \int p(t_{ij} | \mathbf{t}_{-ij}, \Theta) \pi(\Theta | \mathbf{t}_{-ij}) d\Theta \\ &= \int p(t_{ij} | \mathbf{t}_{1:i-1,j}, \Theta) \cdot \frac{\prod_{l \neq j} \prod_{k=1}^{n_l+1} p(t_{kl} | \mathbf{t}_{1:k-1,l}, \Theta) \times \prod_{k \neq i}^{n_j+1} p(t_{kj} | \mathbf{t}_{1:k-1,j}, \Theta) \pi(\Theta)}{\int \prod_{l \neq j} \prod_{k=1}^{n_l+1} p(t_{kl} | \mathbf{t}_{1:k-1,l}, \Theta) \times \prod_{k \neq i}^{n_j+1} p(t_{kj} | \mathbf{t}_{1:k-1,j}, \Theta) \pi(\Theta) d\Theta} d\Theta \\ &= \left\{ \int \frac{1}{p(t_{ij} | \mathbf{t}_{1:i-1,j}, \Theta)} \pi(\Theta | \mathbf{t}_{-ij}) d\Theta \right\}^{-1}. \end{aligned}$$

The LPML is then estimated from the MCMC iterates by

$$\text{LPML} = - \sum_{j=1}^m \sum_{i=1}^{n_j+1} \log \left\{ \frac{1}{s} \sum_{k=1}^s \frac{1}{p(t_{ij} | \mathbf{t}_{1:i-1,j}, \Theta^{(k)})} \right\}, \quad (10)$$

where $\Theta^{(k)} = \{\lambda^{(k)}, \theta^{(k)}, \beta^{(k)}, k = 1, 2, \dots, s\}$ are iterates from MCMC outputs of all the parameters.

By definition,

$$\text{DIC} = 2E[D(\Theta|y)] - D(\hat{\Theta}),$$

where $D(\Theta) = -2 \log[L(\Theta)] + C$, $L(\Theta)$ is the likelihood and C is a constant canceled in model comparison. Conditional expectation $E[D(\Theta|y)]$ is typically estimated by averages of $D(\Theta)$ over posterior samples of Θ ; $\hat{\Theta}$ in $D(\hat{\Theta})$ is commonly chosen as the posterior mean of Θ .

3.3. Relaxing the linearity assumption

In this section, we generalize the linear predictor to a flexible additive structure. For simplicity, consider three covariates in the regression: an intercept, a discrete covariate w_{ij1} , and a continuous covariate w_{ij2} . A generalized additive model with the exponential link assumes

$$\log(D_{ij}) = \beta_0 + \beta_1 w_{ij1} + h(w_{ij2}), \quad i = 1, \dots, n_i, \quad j = 1, \dots, m.$$

We approximate the unknown function $h(x)$ using B -splines, i.e.

$$h(x) = \sum_{l=1}^M b_l B_l(x),$$

where $\{B_l(\cdot)\}$ are quadratic B -spline basis functions, defined in De Boor (2001), with support (a, b) , the observed range of w_{ij2} . Since the space spanned by these functions includes the constant term, we let one B -spline coefficient be zero (Gray, 1992) – we choose $b_M = 0$ – under which $h(x)$ equals the constant zero if and only if all the B -spline coefficients are equal to zero. In the following, define $\mathbf{b} = (b_1, \dots, b_{M-1})'$. Note then $h(x) = \sum_{l=1}^{M-1} b_l B_l(x)$. Define $\beta = (\beta_0, b_1, \dots, b_{M-1})'$. The g -prior on β is $\beta \sim N_M(\mathbf{0}, ng(\mathbf{X}'\mathbf{X})^{-1})$ where \mathbf{X} is the design matrix. The above extension can be fit using the algorithm developed in Section 3.

For equally-spaced knots, $\sum_{l=1}^{M-1} b_l B_l(x) = \beta_1 x$ for some β_1 when $b_{l-1} + b_{l+1} - 2b_l = 0$ for $l = 2, \dots, M-2$. Define $\Delta = (b_1 + b_3 - 2b_2, b_2 + b_4 - 2b_3, \dots, b_{M-3} + b_{M-1} - 2b_{M-2}) = \Phi \mathbf{b}$ where Φ is a $(M-3) \times (M-1)$ matrix. Suppose MCMC iterates for \mathbf{b} are $\mathbf{b}^{(k)}$, $k = 1, \dots, s$. To test whether $h(x)$ is linear in x is equivalent to testing the point null $H_0 : \Phi \mathbf{b} = \mathbf{0}$. Bayes factors against the null hypothesis can be computed using the Savage–Dickey ratio (Verdinelli and Wasserman, 1995),

$$\text{BF} \approx \frac{N_{M-3}(\mathbf{0}|\mathbf{0}, ng(\mathbf{X}'\mathbf{X})^{-1}\Phi')}{N_{M-3}(\mathbf{0}|\mathbf{m}, \mathbf{V})},$$

where $\mathbf{m} = s^{-1} \sum_{k=1}^s \Phi \mathbf{b}^{(k)}$ and $\mathbf{V} = s^{-1} \sum_{k=1}^s (\Phi \mathbf{b}^{(k)} - \mathbf{m})(\Phi \mathbf{b}^{(k)} - \mathbf{m})'$. A larger BF value indicates stronger evidence against the null hypothesis.

