Statistical Inference for Stochastic Processes (2005) 8:311–329
DOI 10.1007/s11203-005-6100-y

On Modeling Change Points in Non-Homogeneous Poisson Processes

FABRIZIO RUGGERI[a,*] and SIVA SIVAGANESAN[b]
[a] *CNR-IMATI, Milano, Italy*
[b] *University of Cincinnati, USA. e-mail: siva@math.uc.edu*

Abstract. Failures in repairable systems are often described by means of non-homogeneous Poisson processes, identified by their intensity and mean value functions. Intervention on the systems are likely to modify their reliability, and changes in intensities and mean value functions are therefore induced. We consider different scenarios in which interventions take places and propose models describing each of them. Bayesian analyses, relying on Markov-chain Monte Carlo methods, are illustrated along with applications to simulated and real, widely-known, data.

Key words: Bayesian inference, power law process, Markov-chain Monte Carlo methods, change points.

1. Introduction

Repairable systems are those systems or machines which, in the event of a failure, can be repaired and returned to regular operation. In some cases, the reliability of a system, after a repair or an "intervention", returns to the same state as before repair. This condition is commonly called "bad-as-old" or "same-as-old".

Failures of such repairable systems are often described by means of non-homogeneous Poisson processes (NHPP), mainly the Power Law process (PLP); see, among others Ascher and Feingold (1984), Bain and Engelhardt (1991), Crow (1974), Crow (1982), Thompson (1988) and, for a review of Bayesian literature on PLP, Bar-Lev et al. (1992). Other NHPPs have been considered in, e.g. Pievatolo et al. (2003) and Pievatolo and Ruggeri (2004). A general review of NHPP's and their applications to the reliability of repairable systems can be found in Rigdon and Basu (2000).

The NHPP, such as PLP typically assume that the reliability of a system evolves continuously over time. When reliability of a repairable system is modeled using such a NHPP, it is implicitly assumed that the reliability

*Author for correspondence: Fax: +39 02 23699538; e-mail: fabrizio@mi.imati.cnr.it

of the system remains the same before and after the repair, i.e., the intervention on the system in the form of a repair does not affect the reliability of the system. In many cases, however, the reliability of a system, after an intervention such as a repair, may be different to that before the intervention, i.e., the reliability may improve or decline due to an intervention.

It is therefore of interest to model the reliability of a system in a way that allows changes in reliability due to interventions on the system. These interventions may involve known ones such as repairs, and also involve interventions that occur within the system that may not be known to the investigator but affect a change in the reliability of the system.

In this paper, we focus on modeling changes in the reliability due to interventions of the system by allowing changes (or discontinuities) in the intensity function of such NHPPs, focusing mainly on the PLP. Specifically, we consider two different types of change point models. In the first, we consider models that allow changes in reliability level after each failure, as the the system is repaired and put to operation, e.g., in software reliability. In the second, we consider a model that allow changes at random points in time, due to breakdown of a component without causing the failure of system or due to interventions by the maintenance squad at unknown time points. Change points in homogeneous Poisson processes (HPPs) were considered by Raftery and Akman (1986) and Green (1995).

In Section 2, we consider a general class of NHPPs, as well as some specific models including the PLP. In Sections 3 and 4, we mainly focus on PLP. In Section 3, we consider both a hierarchical and a dynamic model for the case of change points at each failure; the models are then applied to simulated data sets and compared. In Section 4, we expand on the previous models by allowing change points at an unknown subset of failure times, rather than at each failure time, to allow the possibility that change in reliability does not occur at every failure time. In Section 5, we model changes in reliability at random number of change points and at random locations. Here, we use the reversible jump Markov-chain Monte Carlo (RJMCMC) method and provide the details for general NHPP as well as some specific ones described in Section 2. In Section 6, we illustrate the application of the proposed methods to simulated and real data, and end the paper with some remarks.

2. A class of Non-homogeneous Poisson Processes

The NHPP are identified by their intensity function $\lambda(t;\theta)$ and/or their mean value function $\nu(t;\theta)=\int_0^t \lambda(u;\theta)\mathrm{d}u$. Suppose we observe the system up to time y and let n be the number of failures, occurred at times $t_1 < t_2 < \cdots < t_n$; then the likelihood function is given by

$$L(\theta; \mathbf{t}) = \prod_{i=1}^{n} \lambda(t_i) \exp\left\{ -\int_0^y \lambda(t)\mathrm{d}t \right\},$$

where $\mathbf{t} = (t_1, \ldots, t_n)$.

A general class of NHPPs can be described by their intensity function $\lambda(t; M, \beta) = Mg(t, \beta)$, with $M, \beta > 0$, such that their mean value function is $\nu(t; M, \beta) = MG(t, \beta)$, with $G(t, \beta) = \int_0^t g(u, \beta)\mathrm{d}u$. This class contains well-known processes, such as the Musa–Okumoto, the Cox–Lewis and the Power Law processes.

The first process described in Musa and Okumoto (1984) has been widely used in modeling software reliability; it has intensity function $\lambda(t; M, \beta) = M/(t + \beta)$ and mean value function $\nu(t; M, \beta) = M \log(t + \beta)$. The second process described in Cox and Lewis (1966) is such that $\lambda(t; M, \beta) = M \exp\{\beta t\}$ and $\nu(t; M, \beta) = (M/\beta)[\exp\{\beta t\} - 1]$.

In this paper, we focus on the Power law process (see, e.g. Ascher and Feingold, 1984), but other processes can also be considered likewise. The intensity and mean value functions of a Power law process are given, respectively, by $\lambda(t; M, \beta) = M\beta t^{\beta-1}$ and $\nu(t; M, \beta) = Mt^{\beta},\ M, \beta > 0$. Note that $M = \nu(1; M, \beta)$ denotes the expected number of failures up to time $t = 1$, whereas β determines how reliability decays or grows. For $\beta < 1$, the intensity function decreases over time and the reliability grows, as a consequence. The HPP is obtained by taking $\beta = 1$, whereas $\lambda(t)$ is an increasing concave (straight, convex) curve if $1 < \beta < 2$ ($\beta = 2,\ \beta > 2$).

