*Proceedings of the 2009 Winter Simulation Conference*
*M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds.*

# BAYESIAN NON-PARAMETRIC SIMULATION OF HAZARD FUNCTIONS

Dmitriy Belyi
Elmira Popova
David Morton

Operations Research and Industrial Engineering
The University of Texas at Austin
Austin, TX 78712, U.S.A.

Paul Damien

Red McCombs School of Business
The University of Texas at Austin
Austin, TX 78712, U.S.A.

## ABSTRACT

In Bayesian non-parametric statistics, the extended gamma process can be used to model the class of monotonic hazard functions. However, numerical evaluations of the posterior process are very difficult to compute for realistic sample sizes. To overcome this, we use Monte Carlo methods introduced by Laud, Smith, and Damien (1996) to simulate from the posterior process. We show how these methods can be used to approximate the increasing failure rate of an item, given observed failures and censored times. We then use the results to compute the optimal maintenance schedule under a specified maintenance policy.

## 1 INTRODUCTION

The classical failure rate model typically assumes a parametric distribution of failure times, such as the Weibull distribution. Given a set of observations, researchers using these models estimate the values of the parameters of the assumed distribution (for example, using maximum likelihood estimators). Generally, in Bayesian parametric approaches, researchers assume a particular failure distribution with random parameters, assume prior distributions on these parameters, collect observations, and use Bayes' theorem to compute the posterior distributions (Bernardo and Smith 1994).

One disadvantage of the Bayesian parametric approach can be the assumption of the distribution governing failure times. If the true underlying distribution of failure times is significantly different from the assumed distribution, then the results of analysis may be misleading. In general, it may not be reasonable to assume that an item has a failure distribution that is well-known. For example, Popova et al. (2006) analyze fitting a Weibull distribution to component failure times and show that for many components, the Weibull distribution is not a good fit by the Kolmogorov-Smirnov goodness-of-fit test.

Bayesian non-parametric models attempt to counter this by dropping the assumption on a specific distribution. In the case of estimating the failure rate of an item, a general non-parametric model considers the failure rate to be a random function, takes a continuum of some class of distributions (possibly all distributions) as the state space, and uses Bayesian updates to select a distribution to serve as a model for the failure rate. Dykstra and Laud (1981) discuss using the extended gamma (EG) process as a prior on either monotonically increasing or monotonically decreasing functions. In the case of an increasing failure rate, this means taking the continuum of monotonically increasing functions as the state space. However, it is difficult to analytically compute the posterior process unless the sample size is small, because the posterior process is a complicated expression of multi-dimensional variables.

Laud, Smith, and Damien (1996) propose a Monte Carlo method to simulate from the posterior of the EG process. It is a Gibbs sampler (see Roberts and Polson 1994) that provides a Bayesian solution to non-parametric modeling and analysis of the failure rate using the EG process. This paper will illustrate using this algorithm to effectively model and analyze the increasing failure rate of an item, and show how the results can be applied.

## 2 EXTENDED GAMMA PROCESS

The density of the Gamma distribution, given parameters $\alpha > 0$ and $\beta > 0$ is

$$Gamma(x|\alpha,\beta) = x^{\alpha-1}\frac{e^{-\beta x}}{\Gamma(\alpha)\beta^{\alpha}}. \tag{1}$$

Now, let $\alpha(t)$, $t \geq 0$, be a non-decreasing, left-continuous function such that $\alpha(0) = 0$ (which implies non-negativity). Let $\beta(t)$, $t \geq 0$, be a positive, right-continuous function with left-hand limits existing and bounded away from 0 and $\infty$ (see Laud, Smith, and Damien 1996). Define $Z(t)$, $t \geq 0$, to be a gamma process with parameter $\alpha(\cdot)$. In other words,

- $Z(0) = 0$,
- $Z(t)$ has independent increments,
- for $t > s$, $Z(t) - Z(s) \sim Gamma(\alpha(t) - \alpha(s), 1)$ (increments may be non-stationary).