4. Simulations

We perform simulations to examine the proposed models and the Bayesian nonparametric method. Suppose m systems are included in each simulated sample and each system is maintained up to its 5th failure, yielding a total number of events $m^* = 5m$. The associated event (failure) times are recorded as t_{ij} , $i = 1, 2, \dots, 5$, $j = 1, 2, \dots, m$. At each event time, a type of repair is performed with effectiveness according to the Kijima type I or type II model. The degree of effectiveness D_{ij} is $\text{logit}(\beta' \mathbf{w}_{ij})$ or $\exp(\beta' \mathbf{w}_{ij})$ where \mathbf{w}_{ij} includes $w_{ij0} = 1$, w_{ij1} a Bernoulli(0.5) and w_{ij2} simulated so that D_{ij} follows uniform $(0, 1)$. The true baseline distribution is $0.5 \text{ Weibull}(2, 2) + 0.5 \text{ Weibull}(2, 4)$ for simulations in Tables 1 and 3 and has a corresponding hazard $\exp(t^2/3 + t/3)$ for simulations in Table 2. Coefficients are set to $\beta = (-1, 1, 1)$ or $(1, -1, 1)$ and the sample size is $m^* = 300$ or 500. For each setup, 300 datasets are simulated and fitted with the following model and prior specifications. The baseline distribution F is given the tailfree prior with $J = 5$, $c \sim \Gamma(5, 1)$, and $\theta \sim N(\mu_\theta, \mathbf{V}_\theta)$ where μ_θ is the maximum likelihood estimate of θ based on a parametric Weibull fit of the first failures of all systems and \mathbf{V}_θ the inverse information matrix associated with μ_θ . Optimization routines in R or SAS give θ and \mathbf{V}_θ . For the logistic link, regression parameter β is given the g -prior $N_3(0, gn(\mathbf{W}'\mathbf{W})^{-1})$ with $g^{-1} \sim \text{Exp}(1)$. For the exponential link, the flat prior $\pi(\beta) \propto 1$ is used. Following algorithms in Section 3.1, we run 30 000 iterations for each MCMC chain and thin the posterior samples by taking every fifth of them after a burn-in of 10 000 iterates. Each chain takes a few minutes with a 3.00 GHz processor.

Simulation results are presented in Tables 1–3 including the average of the posterior means over 300 datasets, the sample standard deviation SSD of the posterior means, the average of the estimated standard deviations ESE and 95% the coverage probability CP. Based on the simulation results, the true parameters β are estimated with little bias. As the sample size increases, both SSD and SSE decrease. We also get coverage probabilities close to the nominal level 0.95. For one simulation setup posterior means for the baseline density and survival functions (gray lines) are plotted in Fig. 1, overlaid with the true density or survival functions in black lines.

We also perform simulations to examine the additive model described in Section 3.3. Assume $D_{ij} = \exp(\beta_0 + \beta_1 w_{ij1} - w_{ij2}^2)$ or $\text{logit}(\beta_0 + \beta_1 w_{ij1} - w_{ij2}^2)$ where w_{ij1} is sampled from Bernoulli(0.5) and w_{ij2} from uniform $(-1, 1)$. Data are simulated

Table 1Summary of simulation studies: $f = 0.5 \text{ Weibull}(2, 2) + 0.5 \text{ Weibull}(2, 4)$; link function is logistic.

True	$n = 300$				$n = 500$			
	Point	SSD	ESE	95% CP	Point	SSD	ESE	95% CP
Type I								
$\beta_0 = -1$	-0.92	1.10	1.17	0.95	-1.04	0.75	0.82	0.95
$\beta_1 = 1$	1.05	1.44	1.69	0.95	1.04	1.03	1.19	0.96
$\beta_2 = 1$	0.97	0.62	0.73	0.95	0.97	0.43	0.52	0.95
$\beta_0 = 1$	1.06	1.24	1.56	0.95	1.02	1.07	1.22	0.93
$\beta_1 = -1$	-0.95	1.30	1.72	0.96	-0.96	1.16	1.28	0.94
$\beta_2 = -1$	-0.94	0.55	0.74	0.98	-1.01	0.51	0.60	0.95
Type II								
$\beta_0 = -1$	-0.98	0.81	0.89	0.94	-1.02	0.60	0.60	0.92
$\beta_1 = 1$	0.97	1.19	1.26	0.95	1.00	0.89	0.87	0.93
$\beta_2 = 1$	1.00	0.48	0.56	0.95	1.01	0.39	0.40	0.95
$\beta_0 = 1$	1.02	1.01	1.14	0.95	0.96	0.85	0.80	0.93
$\beta_1 = -1$	-1.08	1.11	1.27	0.96	-0.96	0.89	0.85	0.93
$\beta_2 = -1$	-1.01	0.50	0.58	0.94	-0.99	0.41	0.40	0.94

Table 2Summary of simulation studies: $r(t) = \exp(t^2/3 + t/3)$; link function is logistic.

True	$n = 300$				$n = 500$			
	Point	SSD	ESE	95% CP	Point	SSD	ESE	95% CP
Type I								
$\beta_0 = -1$	-1.00	0.37	0.37	0.96	-1.01	0.29	0.29	0.94
$\beta_1 = 1$	1.02	0.50	0.53	0.96	1.04	0.39	0.40	0.94
$\beta_2 = 1$	1.03	0.27	0.28	0.96	1.02	0.20	0.20	0.95
$\beta_0 = 1$	1.10	0.51	0.54	0.94	1.05	0.38	0.39	0.96
$\beta_1 = -1$	-1.04	0.53	0.55	0.93	-1.02	0.38	0.40	0.95
$\beta_2 = -1$	-1.05	0.26	0.29	0.95	-1.03	0.19	0.21	0.97
Type II								
$\beta_0 = -1$	-1.00	0.36	0.34	0.93	-1.02	0.29	0.27	0.93
$\beta_1 = 1$	0.99	0.45	0.43	0.94	1.02	0.35	0.33	0.93
$\beta_2 = 1$	1.01	0.22	0.22	0.95	1.02	0.18	0.17	0.94
$\beta_0 = 1$	1.04	0.36	0.37	0.93	1.01	0.27	0.27	0.95
$\beta_1 = -1$	-1.04	0.43	0.44	0.94	-1.03	0.31	0.33	0.96
$\beta_2 = -1$	-1.03	0.21	0.22	0.96	-1.01	0.17	0.17	0.94

Table 3Summary of simulation studies: $f = 0.5 \text{ Weibull}(2, 2) + 0.5 \text{ Weibull}(2, 4)$; link function is exponential.