There has been much interest in using PLP where the value of β varies over time. Previous research has used PLP where the value of β is allowed to change at two fixed time-points. This model allows for three different stages of reliability: first and initial stage where reliability improves ($\beta < 1$) up to a time t_1, a second stage where the reliability remains the same ($\beta = 1$) over a period between t_1 and t_2, and a third and final stage where the reliability of the system declines ($\beta > 1$) after time t_2.

We consider different scenarios in which interventions take place and propose models that allow change points in reliability to occur. Specifically, we consider models which: adaptively adjust for change points allowing change in β at each failure time; detect change points allowing change in β at random (failure) times.

3. Changes at each Failure Time

Suppose that interventions occur right after each failure (at times t_i^+'s), modifying the value of β. Changes in M can be considered in a similar, but cumbersome, manner, and so we do not pursue that here. We denote the parameter value at time t_i^+, $i = 1, \ldots, n$, right after a failure, by β_i, identifying the process over $(t_i, t_{i+1}]$. We denote by β_0 the parameter value over

$(t_0, t_1]$. Here we take $t_0 = 0$ and $t_{n+1} = y$, i.e. the endpoints of the observation interval.

3.1. SIMULATED DATA SETS

To illustrate our methodology we use, among other, two simulated data sets obtained using a PLP with change points in the values of β. Plots of failure times versus failure number are given in Figures 1 and 2. In simulated data set 1, there is a change point (in the value of β) after 11th failure, and in Simulated Data Set 2, there are two change points, one after fifth failure and the other after 13th failure. Note, however, from Figure 2 that there also appears to be a (unintended) change point at the 8th failure time.

3.2. A HIERARCHICAL MODEL

When the system is fairly stable over time, it may be reasonable to assume that β_i's are similar. This can be done by using a hierarchical model as follows. In the first stage, given (ϕ, σ^2),

$$\beta_i \text{ are i.i.d. } \mathcal{LN}(\phi, \sigma^2), \quad i = 0, \ldots, n$$

and, in the second stage, ϕ and σ^2 have $\mathcal{N}(\mu, \tau^2)$ and $\mathcal{IG}(\rho, \gamma)$ distributions, respectively. ($\mathcal{LN}, \mathcal{N}$ and $\mathcal{IG}$ denote, respectively, lognormal, normal and inverse gamma distributions.)

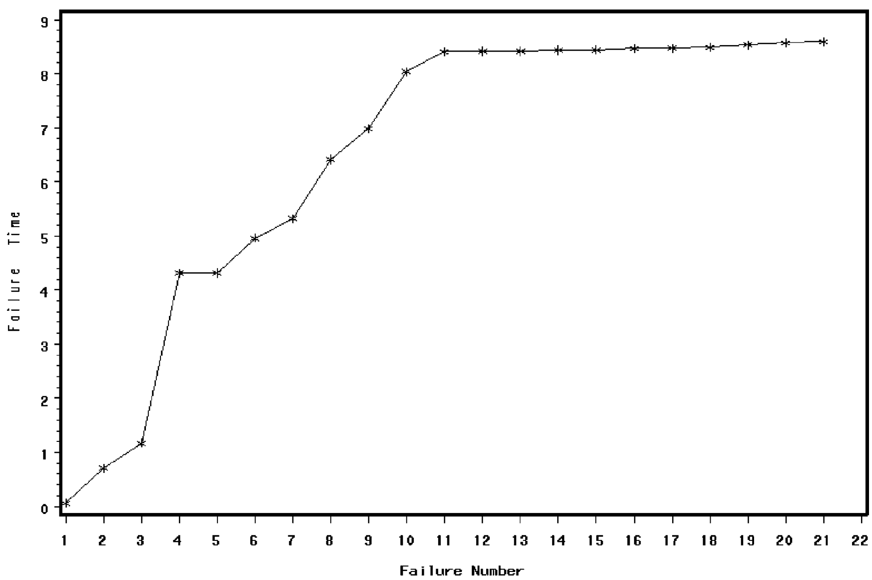

Figure 1. Simulated data set 1: $\beta = 0.7$ for $t < 11$, $\beta = 2$ for $t \geqslant 11$.

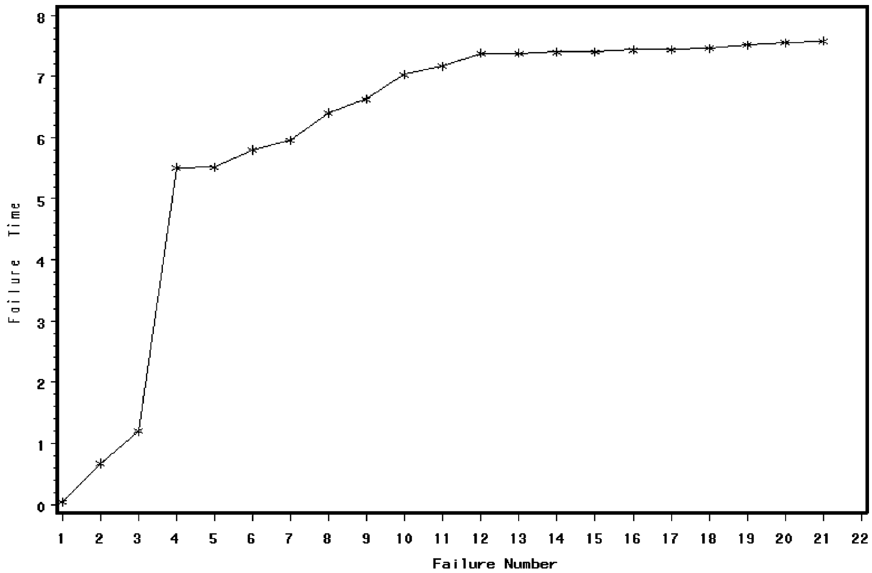

Figure 2. Simulated data set 2: $\beta=0.6$ for $t<5$, $\beta=1$ for $6\leqslant t<13$, $\beta=2$ for $t\geqslant 14$.