Then, the EG process is the process $h(t)$, $t \geq 0$, where

$$h(t) = \int_{[0,t]} [\beta(s)]^{-1} dZ(s). \tag{2}$$

Dykstra and Laud (1981) use the distribution of this process as a non-parametric prior for increasing functions; specifically, they use it as a prior for an increasing failure rate $r(t)$.

### 2.1 Prior and Posterior Processes

Let $0 \leq s_0 < s_1 < ... < s_M$ denote a finite partition of the time axis, dividing the time axis into $M$ intervals of length $s_0, s_1, ..., s_M$. Let $\delta_i$ be the increment of the hazard rate over the $i$th interval $(s_{i-1}, s_i]$, for $i = 1, 2, ..., M$; i.e. $h(s_j) = \sum_{i=1}^{i=j} \delta_i$. Dykstra and Laud (1981) show that the prior distribution for each $\delta_i$ is defined by the parameters of the EG process; i.e.

$$\delta_i \sim Gamma(\alpha(i), \beta(i)). \tag{3}$$

To make notation easier, we will refer to this prior density of $\delta_i$, for $i = 1, 2, ..., M$, as $f^*_{\delta_i}$; in other words,

$$\delta_i \sim f^*_{\delta_i}.$$

From reliability theory, the failure rate $r(t)$ corresponds to the cumulative failure distribution function $F(t)$ such that

$$\begin{aligned} F(t) &= 1 - exp\left[-\int_{[0,t]} r(u)du\right] \\ &= 1 - exp\left[-\int_{[0,t]} h(u)du\right] \\ &\approx 1 - exp\left[-\sum_{j\in[0,t]} \delta_j(s_j - s_{j-1})\right]. \end{aligned}$$

Let $d_i$ be the number of exact observations (failures) in the $i$th interval $(s_{i-1}, s_i]$, for $i = 1, 2, ..., M$. Using Bayes' theorem, Dykstra and Laud (1981) show that the posterior of $\Delta = (\delta_1, \delta_2, ..., \delta_M)$ can be expressed as follows:

$$[\Delta|observations] \propto \prod_{i=1}^{M} exp\left[-d_i \sum_{j=1}^{i-1} \delta_j(s_{i-1} - s_{j-1})\right] \left(1 - exp\left[-(s_i - s_{i-1})\sum_{j=1}^{i} \delta_j\right]\right)^{d_i} f^*_{\delta_j}. \tag{4}$$

Some items may still be operating by the time one stops observations (i.e. they did not fail during the observed time horizon). We account for these items by using them as observations recorded at the time when we stop observing the items. These observations, known in reliability theory literature as right-censored observations (see, for example, Rausand and Høyland 2004), provide additional information that is useful in modeling the failure rate. Right-censored data is accounted for by

updating the parameter $\beta(s_i)$ for every $i$. Consider $n$ right-censored items, and let $x_1^c, x_2^c, ..., x_n^c$ be the censored times. Then, $\beta(s_i)$ is updated as follows:

$$\beta(s_i) = \beta(s_i) + \sum_{j \in [1,...,n], x_j^c \geq s_i} (x_j^c - s_i). \tag{5}$$

### 2.2 Gibbs Sampler

The posterior distribution described in (4) is difficult if not impossible to explicitly compute or numerically estimate, for a reasonably sized sample and a moderately dense grid. Laud, Smith, and Damien (1996) tackle this issue by proposing a Markov Chain Monte Carlo algorithm that samples from the posterior distribution. They re-write the posterior distribution (4) as:

$$[\Delta | observations] \propto \prod_{j=1}^{M} \left(1 - e^{-T_j(\Delta)}\right)^{d_j} \prod_{j=1}^{M} f^*_{\delta_j} e^{-a_j \delta_j}, \tag{6}$$

where

- $a_j = \sum_{i=j+1}^{M} d_i (s_{i-1} - s_{j-1})$, and
- $T_j(\Delta) = (s_j - s_{j-1})$.