True	$n = 300$				$n = 500$			
	Point	SSD	ESE	95% CP	Point	SSD	ESE	95% CP
Type I								
$\beta_0 = -1$	-1.09	0.51	0.49	0.92	-1.01	0.37	0.34	0.92
$\beta_1 = 1$	1.02	0.66	0.75	0.96	1.01	0.57	0.52	0.92
$\beta_2 = 1$	1.09	0.45	0.52	0.96	1.08	0.41	0.38	0.92
$\beta_0 = 1$	0.94	0.81	0.86	0.95	0.92	0.58	0.62	0.94
$\beta_1 = -1$	-1.01	0.71	0.73	0.94	-0.99	0.47	0.51	0.96
$\beta_2 = -1$	-1.07	0.45	0.50	0.96	-1.02	0.34	0.36	0.94
Type II								
$\beta_0 = -1$	-1.10	0.35	0.35	0.93	-1.01	0.28	0.27	0.91
$\beta_1 = 1$	0.94	0.51	0.56	0.96	0.98	0.43	0.43	0.92
$\beta_2 = 1$	1.04	0.34	0.40	0.95	1.04	0.30	0.31	0.94
$\beta_0 = 1$	0.91	0.62	0.69	0.95	1.02	0.55	0.52	0.93
$\beta_1 = -1$	-1.00	0.55	0.62	0.97	-1.02	0.50	0.46	0.92
$\beta_2 = -1$	-1.09	0.40	0.43	0.95	-1.07	0.35	0.33	0.93

from the Kijima type I model with baseline hazard $r(t) = \exp(t^2/3 + t/3)$. We consider a sample size $m^* = 1000$ and each setup has 300 replications. We take $M = 6$ equally spaced quadratic B -splines to model the effect of w_{ij2} and one B -spline coefficient is set to be zero. Table 4 summarizes estimates for coefficients β_0 , β_1 and Fig. 2 plots the point-wise mean, 2.5% and 97.5% quantiles of the estimates for the true function $h(w) = -w^2$.

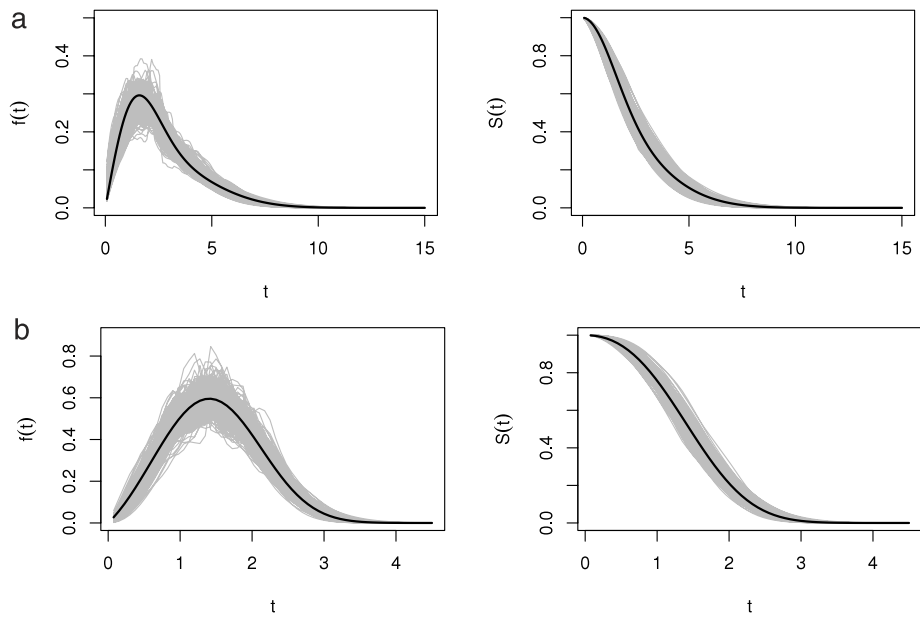


Fig. 1. Density and survival estimates for the simulated data sets with (a) $f = 0.5 \text{ Weibull}(2, 2) + 0.5 \text{ Weibull}(2, 4)$ and (b) $r(t) = \exp(t^2/3 + t/3)$ based on Kijima type I model; the dark lines are the true density or survival functions and the gray lines are the point-wise posterior means.