Note that the likelihood based on the PLP (from the likelihood for a general NHPP given in Section 2), is

$$L(M,\beta,a,\sigma^2;\mathbf{t})=M^n\prod_{i=1}^{n}\beta_{i-1}t_i^{\beta_{i-1}-1}\exp\left\{-M\sum_{i=1}^{n+1}\left(t_i^{\beta_{i-1}}-t_{i-1}^{\beta_{i-1}}\right)\right\}.$$

Assuming a gamma prior $\mathcal{G}(\alpha,\delta)$ for M, one can write down the following full conditionals (given the rest of the parameters we omit).

- $M\sim\mathcal{G}\big(\alpha+n,\delta+\sum_{i=1}^{n+1}[t_i^{\beta_{i-1}}-t_{i-1}^{\beta_{i-1}}]\big)$,
- $\sigma^2\sim\mathcal{IG}(\rho+n/2,\tau+(1/2)\sum_{i=0}^{n}(\log\beta_i-\phi)^2)$,
- $\phi\sim\mathcal{N}(\mu_1,\tau_1^2)$, where $\mu_1=(\mu\sigma^2+\tau^2\sum_{i=0}^{n}\log\beta_i)/(\sigma^2+(n+1)\tau^2)$ and $\tau_1^2=\tau^2\sigma^2/(\tau^2(n+1)+\sigma^2)$,
- $\beta_i\propto t_{i+1}^{\beta_i}\exp\{-M(t_{i+1}^{\beta_i}-t_i^{\beta_i})-[(\log\beta_i-\phi)^2]/(2\sigma^2)\}, i=0,\dots,n-1$,
- $\beta_n\propto\beta_n^{-1}\exp\big\{-M(t_{n+1}^{\beta_n}-t_n^{\beta_n})-[(\log\beta_n-\phi)^2]/(2\sigma^2)\big\}$.

One can use the above full conditionals to simulate from the joint posterior via Metropolis–Hastings and Gibbs sampling.

3.3. A DYNAMIC MODEL

We suppose that at time t_i^+, $i=1,\dots,n$, right after a failure, the parameter β_{i-1}, identifying the process over $(t_{i-1},t_i]$, is modified according to

$$\log \beta_i = \log a + \log \beta_{i-1} + \epsilon_i, \tag{1}$$

where β_i is the value of β over $(t_i, t_{i+1}]$, a is a positive constant and ϵ_i is a normally distributed random variable with mean 0 and variance σ^2 (heteroskedasticity would be possible but it would imply no relevant conceptual difference).

Let β_0 be the value of β over $(0,t_1]$, $\beta = (\beta_0,\dots,\beta_n)$ and $\beta_{(-i)} = (\beta_0,\dots,\beta_{i-1},\beta_{i+1},\dots,\beta_n)$. We set $t_0=0$ and $t_{n+1}=y$, as before.

Consider a gamma prior on M, i.e. $M \sim \mathcal{G}(\alpha,\delta)$, an inverse gamma prior $\mathcal{IG}(\rho,\tau)$ on σ^2, and two lognormal distributions, $\mathcal{LN}(\mu,\sigma^2)$ and $\mathcal{LN}(\phi,\sigma^2)$, on $a|\sigma^2$ and $\beta_0|\sigma^2$, respectively. Note that we could have taken independent priors on both a and β_0 and the posterior distributions, although more cumbersome, would be slightly affected. The dependence upon σ^2, besides its mathematical convenience, can be justified by the relations among the β_i's.

Under the previous choices for the prior, it is possible to obtain the full conditional distributions for M, σ^2 and a, whereas they are known apart from a constant for β_i, $i=0,\dots,n$.

In particular, we obtain

- $M|\beta,\sigma^2,a,\mathbf{t} \sim \mathcal{G}\Big(\alpha+n, \delta+\sum_{i=1}^{n+1}[t_i^{\beta_{i-1}} - t_{i-1}^{\beta_{i-1}}]\Big)$,
- $\sigma^2|M,\beta,a,\mathbf{t} \sim \mathcal{IG}(\rho+n/2+1, \tau+(1/2)\sum_{i=1}^n \log^2(\beta_i/(a\beta_{i-1}))$ $+ (\log a-\mu)^2 + (\log\beta_0-\phi)^2)$,
- $a|M,\beta,\sigma^2,\mathbf{t} \sim \mathcal{LN}([\mu+\log\beta_n - \log\beta_0]/(n+1), \sigma^2/(n+1))$,
- $\beta_n|M,\beta_{(n)},\sigma^2,a,\mathbf{t} \propto (1/\beta_n)\exp\{-M(y^{\beta_n}-t_n^{\beta_n}) - [(\log\beta_n - \log a$ $-\log\beta_{n-1})^2\big]/(2\sigma^2)\}$,
- $\beta_0|M,\beta_{(0)},\sigma^2,a,\mathbf{t} \propto t_1^{\beta_0}\exp\{-Mt_1^{\beta_0} - \big[(\log\beta_1-\log a-\log\beta_0)^2$ $+(\log\beta_0-\phi)^2\big]/(2\sigma^2)\}$,
- $\beta_i|M,\beta_{(i)},\sigma^2,a,\mathbf{t} \propto t_{i+1}^{\beta_i}\exp\{-M(t_{i+1}^{\beta_i}-t_i^{\beta_i}) - [(\log\beta_i-\log a$ $-\log\beta_{i-1})^2 + (\log\beta_{i+1}-\log a-\log\beta_i)^2\big]/(2\sigma^2)\}, i=1,\dots,n-1$.

We can now use a Gibbs algorithm with Metropolis steps to sample the posterior distribution.

3.4. EXAMPLES

We now use the above models on the data sets discussed in Section 3.1. In the following tables we give the estimates of β_i's and their standard deviations. Here, β_i stands for the value of β after the ith failure. Our goal here is simply to see how well the proposed models capture the (known) change points in the simulated data sets, rather than to evaluate the model fit using a formal method. Table I give the estimates of β_i's for Simulated Data Set 1, using the hierarchical and dynamic models. We used vague priors by choosing $\alpha = \delta = .01$, $\rho = \tau = 0.01$, and $\mu = \phi = 0$. When prior information is available, it may be incorporated in the prior distribution (see Campodonico and Singpurwalla, 1995) for an illustration of this in the context of Poisson point process.