Then, they introduce random variables

- $\mathbf{m_j} = (m_{j1}, m_{j2}, ..., m_{jj})$ - independent multinomials, where each $\mathbf{m_j}$ is a $j$-cell multinomial of $d_j$ independent trials, with probability of $k$th cell being $p_k = \frac{\delta_k}{\sum_{i=1}^{j} \delta_i}$, and
- $\mathbf{g_j} = (g_{j1}, g_{j2}, ..., g_{jd_j})$, a collection of independent exponential random variables, with mean $\frac{1}{T_j(\Delta)}$, and truncated at 1.

The posterior distribution (6) is then expressed as:

$$[\Delta | observations, \mathbf{g}, \mathbf{m}] \propto \prod_{j=1}^{M} \delta_j^{\sum_{i=j}^{M} m_{ij}} exp\left[ -\left( a_j + \sum_{i=j}^{M} (s_i - s_{i-1}) \sum_{k=1}^{d_j} g_{ik} \right) \delta_j \right] f^*_{\delta_j}. \tag{7}$$

Laud, Smith, and Damien (1996) then show that we can sample from the posterior of $\Delta$ using the following Gibbs sampling algorithm:

- $[\mathbf{g} | observations, \Delta, \mathbf{m}] \propto$ truncated exponentials,
- $[\mathbf{m} | observations, \Delta, \mathbf{g}] \propto$ multinomials,
- $[\Delta | observations, \mathbf{m}, \mathbf{g}] \propto$ Gamma with a rejection algorithm using $f^*_{\delta_j}$ as the importance sampling function.

For more detailed derivations, see Laud, Smith, and Damien (1996), Dykstra and Laud (1981).

## 3 SIMULATION

We implemented this algorithm on a Xeon(TM) 3.00 GHz CPU with 2 Gb of RAM, using Microsoft Visual C++ .Net (Microsoft Development Environment 2003). We used this algorithm to simulate from hazard rates of real data from physical components, as well as simulated data generated from a pre-defined failure distribution for a more in-depth analysis. This algorithm runs fairly efficiently, although this may require a careful set-up of the rejection algorithm when simulating from $\Delta$ (see Laud, Smith, and Damien 1996), and depends on the size of $\Delta$. Table 1 shows how long the algorithm took to generate certain quantities of $\Delta$. The algorithm also seems to converge fairly quickly. Figure 1 shows an example of convergence for the mean of the simulated variates of a particular $\delta$; typically, the mean converges after a couple of hundred iterations. Nevertheless, we use a much larger "burn-in" period of 1000 or more. Figure 2 shows the distribution of that particular $\delta$. Not surprisingly, this distribution (as well as the distribution for other $\delta$) look like a Gamma. We explored approximating the Gamma/rejection algorithm simulation step with a simple Gamma, using the same parameter values. The algorithm

Table 1: Times required to generate quantities of $\Delta$

| Quantity of $\Delta$ | Size of $\Delta$ | Time (Seconds) |
|---|---|---|
| 1000 | 11 | 168 |
| 1000 | 22 | 562 |
| 5000 | 11 | 780 |
| 5000 | 22 | 1561 |
| 10000 | 11 | 1859 |
| 10000 | 22 | 3259 |