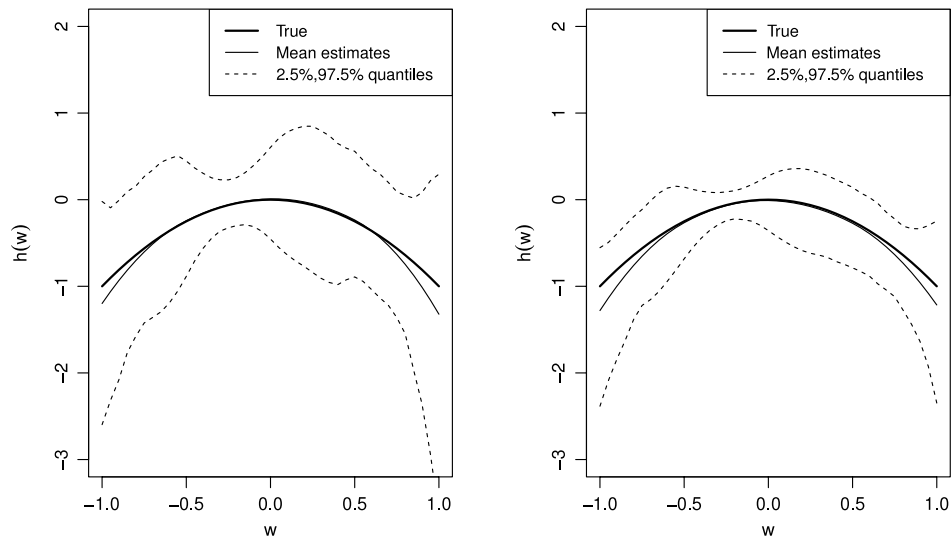


Fig. 2. Mean (gray-solid lines), 2.5% and 97.5% quantiles (gray-dashed lines) of the estimates for $h(w)$ based on the simulated datasets from the type I model with logistic (left) and exponential (right) links. The black solid lines are the true function $h(w) = -w^2$.

Table 4
Summary of simulation studies: $r(t) = \exp(t^2/3 + t/3)$; $h(w_{ij2}) = \sum_{l=1}^5 b_l B_{l3}(w_{ij2})$.

True	Logistic				Exponential			
	Point	SSD	ESE	95% CP	Point	SSD	ESE	95% CP
$\beta_0 = -1$	-1.09	0.38	0.37	0.93	-1.06	0.23	0.24	0.96
$\beta_1 = 1$	1.07	0.23	0.21	0.91	-1.07	0.17	0.17	0.92

5. Data analysis

We first consider the dataset analyzed in Lawless and Nadeau (1995) that gives the times of replacing valve seats on 41 diesel engines in a service fleet. A few successive repairs recorded on the same day are deleted. We assume the end of history-time is independent of the event process, as concluded in Lawless and Nadeau (1995). On the left panel of Fig. 3,

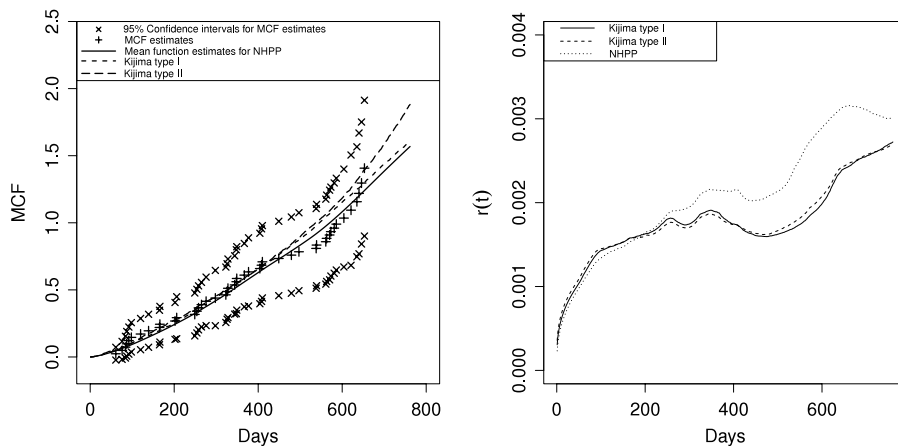


Fig. 3. Plots for valve seats maintenance data; left panel is the MCF plot for data where '+' is the empirical point estimate and 'x' is its 95% confidence interval, overlaid with estimates of the mean function for NHPP (solid line), Kijima type I (short-dashed line) and Kijima type II (long-dashed line); right panel is the estimated baseline hazard function for Kijima type I (solid line), Kijima type II (dashed line) and NHPP (dotted line).

Table 5
Summaries of β_0 for Kijima type I and type II models for the valve seats maintenance data.

Kijima model	J	$\pi(c)$	$\pi(\beta_0)$	$\hat{\beta}_0(95\% \text{ CI.})$	$P(\beta_0 > 0)$	LPML	DIC
Type I	4	$\Gamma(5, 1)$	$N(0, 3^2)$	1.11 (−1.73, 3.06)	0.92	−334.0	665.4
	5	$\Gamma(5, 1)$	$N(0, 3^2)$	1.07 (−1.79, 2.98)	0.93	−334.9	665.2
	4	$\Gamma(5, 1)$	$N(0, 2^2)$	1.08 (−1.47, 2.76)	0.93	−334.0	665.0
	5	$\Gamma(5, 1)$	$N(0, 2^2)$	1.04 (−1.48, 2.61)	0.93	−334.1	664.0
	5	$\Gamma(10, 1)$	$N(0, 2^2)$	0.96 (−1.78, 2.85)	0.88	−334.9	668.0
Type II	4	$\Gamma(5, 1)$	$N(0, 3^2)$	0.80 (−2.69, 2.89)	0.89	−334.7	667.2
	5	$\Gamma(5, 1)$	$N(0, 3^2)$	0.79 (−2.46, 2.66)	0.89	−334.5	666.2
	4	$\Gamma(5, 1)$	$N(0, 2^2)$	0.78 (−1.36, 2.23)	0.90	−334.6	666.3
	5	$\Gamma(5, 1)$	$N(0, 2^2)$	0.84 (−1.43, 2.39)	0.91	−334.5	665.7
	5	$\Gamma(10, 1)$	$N(0, 2^2)$	0.64 (−1.77, 2.27)	0.85	−335.1	669.5