Note that the dynamic model estimates of the parameters β_i capture the true values better than the hierarchical model, which is reasonable as the

Table I. Failure number (i) and posterior mean and std. dev. of β_i for Simulated Data Set 1: $\beta = 0.7$ for $t < 11$, $\beta = 2$ for $t \geqslant 11$, using the hierarchical and the dynamic models

	Hierarchical model		Dynamic model	
Failure number	Mean (β_i)	Std. Dev.	Mean (β_i)	Std. Dev.
1	2.58	1.67	0.86	0.57
2	1.8	1.26	0.81	0.48
3	0.66	0.35	0.68	0.33
4	2.9	0.89	1.1	0.67
5	1.01	0.41	0.92	0.37
6	1.14	0.43	0.94	0.38
7	0.86	0.35	0.85	0.31
8	1.01	0.37	0.93	0.33
9	1.81	0.55	0.99	0.31
10	2.17	0.6	1.92	0.54
11	2.04	0.59	2.28	0.47
12	4.35	0.94	3.68	0.98
13	1.88	0.56	2.26	0.43
14	2.55	0.68	2.57	0.55
15	1.77	0.54	2.08	0.44
16	2.45	0.65	2.41	0.54
17	1.89	0.56	2.01	0.46
18	1.6	0.51	1.67	0.45
19	1.16	0.54	1.35	0.51

values of β used in the Simulated Data Set 1 belong to two separate time periods of equal values, a phenomenon not consistent with the assumption under hierarchical model. Note that the dynamic model also seems to adapt to a temporary (unintended) change near the fourth failure time.

Table II gives the estimates of β_i's for Simulated Data 2 using the dynamic model. We note that the estimates remain fairly steady over the duration with no (true) change points and adapt well to the (known) change points.

Suppose that the interventions occur at random points, instead of right after each failure as assumed above. Assuming, for instance, that intervention times T_i's follow a HPP, we can update β's as in (1), except that now the changes occur at T_j's instead of t_i's.

4. Changes at a Random Number of Failures

Often, it may be desirable to allow change points in the values of β only as needed, rather than at every failure time as done in Section 3. This may

Table II. Failure number (i) and posterior mean and std. dev. of β_i for Simulated Data Set 2: $\beta=0.6$ for $t<5$, $\beta=1$ for $6\leqslant t<13$, $\beta=2$ for $t\geqslant 14$, using the dynamic model

Failure number	Mean (β_i)	Std. Dev.
1	0.82	0.55
2	0.78	0.43
3	0.68	0.30
4	1.28	0.77
5	1.13	0.39
6	1.20	0.41
7	1.11	0.36
8	1.18	0.38
9	1.17	0.36
10	1.37	0.39
11	1.51	0.37
12	3.02	1.16
13	2.19	0.44
14	2.52	0.56
15	2.09	0.44
16	2.39	0.54
17	2.03	0.47
18	1.70	0.46
19	1.41	0.53

be achieved by modeling whether a change occurs at a failure time as a Bernoulli random event. Thus, we define $\mathbf{Z}=(Z_1,\ldots,Z_n)$, such that $Z_i=1$ implies that the ith failure time t_i is a change point, whereas $Z_i=0$ implies that it is not a change point. Thus, we may write

$$\log\beta_i=\log\beta_{i-1}+\epsilon_i$$

with ϵ_i having $\mathcal{N}(0,\sigma^2)$ if $Z_i=1$ and point mass δ_0 at 0, otherwise.

Instead, we prefer a computationally simpler form based on George and McCulloch (1993), and let ϵ_i have $\mathcal{N}(0,\sigma^2)$ if $Z_i=1$, and ϵ_i have $\mathcal{N}(0,w^2\sigma^2)$, if $Z_i=0$, where w is very small. Here, $P(Z_i=1)=p$ may be fixed, or a prior could be given on it.

We consider a lognormal model $\mathcal{LN}(\phi,\sigma^2)$ for $\beta_0|\sigma^2$ (we prefer a larger value of variance, with respect to $w^2\sigma^2$, to draw the first β), and an inverse gamma prior $\mathcal{IG}(\rho,\tau)$ for σ^2.

Using this model, the conditional posterior $p(\theta|\mathbf{t},\mathbf{Z})$ is proportional to

$$M^n\prod_{i=1}^n\beta_{i-1}t_i^{\beta_{i-1}-1}\exp\left\{-M\sum_{i=0}^n\left[t_{i+1}^{\beta_i}-t_i^{\beta_i}\right]\right\}$$
$$\cdot\prod_{i=1}^n p^{Z_i}(1-p)^{1-Z_i}\cdot\prod_{i=1}^n(\sqrt{2\pi}\sigma_i)^{-1}\exp\{-[\log\beta_i-\log\beta_{i-1}]^2/(2\sigma_i^2)\}$$
$$\cdot(\sqrt{2\pi}\sigma\beta_0)^{-1}\exp\{-[\log\beta_0-\phi]^2/(2\sigma^2)\},$$

where

$$\sigma_i=\sigma\quad\text{if } Z_i=1\quad\text{and}\quad\sigma_i=w\sigma\quad\text{if } Z_i=0.$$

Consider a gamma prior on M, i.e. $M\sim\mathcal{G}(\alpha,\delta)$, an inverse gamma prior $\mathcal{IG}(\rho,\tau)$ on σ^2.