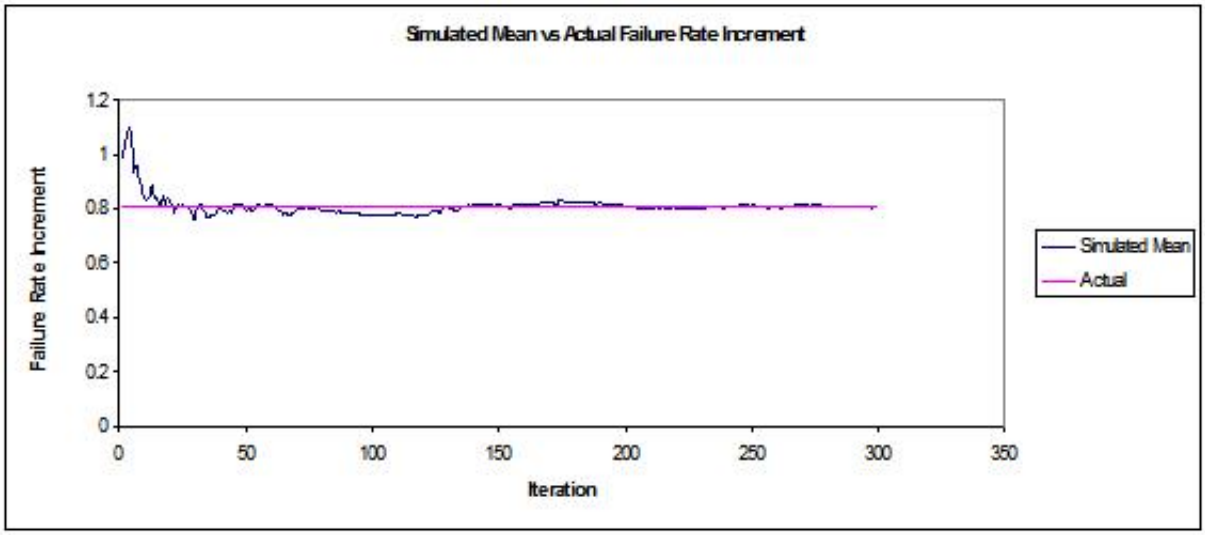


Figure 1: Convergence of the simulated mean to the actual $\delta$

does work faster without the rejection step; unfortunately, the resultant model doesn't match the true failure rate as well as the model with the rejection step. We suspect that there may be a way to do this successfully, perhaps with some slight parameter manipulations. Figure 3 shows the simulated and actual hazard rates for a simulated data set. We see that the Gamma approximation algorithm (at least in the mean) does not match the actual failure rate as well as the Gamma/rejection algorithm. One doesn't have to use the means of the components of $\Delta$ to approximate the desired failure rate. Figure 4 shows the results of using the medians on a real data set. We do not know the actual failure rate of the data set, so we use the empirical failure rate approximation as our benchmark. Also, one of the advantages of simulation is the ability to look at certain percentiles of the failure rate, rather than only at the empirical failure rate. For real data, it is often the case that we do not have many observations, and the empirical failure rate can differ from the actual failure rate. Simulation allows us to consider various percentiles of the failure rate increments, and lets us use more conservative failure rate estimates in our calculations.

### 3.1 Maintenance Optimization

We use the simulation of the failure rate to optimize maintenance scheduling for a single item. Consider the problem of scheduling preventive maintenance (PM) for an item with a strictly increasing, stationary, and continuous failure rate between times 0 and $L$, where $L$ is the time horizon (i.e., we don't care what happens to the item after time $L$). We assume that the item is in a "new" state at time 0, and is again restored to a "new" state by PM. Should an item fail, it is restored to its condition just before the failure (i.e., "good-as-old" state) by corrective maintenance (CM). During a failure, an item may trigger a larger system failure or "trip" with probability $P_{trip}$. Should this happen, instead of a simple CM we must