a nonparametric estimate for the mean cumulative function (MCF) (Nelson, 1995) is plotted, where each '+' is a point estimate of MCF at its corresponding time and 'x's are its associated 95% confidence interval. Note that MCF is defined as $E[N(t)]$. When the event process is NHPP, MCF equals the mean function $\int_0^t r(s)ds$. The dataset has been fitted many times in the literature assuming NHPP. In this work, Kijima type I, Kijima type II, and NHPP models ($D_{ij} = 1$) are fitted for the data using the proposed method. The baseline distribution F is given a tailfree prior: $J = 4$ or 5 ; $c \sim \Gamma(5, 1)$ or $\Gamma(10, 1)$; $\theta \sim N_2(\mu_\theta, \mathbf{V}_\theta)$ where μ_θ and \mathbf{V}_θ are obtained from a fit of the Poisson process assuming Weibull for the baseline distribution. For the Kijima models, we choose exponential link for the age reduction factor, i.e. $D_{ij} = \exp(\beta_0)$. Two sets of priors are considered for β_0 : $N(0, 2^2)$ and $N(0, 3^2)$.

Based on the results in Table 5, both Kijima type I and type II models show high probabilities for D_{ij} being greater than 1, i.e. posterior $P(\beta_0 > 0)$. There is little difference in both estimation of β_0 and goodness-of-fit measures (LPML, DIC) when increasing the tailfree level J . The prior favoring lower c shows slightly better values for LPML and DIC and has some effects on estimation of β_0 due to a less weight of the centering Weibull family. The prior on β_0 with larger variance results in wider 95% credible intervals but the point estimate of β_0 remains stable. Under $J = 5$, $c \sim \Gamma(5, 1)$ and $\pi(\beta_0) \sim N(0, 2^2)$, the estimated baseline hazards $\hat{r}(t)$ (right panel of Fig. 3) are nondecreasing in general and with slight decreases around 500 days. Interpretation of the age reduction is then related to effectiveness of the repairs, i.e. the repairs have high probability of being worse than "bad as old", explaining to some extent the rapid increase of failures around 600 days. To compare data fits, estimates of $E(N(t))$ for NHPP, Kijima type I and type II models are also plotted on the left panel in Fig. 3 where estimates for Kijima types I and II are based on simulated failure times using the posterior means of β_0 and the baseline survival function (Krivtsov, 2000; Veber et al., 2008). The plot shows that MCF estimates based on all three of NHPP, Kijima type I, and type II interpolate the nonparametric MCF estimates well, except for slight differences in the tail. Under $J = 5$, $c \sim \Gamma(5, 1)$, the nonparametric fit of the NHPP model has LPML and DIC −336.0 and 666.6 respectively. Little difference can be found in the goodness-of-fit measures, compared to those in Table 5. Nevertheless, our models provide information on the age reduction factor which is helpful for understanding the repairs and future modeling. We also fit parametric Kijima type I and II models assuming Weibull baseline distribution and obtain LPML as −334.6 and −334.7 and DIC as 669.4 and 669.6 respectively. Therefore, the simpler Weibull family would be an adequate choice for the baseline distribution.

Table 6

Summary of the coefficients for Kijima type I and type II models for Syringe-driver maintenance data; $J = 5$, link function is exponential, estimates are posterior means, and 95% CIs are credible intervals.

Coefficient	$\pi(c)$	Type I $\hat{\beta}$ (95% CI)	Type II $\hat{\beta}$ (95% CI)
	$\Gamma(10, 1)$	$\beta \sim N(\mathbf{0}, \text{gn}(\mathbf{W}'\mathbf{W})^{-1})$; –0.13 (–1.96, 1.13) 1.29 (0.46, 2.31) –0.20 (–0.78, 0.25)	$g \sim \text{Exp}(1)$ –0.36 (–0.65, –0.06) 0.05 (–0.04, 0.19) –0.02 (–0.16, 0.11)
β_0 (pm)	$\Gamma(10, 1)$	$\beta \sim N(\mathbf{0}, 3^2 \mathbf{I}_3)$ 0.04 (–3.47, 2.18)	–0.36 (–0.66, –0.07)
β_1 (cm)		2.19 (0.67, 4.44)	0.05 (–0.05, 0.19)
β_2 (cost)		–0.28 (–1.03, 0.31)	–0.02 (–0.14, 0.11)
	$\Gamma(5, 1)$	$\beta \sim N(\mathbf{0}, 3^2 \mathbf{I}_3)$ –0.06 (–4.4, 2.55) 2.25 (0.65, 5.17) –0.29 (–1.16, 0.29)	–0.34 (–0.63, –0.06) 0.04 (–0.05, 0.16) 0.01 (–0.12, 0.13)
β_0 (pm)	$\Gamma(5, 1)$		
β_1 (cm)			
β_2 (cost)			

Table 7

Goodness-of-fit measures for Kijima type I and type II model for Syringe-driver maintenance data; $J = 5$, and link function is exponential.