Under the previous choices for the prior, it is possible to obtain the full conditional distributions for M, and σ^2, whereas they are known apart from a constant for β_i, $i=0,\ldots,n$. In particular, we obtain, conditional on the rest of the parameters:

- $M\sim\mathcal{G}(\alpha+n,\delta+\sum_{i=0}^n[t_{i+1}^{\beta_i}-t_i^{\beta_i}])$
- $\sigma^2\sim\mathcal{IG}(\rho+(1/2)(\#\{Z_i=1\})+1,\tau+(1/2)\left[\sum_{i:Z_i=1}(\log(\beta_i/\beta_{i-1}))^2\right.$ $\left.+\sum_{i:Z_i=0}(\log(\beta_i/\beta_{i-1}))^2/w^2+(\log\beta_0-\phi)^2\right])$,
- $\beta_0\propto t_1^{\beta_0}\exp\{-Mt_1^{\beta_0}-(\log\beta_1-\log\beta_0)^2/(2\sigma_1^2)$ $-(\log\beta_0-\phi)^2/(2\sigma^2)\}$,
- $\beta_i\propto t_{i+1}^{\beta_i}\exp\{-M[t_{i+1}^{\beta_i}-t_i^{\beta_i}]-(\log\beta_i-\log\beta_{i-1})^2/(2\sigma_i^2)$ $-(\log\beta_{i+1}-\log\beta_i)^2/(2\sigma_{i+1}^2)\}$, $\quad i=1,\ldots,n-1$,

Table III. Failure number (i) and probability of change point at i for simulated data set 1: $\beta = 0.7$ for $t < 11$, $\beta = 2$ for $t \geqslant 11$

Failure number	Prob. change point
2	0.07
3	0.14
4	0.07
5	0.05
6	0.04
7	0.02
8	0.02
9	0.01
10	0.02
11	0.94
12	0.01
1	0.01
3	0.01
14	0.01
15	0.01
16	0.01
17	0.01

- $\beta_n \propto (1/\beta_n) \exp\{-M \left[t_{n+1}^{\beta_n} - t_n^{\beta_n}\right] - (\log \beta_n - \log \beta_{n-1})^2/(2\sigma_n^2)\}$,
- $Z_i \sim$ Bernoulli(p_i),

where

$$p_i = \frac{p f_{LN}(\beta_i; \beta_{i-1}, \sigma^2)}{\left[p f_{LN}(\beta_i; \beta_{i-1}, \sigma^2) + (1-p) f_{LN}(\beta_i; \beta_{i-1}, w^2)\right]}$$

and $f_{LN}(\cdot; \mu, \sigma^2)$ is the pdf of the lognormal distribution $\mathcal{LN}(\mu, \sigma^2)$.

The Beta $\mathcal{B}e(\mu, \nu)$ on p is updated to $\mathcal{B}e(\mu + \sum Z_i, \nu + \sum(1 - Z_i))$. We can sample the posterior using a Metropolis within Gibbs algorithm as earlier. Tables III and IV show that the method is able to give high probability of being a change point to the actual ones.

In the examples below, we used vague priors with $\alpha = \delta = 0.01$, $\rho = 3$, $\tau = 2$, and $\phi = 0$, $p = 0.5$, and used $w^2 = 0.001$. The choice of w^2 affected the answers to the extent that choosing a larger value for w^2 still identified the same change points but the associated probability was somewhat smaller. In practice, the choice of w^2 may be chosen to reflect the size of change one deems warrants a change point.

Table IV. Failure number (i) and probability of change point at i for simulated data set 2: $\beta = 0.6$ for $t < 5$, $\beta = 1$ for $6 \leqslant t < 13$, $\beta = 2$ for $t \geqslant 14$

Failure number	Prob. change point
2	0.11
3	0.11
4	0.07
5	0.81
6	0.22
7	0.04
8	0.03
9	0.03
10	0.03
11	0.04
12	0.04
13	0.86
14	0.51
15	0.04
16	0.04
17	0.02
18	0.05

5. Changes at a Random Number of Points

Our main premise is that change points in reliability occur when there is an intervention. In the above, we considered intervention due to a failure. However, it is plausible that certain other type of interventions that may not be directly observable, such as an internal failure, may occur at random, causing a change in reliability. Thus, it may be desirable to allow change points at random time points. We follow Green (1995), where a reversible jump algorithm has been used to detect change points in a Poisson process whose intensity is a step function. We consider NHPPs with intensity $\lambda(t; M, \beta) = Mg(t, \beta)$, as described in Section 2, and parameters M and β modified after each change point.

We refer to the paper by Green for a thorough illustration of the reversible jump technique and discussion on the choice of the moves in the step function case. Here we mainly stress the differences with his approach and outline the main features of the method we follow.

5.1. A PRIOR MODEL

We suppose that k, the number of change points, is drawn from a truncated Poisson distribution $p(k)$, $k=0,\ldots,K$, with parameter ω. We choose K to be significantly smaller than the number of observed failures. A truncated geometric distribution could have been chosen as well, affecting the selection of the possible moves to be described later, but the results should be only very slightly affected by the choice between the two priors.

Conditional upon k, the change points T_j are such that $0<T_1<\cdots<T_k<y$, whereas the parameters of the intensity function $\lambda(t;M,\beta)$ take the values M_j and β_j on the subinterval (T_j,T_{j+1}), $j=0,\ldots,k$ (assuming $T_0=0$ and $T_{k+1}=y$).

As in Green (1995), we assume that the k change points $T_1,\ldots,T_k$ are distributed as the even-numbered order statistics from $2k+1$ independent, uniformly distributed r.v.'s on $(0,y]$. The choice of $2k+1$ r.v.'s instead of k is justified by the desire of avoiding "small" subintervals.

We suppose that both the parameters $\beta_0,\ldots,\beta_k$ and $M_0,\ldots,M_k$ are independently drawn from Gamma priors, the former ones from $\mathcal{G}(\alpha,\delta)$ whereas the latter ones from $\mathcal{G}(\epsilon,\phi)$.

5.2. REVERSIBLE JUMPS

As described in Green (1995) transitions from a set of parameter values to another occur according to reversible moves. Here we consider four possible moves:

- **[P]** change to the parameters M and β at a randomly chosen change point T_j;
- **[L]** change to the location of a randomly chosen change point;
- **[B]** "birth" of a new change point at a randomly chosen location in $(0,y]$;
- **[D]** "death" of a randomly chosen change point.

At each transition, we attempt one of the four possible moves, choosing randomly one of them. Depending only on the number k of change points, the probability of choosing the move **[P], [L], [B]** or **[D]** is given by π_k,η_k,b_k and d_k, respectively. Note that $\pi_k+\eta_k+b_k+d_k=1$ for all k, $d_0=\eta_0=\pi_0=0$ and $b_K=0$, where K is the maximum number of allowed change points.