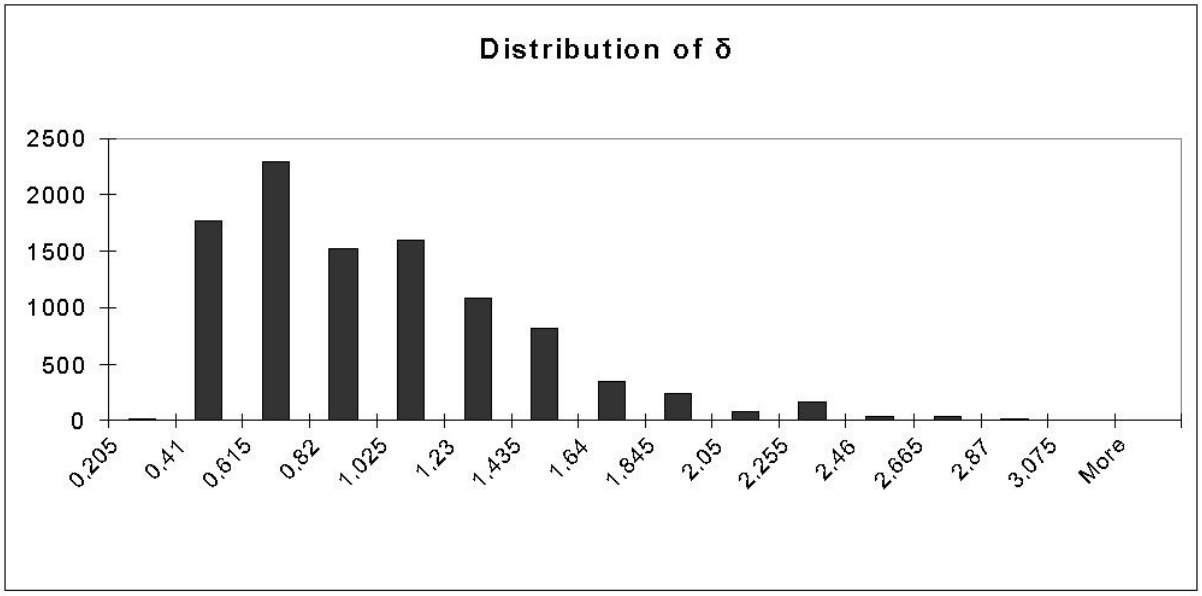


Figure 2: Distribution of a single $\delta$

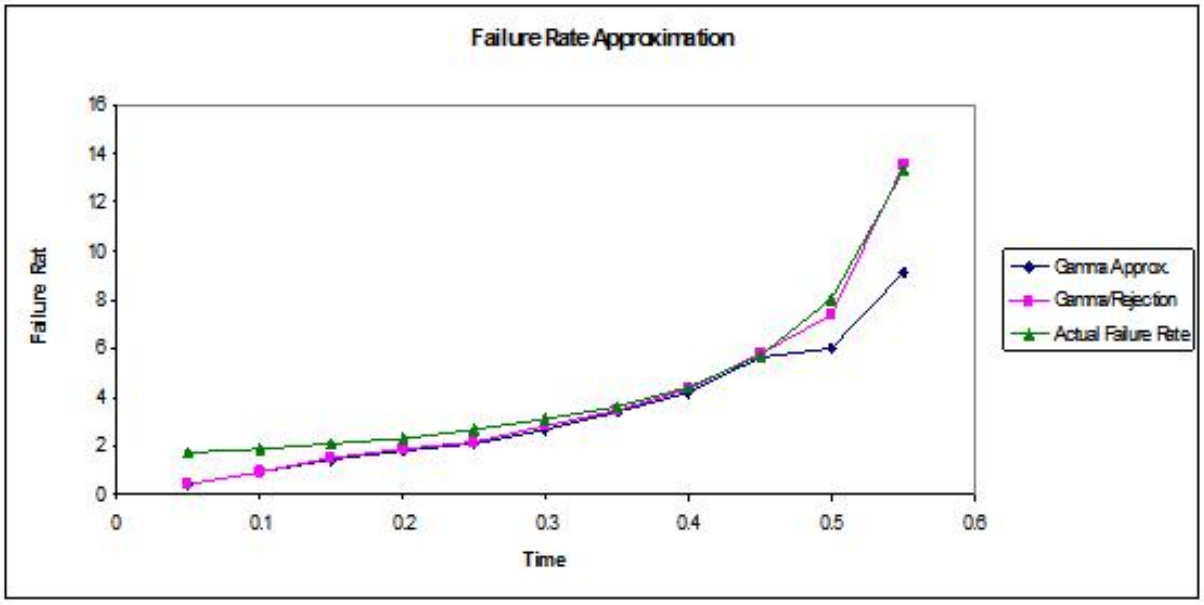


Figure 3: Comparison of failure rate estimates using means

address the entire system "trip" and incur the cost of production loss, which is more expensive than CM. We also assume that PM, CM, and restoration of a downed system are performed instantaneously. Let $\overline{C} = [P_{trip}C_{trip} + (1 - P_{trip})C_m]$, and let $E[N(t_1,t_2)]$ denote the expected number of failures observed between times $t_1$ and $t_2$. If we wish to minimize the expected cost of a maintenance schedule for a general number of PM, we can formulate the problem as:

$$\min_{n\geq 1}(n-1)C_{pm} + \min_{T_1,\ldots,T_n,\sum_i T_i = L} \sum_{i=1}^{n} \bar{C}E[N(0,T_i)].$$

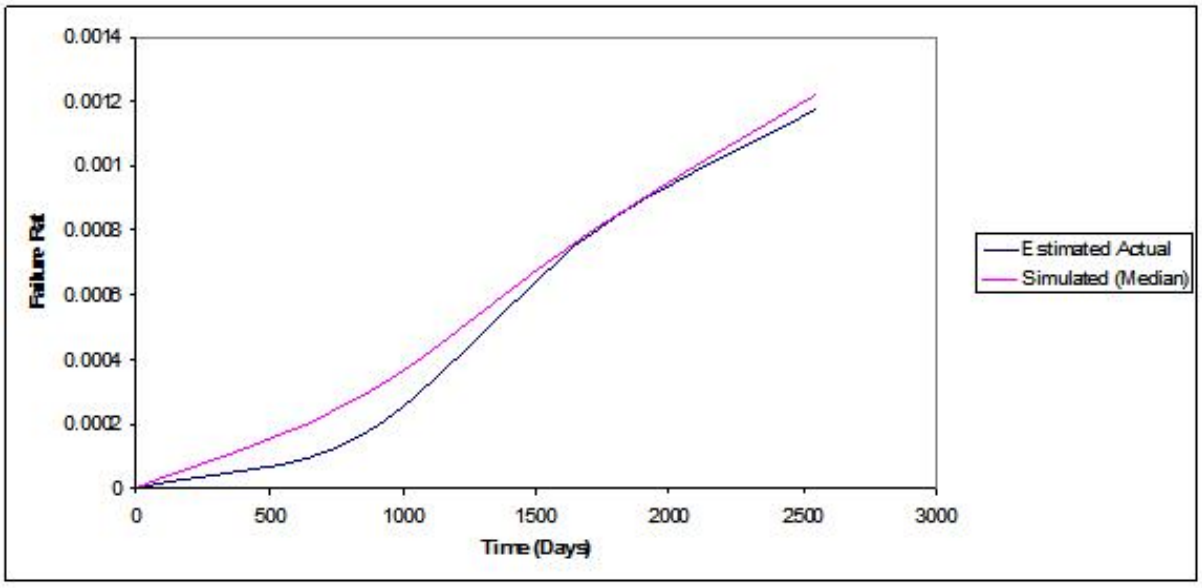


Figure 4: Simulated vs empirical failure rate using medians

Table 2: Estimated parameters of the supply pump

| Parameter | Value ($) |
|---|---|
| $C_{pm}$ | 18998 |
| $C_{cm}$ | 97997 |
| $C_{trip}$ | 4349781 |
| $P_{trip}$ | 0.0073 |

Barlow and Hunter (1960) show that under these conditions, $E[N(0,T_i)] = \int_0^{T_i} z(u)du$. If we have a sufficiently simple analytical form for the failure rate, we can explicitly evaluate this integral. However, if the failure rate does not have a good analytical form, or is unknown, we solve this problem using numerical integration. Given failure data for a specific item, we use the simulation to approximate the failure rate, and use the approximation in the numerical integration. This lets us efficiently solve the minimization problem and derive the optimal schedule.