	$\pi(c)$	$\pi(\beta)$	Type I	Type II
LPML	$\Gamma(10, 1)$	g-prior	–313.2	–319.2
LPML	$\Gamma(10, 1)$	$N(\mathbf{0}, 3^2 \mathbf{I}_3)$	–312.2	–319.1
LPML	$\Gamma(5, 1)$	$N(\mathbf{0}, 3^2 \mathbf{I}_3)$	–311.4	–316.8
DIC	$\Gamma(10, 1)$	g-prior	624.2	631.3
DIC	$\Gamma(10, 1)$	$N(\mathbf{0}, 3^2 \mathbf{I}_3)$	621.8	631.1
DIC	$\Gamma(5, 1)$	$N(\mathbf{0}, 3^2 \mathbf{I}_3)$	616.4	627.2

The second dataset includes failures and repairs of 12 syringe-driver pumps (Baker, 1991; Singh, 2011). Most systems are maintained up until 106 months and we assume that this censoring time is independent of the failure processes. The pumps receive preventive maintenances (pm mode) and corrective maintenances (cm mode) with 48 pms and 94 cms in total. The cost for each cm is also available and we consider it as a covariate interacting with corrective maintenance mode only. Denote \mathbf{w} as the covariate vector including: $w_0 = 1$ if the maintenance is pm and 0 otherwise, $w_1 = 1$ if the maintenance is cm and 0 otherwise and w_2 is the cost of the cm repair. Kijima type I and type II models are fitted to the data. The baseline distribution F is given a tailfree prior: $J = 5$; $c \sim \Gamma(10, 1)$ or $\Gamma(5, 1)$; $\theta \sim N_2(\mu_\theta, \mathbf{V}_\theta)$ where μ_θ and \mathbf{V}_θ are obtained from a parametric fit of the first events of the 12 pumps. The coefficient vector $\beta = (\beta_0, \beta_1, \beta_2)$ is considered with g-prior and other Gaussian priors. We assume the exponential link for both models.

Table 6 shows the point estimates (posterior means) and 95% credible intervals for β . For the Kijima type I model, the g-prior results in narrower credible intervals for the intercept β_0 and shrinks β_1 toward zero. For Kijima type II model, there is little difference in the estimates and credible intervals by using different priors on β . The effective age reduction factor due to preventive maintenance in type II models is significantly less than 1, by exponentiating the estimate of the intercept β_0 , indicating that preventive maintenances are better than “bad as old” repairs. Baker (1991) also observes that the preventive maintenances are very effective in maintaining the systems. Type I models does not show strong evidence for β_0 less than zero. Across all fitting specifications, the corrective repairs perform significantly worse than the preventive repairs and the cost of the corrective repairs shows no significant effect. Table 7 presents estimates of Goodness-of-fit measures LPML and DIC. Since larger LPML or lower DIC indicates a better fit of the data, the results show that type I models fit the data slightly better than type II models; $\Gamma(5, 1)$ yields a better fit than $\Gamma(10, 1)$.

We also fit the Kijima type I and II models with the underlying Weibull baseline ($c \rightarrow \infty$) and $\beta \sim N(\mathbf{0}, 3^2 \mathbf{I}_3)$. The type I model has LPML –314.6 and DIC 628.7. The type II model has LPML –319.1 and DIC 636.6. The nonparametric Bayes method gives slightly better fits to the data than the parametric Weibull models. The differences are not significant, suggesting adequacy of the Weibull assumption for the data. Fig. 4 also shows that the parametric estimates for the baseline density and hazard functions (smooth lines) stay in the credible intervals (dashed lines) of the nonparametric estimates.

Finally, we fit the Kijima type I model with $h(w_2)$ modeled by a B-spline as outlined in Section 3.3. Let $M = 6$ and set one B-spline coefficient as zero for model identifiability. Now $\beta = (\beta_0, \beta_1, b_1, \dots, b_5)$ and let $\beta \sim N(\mathbf{0}, \text{ng}(\mathbf{X}'\mathbf{X})^{-1})$ where \mathbf{X} is the new design matrix and g is fixed at 0.5 or 1. For the tailfree prior, let $J = 5$ and $c \sim \Gamma(10, 1)$. Based on Table 8, point estimates for β_0 and β_1 are close to those fitted with linearity assumption for $h(w_2)$ but credible intervals are much wider due to an increased number of parameters. Fig. 5 plots the point-wise estimates (solid lines) and 95% credible intervals (dashed lines) for $h(w_2)$ showing no significant nonlinear trend. Bayes factors for the test of linearity of $h(w_2)$ are less than one and hence the linear assumption is preferred.