We consider $b_k=c\min\{1,p(k+1)/p(k)\}$ and $d_k=\min\{1,p(k)/p(k+1)\}$, so that $b_k p(k)=d_{k+1}p(k+1)$. Green suggests choosing $\pi_k=\eta_k$ for $k\neq 0$ and c as large as possible so that $b_k+d_k\leqslant 0.9$, for all $k=0,\ldots,K$. We follow his suggestions.

5.2.1. *Move of Type* **[P]**

We first randomly choose the index $j \in \{0, \dots, k\}$ identifying the parameters M_j and β_j to be changed; then we propose new values $M_j^{'}$ and $\beta_j^{'}$ drawn such that both $\log(M_j^{'}/M_j)$ and $\log(\beta_j^{'}/\beta_j)$ are uniformly distributed on $[-0.5, 0.5]$.

Consider the likelihood ratio $p(\mathbf{t}|\mathbf{M}_{(j)}, \beta_{(j)}, M_j^{'}, \beta_j^{'})/p(\mathbf{t}|\mathbf{M}_{(j)}, \beta_{(j)}, M_j, \beta_j)$ for a general NHPP with intensity $\lambda(t; M, \beta) = Mg(t, \beta)$, as described in Section 2. Under this move, the likelihood ratio depends only on what happens in $[T_j, T_{j+1})$ since everything else cancels out; thus the ratio becomes

$$LR_P = (M_j^{'}/M_j)^{|I_j|} \prod_{t_i \in I_j} [g(t_i, \beta_j^{'})/g(t_i, \beta_j)] \tag{2}$$

$$\cdot \exp\left\{-M_j^{'}[G(T_{j+1}, \beta_j^{'}) - G(T_j, \beta_j^{'})] + M_j[G(T_{j+1}, \beta_j) - G(T_j, \beta_j)]\right\}, \tag{3}$$

where $I_j = \{t_i : t_i \in [T_j, T_{j+1})\}$, $|I_j|$ is the size of I_j and the product over the t_is in I_j equals one when $|I_j| = 0$.

As an example, the likelihood ratio for the PLP becomes

$$LR_P = (M_j^{'}\beta_j^{'}/M_j\beta_j)^{|I_j|} \prod_{t_i \in I_j} t_i^{\beta_j^{'} - \beta_j} \exp\left\{-M_j^{'}[T_{j+1}^{\beta_j^{'}} - T_j^{\beta_j^{'}}] + M_j[T_{j+1}^{\beta_j} - T_j^{\beta_j}]\right\}.$$

Similar results are obtained when considering other processes.

The acceptance probability for the move turns out to be

$$\min\{1, LR_P \cdot (M_j^{'}/M_j)^{\epsilon} \exp\{-\phi(M_j^{'} - M_j)\} \cdot (\beta_j^{'}/\beta_j)^{\alpha} \exp\{-\delta(\beta_j^{'} - \beta_j)\}\}.$$

5.2.2. *Move of Type* **[L]**

We first randomly choose the index $j \in \{1, \dots, k\}$ identifying the change point T_j to be moved; then we propose a new location $T_j^{'}$ drawing it from a uniform distribution on $[T_{j-1}, T_{j+1}]$.

Consider the likelihood ratio $p(\mathbf{t}|\mathbf{M}, \beta, \mathbf{T}_{(j)}, T_j^{'})/p(\mathbf{t}|\mathbf{M}, \beta, \mathbf{T}_{(j)}, T_j)$ for a general NHPP with intensity $\lambda(t; M, \beta) = Mg(t, \beta)$, as described in Section 2. Under this move, the likelihood ratio depends only on what happens between T_j and $T_j^{'}$ since everything else cancels out; thus the ratio becomes

$$LR_L = \{(M_{j+1}/M_j)^{|I_j|} \prod_{t_i \in I_j} \left[g(t_i, \beta_{j+1})/g(t_i, \beta_j)\right]\}^{\mathrm{sgn}(T_j - T_j^{'})} \tag{4}$$

$$\cdot \exp\{-M_{j+1}\left[G(T_j, \beta_{j+1}) - G(T_j^{'}, \beta_{j+1})\right] + M_j\left[G(T_j, \beta_j) - G(T_j^{'}, \beta_j)\right]\}, \tag{5}$$

where $I_j = \{t_i : t_i \in (\min\{T_j, T_j^{'}\}, \max\{T_j, T_j^{'}\}]\}$, $\mathrm{sgn}(a)$ equals $a/|a|$ for $a \neq 0$ and 0 otherwise, and the product over the t_is in I_j equals one when $|I_j| = 0$.

As an example, the likelihood ratio for the PLP becomes

$$LR_L = \left\{ (M_{j+1}\beta_{j+1}/M_j\beta_j)^{|I_j|} \prod_{t_i \in I_j} t_i^{\beta_{j+1}-\beta_j} \right\}^{\mathrm{sgn}(T_j - T_j')} \cdot$$

$$\exp\{-M_{j+1}[T_j^{\beta_{j+1}} - T_j'^{\beta_{j+1}}] + M_j[T_j^{\beta_j} - T_j'^{\beta_j}]\}.$$

The acceptance probability for the move turns out to be

$$\min\{1, LR_L \cdot [(T_{j+1} - T_j')(T_j' - T_{j-1})]/[(T_{j+1} - T_j)(T_j - T_{j-1})]\}.$$

5.2.3. *Moves of Type* **[B]** *and* **[D]**

For a birth ([B]), we draw a new position T^* from a uniform distribution on $(0, y]$. It lies in $(T_j, T_{j+1}]$, for some j, where the parameters are β_j and M_j. The new parameters, (β_j', M_j') in $(T_j, T^*]$ and (β_{j+1}', M_{j+1}') in $(T^*, T_{j+1}]$, are chosen to preserve the weighted geometric mean, as in Green (1995),

$$(T^* - T_j)\log\beta_j' + (T_{j+1} - T^*)\log\beta_{j+1}' = (T_{j+1} - T_j)\log\beta_j \tag{6}$$

and the expected number of failures in $(T_j, T_{j+1}]$,

$$\begin{aligned} &M_j'[G(T^*, \beta_j') - G(T_j, \beta_j')] + M_{j+1}'[G(T_{j+1}, \beta_{j+1}') - G(T^*, \beta_{j+1}')] \\ &\quad = M_j[G(T_{j+1}, \beta_j) - G(T_j, \beta_j)]. \end{aligned} \tag{7}$$