As an example, we show the use of this procedure on a real data set from the South Texas Nuclear Operating Company (STPNOC) located in Bay City, Texas, USA. We analyze a hydraulic supply pump, a component of the Electro-Hydraulic Control System (EHC) which is a part of the nuclear power generation reactor. Using 223 exact observations with a time horizon of about 880 days, we model the failure rate using the means of the simulated $\delta$s. We can also use the estimated failure rate to simulate from other important functions, such as the survival function (see Rausand and Høyland 2004 for key relationships between these functions). Figure 5 shows how the simulation model compares to the empirical estimated survival function. Using parameters in Table 2, we arrive at the optimal schedule, where we will perform a single PM at approximately 440 days (halfway through the time horizon). The expected cost of the optimal schedule is approximately $84,290; the expected cost of not performing any PM during the time horizon is approximately $102,050. For more details on the algorithm, see Belyi et al. (2009).

## 4 CONCLUSION

We use the Gibbs sampling algorithm proposed in Laud, Smith, and Damien (1996) to simulate from an increasing failure rate. This algorithm considers failure observations as well as censored times, and seems to perform well and efficiently. The

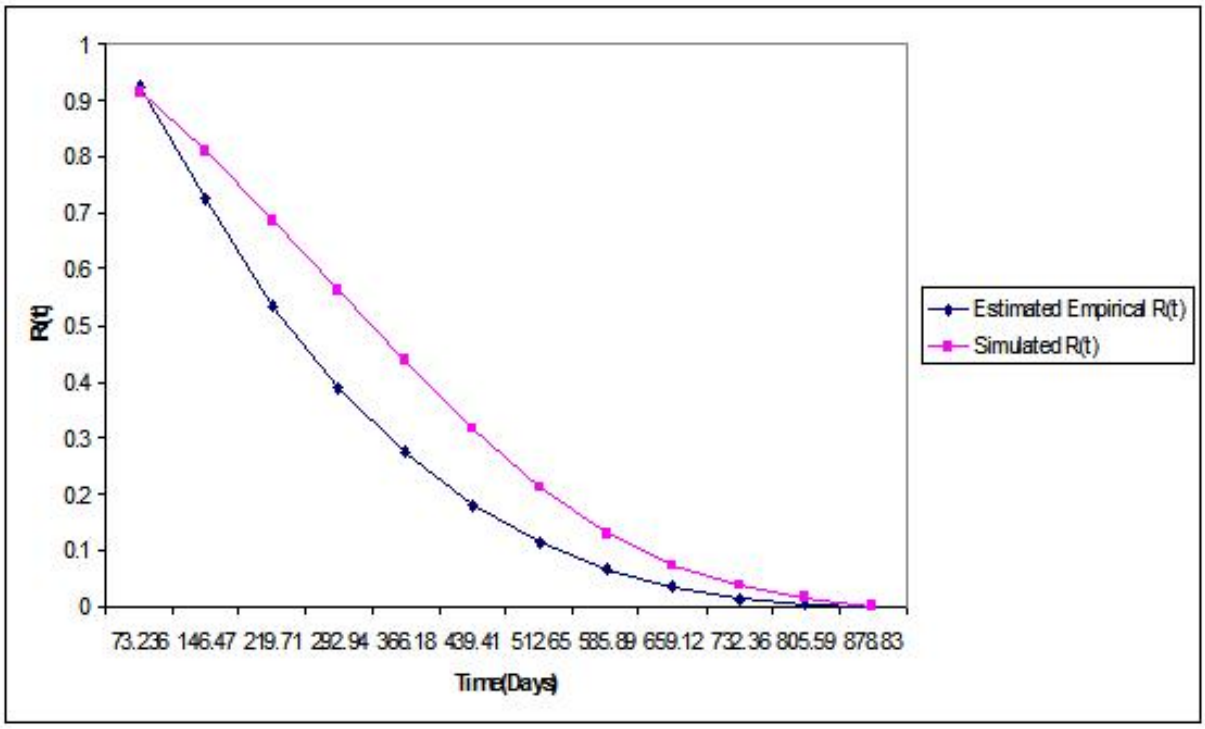


Figure 5: Simulated vs estimated empirical survival function of the supply pump, using means of simulated $\delta$s