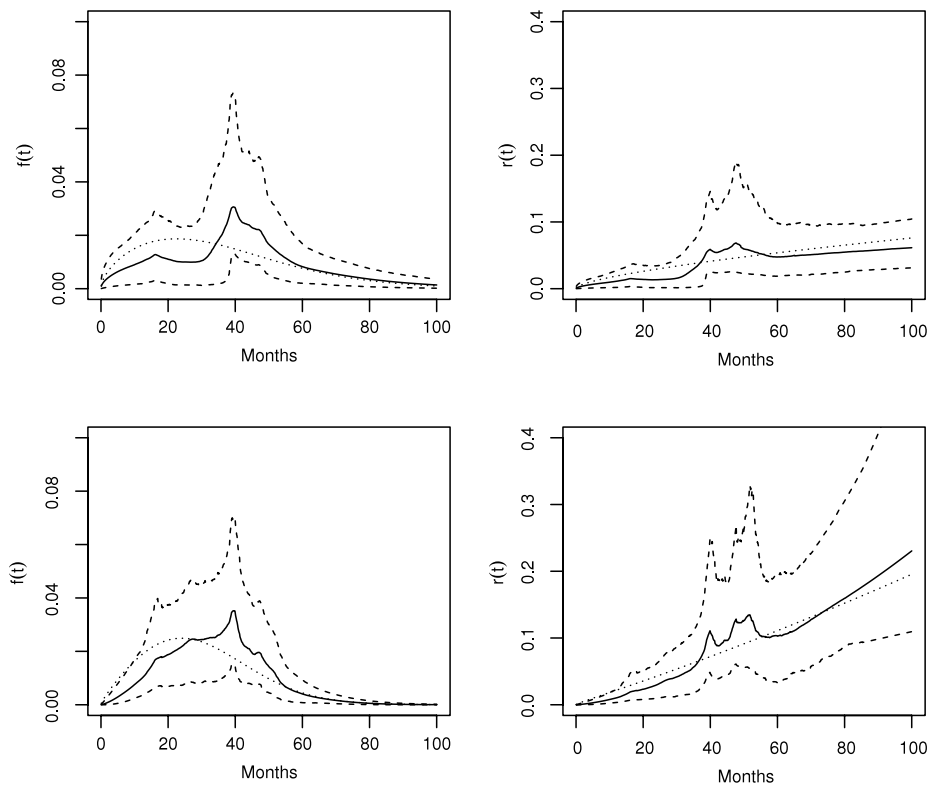


Fig. 4. Plots for Syringe-driver maintenance data for Kijima type I (left) and type II (right) model; solid lines are baseline density (hazard) estimates and dashed lines are 95% credible intervals. Smooth estimates (dotted lines) are fitted from the parametric Weibull fit.

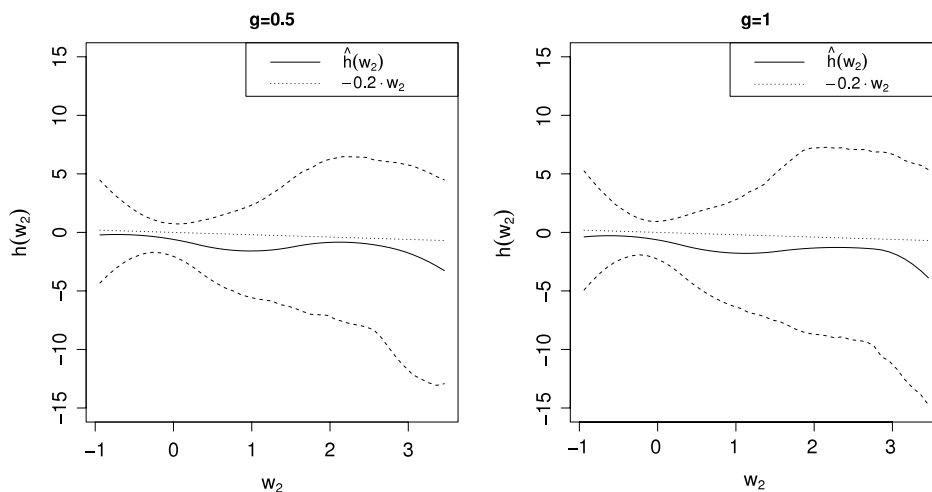


Fig. 5. Estimate (solid) of $h(w_2)$ and its 95% credible intervals (dashed) for type I model with logistic (exponential) link on the left (right) panel; dotted lines are linear functions, i.e. $h(w_2) = -0.2w_2$ on the left panel.

Table 8

Summary of the coefficients for Kijima type I model for Syringe-driver maintenance data; $h(w_2)$ is approximated by a B -spline, estimates are posterior means, and 95% CI are credible intervals.

Coefficient	$g = 0.5$ $\hat{\beta}$ (95% CI)	$g = 1$ $\hat{\beta}$ (95% CI)
β_0 (pm)	-0.12 ($-1.5, 1.17$)	-0.04 ($-2.12, 4.63$)
β_1 (cm)	2.02 ($-2.02, 8.06$)	2.53 ($-5.37, 9.43$)
β_2 (cost)	–	–

6. Discussion

We proposed a new semiparametric regression model for recurrent events arising from maintenances of repairable systems where effectiveness of repairs characterized by covariates are taken into account in the joint modeling. We generalized the Kijima effective age models (Kijima, 1989) by regressing the age reduction factors on covariates. The baseline distribution is flexibly modeled using a tailfree prior, which generalizes the commonly-used Weibull family allowing for data-driven flexibility. Logistic and exponential links are proposed for the regression and efficient, adaptive, and easy-to-implement MCMC is described. The proposed method was illustrated using simulations and two data analyses. We found useful and interesting interpretations of regression coefficients when examining the effect of covariates on the effectiveness of repairs. When the link is the exponential function, the proposed semiparametric regression model provides an easy test for the common assumption of minimal repair (“bad as old repair”) which is also appealing to practitioners.

The regression parameters are interpretable since an increase or decrease of effective age is closely related to intensity of the system. However when the hazard of the system is not monotone, the interpretation becomes more difficult. Finally, we note that it is also straightforward to generalize our model to include heterogeneous systems by including random system effects in the linear predictor or times the intensity function.