The condition (7) becomes

$$\begin{aligned} &M_j'[T^{*\beta_j'} - T_j^{\beta_j'}] + M_{j+1}'[T_{j+1}^{\beta_{j+1}'} - T^{*\beta_{j+1}'}] \\ &\quad = M_j[T_{j+1}^{\beta_j} - T_j^{\beta_j}] \end{aligned}$$

for the PLP,

$$\begin{aligned} &(M_j'/\beta_j')[\exp\{T^*\beta_j'\} - \exp\{T_j\beta_j'\}] \\ &\quad + (M_{j+1}'/\beta_{j+1}')[\exp\{T_{j+1}\beta_{j+1}'\} - \exp\{T^*\beta_{j+1}'\}] \\ &\quad = (M_j/\beta_j)[\exp\{T_{j+1}\beta_j\} - \exp\{T_j\beta_j\}] \end{aligned}$$

for the Cox–Lewis process and

$$\begin{aligned} &M_j' \log[(T^* + \beta_j')/(T_j + \beta_j')] + M_{j+1}' \log[(T_{j+1} + \beta_{j+1}')/(T^* + \beta_{j+1}')] \\ &\quad = M_j \log[(T_{j+1} + \beta_j)/(T_j + \beta_j)] \end{aligned}$$

for the Musa–Okumoto process.

We define the perturbations of the parameters to be such that

$$(\beta'_{j+1}/\beta'_j) = (1-u)/u \quad \text{and} \quad (M'_{j+1}/M'_j) = (1-w)/w, \tag{8}$$

with u and w independent and drawn uniformly from [0, 1].

The acceptance probability for a "birth" move turns out to be

$$\min\{1, LR_B \cdot \text{Prior ratio} \cdot \text{Proposal ratio} \cdot \text{Jacobian}\},$$

where the likelihood ratio LR_B is similar to the previous ones, the prior ratio is

$$\frac{p(k+1)}{p(k)} \frac{2(k+1)(2k+3)}{y^2} \frac{(T^*-T_j)(T_{j+1}-T^*)}{T_{j+1}-T_j} (\delta^\alpha / \Gamma(\alpha))[\beta'_j \beta'_{j+1}/\beta_j]^{\alpha-1}$$

$$\cdot \exp\{-\delta(\beta'_j + \beta'_{j+1} - \beta_j)\}(\phi^\epsilon / \Gamma(\epsilon))[M'_j M'_{j+1}/M_j]^{\epsilon-1} \exp\{-\phi(M'_j + M'_{j+1} - M_j)\}$$

the proposal ratio is

$$\frac{d_{k+1} y}{b_k(k+1)},$$

whereas the Jacobian is given by

$$\frac{(\beta'_j + \beta'_{j+1})^2}{\beta_j} \frac{(M'_j + M'_{j+1})^2}{M_j}. \tag{9}$$

The proof of the result on the Jacobian is as follows. We set, just for simplicity, $a = \beta'_j, b = \beta'_{j+1}, c = M'_j, d = M'_{j+1}$ and $\Psi = (T^* - T_j)/(T_{j+1} - T_j)$.

It can be shown that the equations (6) and (8) imply

$$u = a/(a+b), \qquad w = c/(c+d), \qquad \beta_j = a^\Psi b^{1-\Psi}.$$

Equation (7) can be written as

$$M_j = cf_1(a, T_j, T^*, T_{j+1}) + df_2(b, T_j, T^*, T_{j+1}). \tag{10}$$

The inverse of the Jacobian is given by considering the derivatives of u, w, β_j andM_j with respect to a, b, c and d, i.e.

$$\begin{vmatrix} b/(a+b)^2 & -a/(a+b)^2 & 0 & 0 \\ 0 & 0 & d/(c+d)^2 & -c/(c+d)^2 \\ \Psi(b/a)^{1-\Psi} & (1-\Psi)(a/b)^\Psi & 0 & 0 \\ \frac{\partial M_j}{\partial a} & \frac{\partial M_j}{\partial b} & \frac{\partial M_j}{\partial c} & \frac{\partial M_j}{\partial d} \end{vmatrix},$$

which becomes, after some computations,

$$\frac{\beta_j}{(\beta'_j + \beta'_{j+1})^2 (M'_j + M'_{j+1})^2} \left[c\frac{\partial M_j}{\partial c} + d\frac{\partial M_j}{\partial d} \right],$$

Table V. Number of change points for simulated data set 1

k	0	1	2	3	4
Prob.	0	0	0.71	0.24	0.05

which is the inverse of (9) since

$$M_j = c\frac{\partial M_j}{\partial c} + d\frac{\partial M_j}{\partial d},$$

because of (10).

The "death" move [D] is given by randomly choosing a change point T_{j+1} and moving from the triple (T_j, T_{j+1}, T_{j+2}) to the pair $(T_j^{'}, T_{j+1}^{'})$ (and the corresponding parameters), so that the following conditions hold:

$$(T_{j+1} - T_j)\log\beta_j + (T_{j+2} - T_{j+1})\log\beta_{j+1} = (T_{j+1}^{'} - T_j^{'})\log\beta_j^{'},$$

$$M_j[G(T_{j+1}, \beta_j) - G(T_j, \beta_j)] + M_{j+1}[G(T_{j+2}, \beta_{j+1}) - G(T_{j+1}, \beta_{j+1})]$$

$$= M_j^{'}[G(T_{j+1}^{'}, \beta_j^{'}) - G(T_j^{'}, \beta_j^{'})].$$

The acceptance probability for the "death" move is similar, with the due changes.

6. Detection of the Number of Change Points

We now consider simulated and well-known actual data and apply them the reversible jump MCMC method proposed in Section 6. We just focus on the issue of detecting the number of change points, actually the posterior distribution on their number.