simulation also allows for a probabilistic approach to modeling the failure rate by allowing us to consider percentiles of the simulated variates. This algorithm allows us to compute the estimated number of failures between preventive maintenances in a maintenance schedule optimization problem. This technique enables us to solve this problem efficiently. In future research, we will explore a similar algorithm for simulation from non-monotone functions. This will allow us to sample from failure rates that aren't strictly increasing, such as "bathtub" failure rates. This will allow us to solve maintenance optimization problems for items with such failure rates efficiently.

## ACKNOWLEDGMENTS

This research has been partially supported by South Texas Project Nuclear Operating Company under grant B02857, and the National Science Foundation under grants CMMI-0457558 and CMMI-0653916.

## REFERENCES

Barlow, R., and L. Hunter. 1960. Optimum preventive maintenance policies. *Operations Research* 8:90–100.

Belyi, D., E. Popova, D. Morton, P. Damien, E. Kee, and D. Richards. 2009. Bayesian nonparametric analysis of single item preventive maintenance strategies. In *17th International Conference on Nuclear Engineering*. Brussels, Belgium.

Bernardo, J., and A. Smith. 1994. *Bayesian theory*. John Wiley and Sons.

Dykstra, R., and P. Laud. 1981. Bayesian nonparametric approach toward reliablity. *Annals of Statistics* 9:356–367.

Laud, P., A. Smith, and P. Damien. 1996. Monte Carlo methods for approximating a posterior hazard rate process. *Statistics and Computing* 6:77–83.

Popova, E., W. Yu, E. Kee, A. Sun, D. Richards, and R. Grantom. 2006. Basic factors to forecast maintenance cost and failure processes for nuclear power plants. *Nuclear Engineering and Design* 236:1641–1647.

Rausand, M., and A. Høyland. 2004. *System reliability theory: Models, statistical methods, and applications*. John Wiley & Sons, Inc.

Roberts, G., and N. Polson. 1994. On the geometric convergence of the Gibbs sampler. *Journal of the Royal Statistical Society. Series B (Methodological)* 56:377–384.

## AUTHOR BIOGRAPHIES

**DMITRIY BELYI** is a Ph.D. student in Operations Research at The University of Texas at Austin expecting to graduate in May 2010. His research interests are in the area of stochastic optimization, applied statistics, and simulation. His email address is <dmitriy.belyi@gmail.com>.

**PAUL DAMIEN** holds a PhD in Mathematics from Imperial College, London. His interests include Bayesian methods, and their application to engineering, finance, and economics. He has made numerous contributions to Bayesian computation, Bayesian nonparametric modeling, and applied modeling. A Fellow of the Royal Statistical Society of England, he is currently the B.M. Rankin Jr. Professor of Business at the McCombs School of Business at UT-Austin. His e-mail address is <paul.damien@mccombs.utexas.edu>.

**ELMIRA POPOVA** is an Associate Professor in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. Her research interests are in uncertainty modeling of complex systems, advanced Bayesian analysis, Markov Chain Monte Carlo methods, and stochastic optimization. Her e-mail address is <elmira@mail.utexas.edu>.

**DAVID P. MORTON** is Engineering Foundation Professor in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. His research interests include computational stochastic programming, including simulation-based approximations in stochastic programming. His email address is <morton@mail.utexas.edu>, and his web page is <http://www.me.utexas.edu/ orie/Morton.html>.