References

- Baker, R.D., 1991. Testing the efficiency of preventive maintenance. *J. Oper. Res. Soc.* 42, 493–503.
- Bové, D.S., Held, L., 2011. Hyper-g priors for generalized linear models. *Bayesian Anal.* 6, 1–24.
- Cox, D.R., 1972. Regression models and life tables (with discussion). *J. R. Stat. Soc. Ser. B* 34, 187–220.
- De Boor, C., 2001. A Practical Guide to Splines. In: *Applied Mathematical Sciences*, vol. 27. Springer.
- Dorado, C., Hollander, M., Sethuraman, J., 1997. Nonparametric estimation for a general repair model. *Ann. Statist.* 25, 1140–1160.
- Doyen, L., Gaudoin, O., 2004. Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliab. Eng. Syst. Saf.* 84, 45–56.
- Escobar, M.D., West, M., 1995. Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* 90, 1301–1312.
- Ferguson, T., 1973. A Bayesian analysis of some nonparametric problems. *Ann. Statist.* 1, 209–230.
- Ferguson, T., 1974. Prior distribution on the spaces of probability measures. *Ann. Statist.* 2, 615–629.
- Fouskakis, D., Ntzoufras, I., Draper, D., 2009. Bayesian variable selection using cost-adjusted BIC, with application to cost-effective measurement of quality of health care. *Ann. Appl. Stat.* 3, 663–690.
- Freedman, D., 1963. On the asymptotic distribution of Bayes' estimates in the discrete case. *Ann. Math. Statist.* 34, 1386–1403.
- Geisser, S., Eddy, W., 1979. A predictive approach to model selection. *J. Amer. Statist. Assoc.* 74, 153–160.
- Gelfand, A.E., Dey, D.K., 1994. Bayesian model choice: asymptotics and exact calculations. *J. R. Stat. Soc. Ser. B* 56, 501–514.
- Gray, R.J., 1992. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *J. Amer. Statist. Assoc.* 87, 942–951.
- Haario, H., Saksman, E., Tamminen, J., 2005. Componentwise adaptation for high dimensional MCMC. *Comput. Statist.* 20, 265–273.
- Hanson, T., 2006. Inference for mixtures of finite Polya tree models. *J. Amer. Statist. Assoc.* 101, 1548–1565.
- Hanson, T., Johnson, W., 2002. Modeling regression error with a mixture of Polya trees. *J. Amer. Statist. Assoc.* 97, 1020–1033.
- Hanson, T., Yang, M., 2007. Bayesian semiparametric proportional odds models. *Biometrics* 63, 88–95.
- Ishwaran, H., James, L.F., 2004. Computational methods for multiplicative intensity models using weighted gamma processes. *J. Amer. Statist. Assoc.* 99, 175–190.
- Jara, A., Hanson, T., 2011. A class of mixtures of dependent tailfree processes. *Biometrika* 98, 553–566.
- Kijima, M., 1989. Some results for repairable systems with general repair. *J. Appl. Probab.* 26, 89–102.
- Krivtsov, V., 2000. A Monte Carlo approach to modeling and estimation of the generalized renewal process in repairable system reliability analysis. Dissertation for the Degree of Doctor of Philosophy, University of Maryland.
- Kuo, L., Ghosh, S.K., 1997. Bayesian Nonparametric Inference for Nonhomogeneous Poisson Processes. Technical Report 9718. University of Connecticut, Department of Statistics.
- Lavine, M., 1992. Some aspects of Polya tree distributions for statistical modeling. *Ann. Statist.* 20, 1222–1235.
- Lawless, J.F., Nadeau, C., 1995. Some simple robust methods for the analysis of recurrent events. *Technometrics* 37, 158–168.
- Lawless, J.F., Thiagarajah, K., 1996. A point process model incorporating renewals and time trends, with application to repairable systems. *Technometrics* 38, 131–138.
- Lindqvist, B.H., 2006. On the statistical modeling and analysis of repairable systems. *Statist. Sci.* 21, 532–551.
- Lo, A.Y., 1992. Bayesian inference for Poisson process models with censored data. *J. Nonparametr. Stat.* 2, 71–80.
- Marin, J.-M., Robert, C.P., 2007. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Springer-Verlag, New York.
- Nelson, W., 1995. Confidence limits for recurrence data applied to cost or number of product repairs. *Technometrics* 37, 147–157.
- Peña, E., Slate, E., González, J., 2007. Semiparametric inference for a general class of models for recurrent events. *J. Statist. Plann. Inference* 137, 1727–1747.
- Singh, I., 2011. Risk-averse periodic preventive maintenance optimization. Doctoral Dissertation, The University of Texas at Austin.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B* 64, 583–639.
- Taddy, M., Kottas, A., 2012. Mixture modeling for marked Poisson processes. *Bayesian Anal.* 7, 335–362.
- Tierney, L., 1994. Markov chains for exploring posterior distributions (with discussion 1728–1762). *Ann. Statist.* 22, 1701–1728.
- Veber, B., Nagode, M., Fajdiga, M., 2008. Generalized renewal process for repairable systems based on finite Weibull mixture. *Reliab. Eng. Syst. Saf.* 93, 1461–1472.
- Verdinelli, I., Wasserman, L., 1995. Computing Bayes factors using a generalization of the Savage–Dickey density ratio. *J. Amer. Statist. Assoc.* 90, 614–618.
- Walker, S., Mallick, B.K., 1999. A Bayesian semiparametric accelerated failure time model. *Biometrics* 55, 477–483.
- Yuan, F., Uday, K., 2012. A general imperfect repair model considering time-dependent repair effectiveness. *IEEE Trans. Reliab.* 61, 95–100.
- Zellner, A., 1983. Applications of Bayesian analysis in econometrics. *Statistician* 32, 23–34.
- Zellner, A., Siow, A., 1980. Posterior odds ratios for selected regression hypotheses. In: Bernardo, J.M., DeGroot, M.H., Lindley, D.V., Smith, A.F.M. (Eds.), *Bayesian Statistics: Proceedings of the First International Meeting Held in Valencia*. pp. 585–603.