EXAMPLE 6.1. We consider the Simulated Data Set 1. As shown in Table V, it is worth mentioning that the method rules out the possibility of no change point, favoring a model with two change points. This is expected (and confirm those in Sections 3 and 5) since data were drawn from a NHPP with one change points and a second one, albeit temporary, appeared after the data simulation. It is interesting that the model with three change points gets a 0.24 probability. Therefore, the method is quite satisfactory in detecting the number of change points and, actually, gives a new insight to be further explored with the search of actual change points.

EXAMPLE 6.2. The data in Table VI are from Rigdon and Basu (1989) and they are the failure times of an aircraft engine.

Table VI. Failure times in hours for aircraft engine data

i = Failure number	t_i = Failure time
1	55
2	166
3	205
4	341
5	488
6	567
7	731
8	1308
9	2050
10	2453
11	3115
12	4017
13	4596

Table VII. Number of change points for aircraft engine data

k	0	1	2	3
Prob.	0.61	0.30	0.09	0

We remind that the β's and M's parameters have Gamma priors, i.e. $\mathcal{G}(\alpha, \delta)$ and $\mathcal{G}(\epsilon, \phi)$, respectively. In this example we used $\alpha = \delta = \epsilon = \phi = 0.3$, whereas we considered a truncated (up to $K = 10$) Poisson distribution, with parameter $\omega = 4$, as the prior for the number k of components.

The posterior probability distribution of k is given in Table VII. When $k = 1$, the posterior distribution of the two β_1 and β_2 (the β parameters of the PLP in the two intervals) have means 0.46 and 0.34, respectively, and standard deviations 0.15 and 0.14, respectively. This seems, along with the distribution of k, to indicate that the model used in this approach tends to favor no change points.

EXAMPLE 6.3. A well-known data set for change point analysis is given by the dates of serious coal-mining disasters, between 1851 and 1962, studied (e.g., in Raftery and Akman (1986).

As in Example 6.2, we used $\alpha = \delta = \epsilon = \phi = 0.3$ as values of the Gamma priors on β's and M's, and $\omega = 3$ as the parameter of the truncated (up to $K = 10$) Poisson prior on k. Visual inspection of the draws from the

Table VIII. Number of change points for coal-mining disasters data

k	0	1	2	3
Prob.	0.01	0.85	0.14	0.09

RJMCMC in the from of trace plots, for the location of change points and the PLP parameters, indicated satisfactory convergence.

The posterior probability distribution of k is given in Table VIII. It is worth mentioning that Raftery and Akman (1986) have found out one change point in the data. Our model confirms the existence of one change point and allows for a possible second change point. The posterior median of the the change point (conditional on a single change point) is in March 1892, and the 95% equal tail credible interval is April 1886 to June 1896. This is quite in agreement with the results reported in Raftery and Akman (1986). One could also provide the probability of change point in any given time interval, e.g., between two consecutive failures, conditional on the number of change points or unconditionally, by taking the appropriate relative frequencies from the RJMCMC output.

For large data sets such as this, the model where change points are included as needed and which does not require the introduction of many latent variables, would be preferable to the model considered in Sections 3 and 4.

7. Discussion

We have fitted various models for change points to reflect different scenarios that may cause change points in reliability to occur. These methods may be used in general NHPP, including the PLP which was the main focus in this paper. Which models is appropriate to use (between the models in Sections 4 and 5) in a specific case depends partly on the underlying phenomenon that is believed to cause the change point. For instance, if a change point may only be associated with a failure as may be the case in software reliability, the model in Section 4, would be more appropriate. On the other hand, if it is not clear which model may be (more) appropriate, we would recommend the more general model of Section 5. While carrying out a model selection using Bayes factors, or any other informal approach may be of interest, we feel that it is beyond the scope of this paper, and hope to address this elsewhere. We, however, note that these models themselves may be viewed as inherently mixed models, and hence the inference derived can be thought of as resulting from some suitable model averaging.

References

Ascher, H. E. and Feingold, H.: *Repairable Systems Reliability*, Marcel Dekker, New York, 1984.

Bain, L. J. and Engelhardt, M.: *Statistical Analysis of Reliability and Life-testing Models*, Marcel Dekker, New York, 1991.

Bar-Lev, S., Lavi, I. and Reiser, B.: Bayesian inference for the power law process, *Ann. Institute Math. Stat.* **44** (1992) 623–639.

Campodonico, S. and Singpurwalla, N.: Inference and predictions from poisson point processes incorporating expert knowledge, *J. Am. Stat. Assoc.* **90** (1995), 220–226.

Cox, D. R. and Lewis, P. A.: *Stat. Anal. Events*, Methuen, London, 1966.

Crow, L. H.: Reliability analysis for complex repairable systems. In: F. Proschan and D. J. Serfling (eds), *Reliability and Biometry* SIAM, Philadelphia, 1974.

Crow, L. H.: Confidence interval procedures for the Weibull process with applications to reliability growth. *Technometrics*, **24** (1982), 67–72.

George, E. I. and McCulloch, R. E.: Variable selection via Gibbs sampling *J. Am. Stat. Assoc*, **88** (1993), 881–890.

Green, P.: Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82** (1995), 711–732.

Musa, J. D. and Okumoto, K.: A logarithmic Poisson execution time model for software reliability measurement. *Proceedings Seventh International Conference on Software Engineering*, Orlando, Florida (1984), pp. 230–238.

Pievatolo, A. and Ruggeri, F.: Bayesian reliability analysis of complex repairable systems. *Appl. Stochastic Models Business and Industry*, **20** (2004), 253–264.

Pievatolo, A., Ruggeri, F. and Argiento, R.: Bayesian analysis and prediction of failures in underground trains. *Qual. Reliability Engineering Inter.* **19** (2003), 327–336.

Raftery, A. E. and Akman, V. E.: Bayesian analysis of a Poisson process with a change point. *Biometrika* **73** (1986), 85–89.

Rigdon, S. E. and Basu, A. P.: The Power Law process: a model for the reliability of repairable systems. *J. Qual. Technol.* **10** (1989), 251–260.

Rigdon, S. E. and Basu, A. P.: *Statistical Methods for the Reliability of Repairable Systems*, Wiley, New York, 2000.

Thompson, W. A.: *Point Process Models with Applications to Safety and Reliability*, Chapman and Hall, London, 1988.