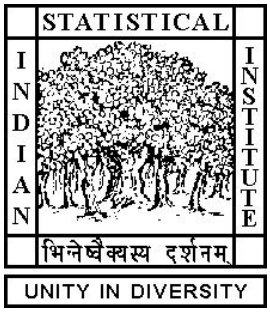
Estimation of Reliability with Semi-parametric Frailty Modeling of Degradation

Technical Report No. ASU/2016/1
Dated: 8 February, 2016

Prajamitra Bhuyan
Applied Statistics Unit,
Indian Statistical Institute,
Kolkata 700108
praja_r@isical.ac.in

Debasis Sengupta
Applied Statistics Unit,
Indian Statistical Institute,
Kolkata 700108
sdebasis@isical.ac.in



Estimation of Reliability with Semi-parametric Frailty Modeling of Degradation

Prajamitra Bhuyan and Debasis Sengupta
Applied Statistics Unit, Indian Statistical Institute, Kolkata 700108.

Abstract

In many real life scenarios, stress accumulates over time and the system fails as soon as the accumulated stress or degradation equals or exceeds a critical threshold. For some devices, it is possible to obtain measurements of degradation over time, and these measurements may contain useful information about product reliability. In this paper, we propose a semi-parametric random effect (frailty) model for degradation path, and a method of estimating this path as well as the reliability. Consistency of the estimator under general conditions is established. Simulation results show superiority of the performance of the proposed method over a parametric competitor. The method is illustrated through the analysis of a real data set.

Keywords: Accelerated failure time, Crack propagation, Kernel function, Monotonic spline, Random effects, SEMOR, Shape invariant model.

1 Introduction

For systems that are designed to achieve high reliability, it is difficult to estimate system reliability with data consisting of a small number of failures, obtained from traditional life tests that record only time to failure. For many such systems, reliability depends on the dynamic balance between stress and strength, where stress accumulates over time. As for example, a vehicle axle fails when the depth of a crack has exceeded a critical level (Nakagawa, 2007, p. 2). Measurements on accumulated stress or degradation, taken over time, contain information about reliability of the unit. By harnessing this information, one can hope to achieve better specification of reliability even with data from a relatively small number of units (Lu and Meeker, 1993).

In designed experiments, systems are inspected at prefixed time points and current status of the systems along with measurements of accumulated stress are recorded. For such systems, failure can be defined in terms of a specified level of strength $s(t)$ at time t, and the reliability at time t is given by $R(t) = P[X(t) < s(t)]$, where $X(t)$ denotes the accumulated stress at time t (See Bhuyan and Dewanji, 2015). Sobczyk and Trbicki (1989) emphasized the importance of statistically designed experiments for estimation of the model parameters for analyzing fatigue crack data, modelled as a cumulative random process. Bhuyan and Dewanji (2017) considered the problem of estimating reliability of a system under deterministic strength as a function of time and cumulative damage due to shocks arriving according to a point process.

There have been much work on the modeling of degradation leading to failure (Gorijan et al., 2009; Nikulin et al., 2010). It is common in many of these models to choose a fixed threshold $s(t) = s$ for the degradation $X(t)$ for all $t \geq 0$. For these models, the reliability at time t is given by

$$R(t) = P\left[X(t) < s\right]. \tag{1}$$

Lu and Meeker (1993) considered a for the degradation as a function of time, given by

$$X(t) = \mu(t, \theta, \phi), \quad t \geq 0, \tag{2}$$

where ϕ is a vector of fixed effect parameters, $\boldsymbol{\theta}$ is a vector of random effect parameters and the degradation is measured with additive error at specified times. They considered a data set consisting of sample paths measuring fatigue crack lengths at equispaced time points for a number of metallic

specimens under test (See Figure 1). Failure is defined by the event of critical crack length exceeding a constant level of 1.6 inches (that is, $s(t) = 1.6$ inches). Lawless and Crowder (2004) regarded $X(t)$ as observable without error and modeled it as a Gamma process, with the scale parameter explained by a random effect. Their analysis of the above data set revealed lack of fit, which they attributed to a simplifying assumption that they were forced to make for the sake of computational tractability. Park and Padgett (2005) considered the same data and modeled $X(t)$ alternatively as a Gamma process and also as a geometric Brownian motion, while ignoring the heterogeneity present among the degradation paths. In order to estimate the reliability function $R(t)$, they assumed a specific parametric model for the degradation path $\mu(\cdot)$ and developed a two stage estimation methodology for the associated parameters. In general, any parametric methodology is potentially sensitive to the assumed model and the resulting estimator may be biased whenever the choice of the parametric form is incorrect.

We consider a special case of (2) motivated by the above data set, and propose a semi-parametric estimation procedure which does not require prior knowledge of the functional form of the degradation path $\mu(\cdot)$ and the distribution of the random effect parameter θ. In Section 2, we describe a simplified model and propose a method of estimating the reliability function $R(\cdot)$ given by (1). Small sample properties of the proposed estimator are investigated through a simulation study in Section 3. Illustrative analysis of the data set used by Lu and Meeker (1993) is presented in Section 4. We provide some concluding remarks in Section 5. Consistency of the proposed estimator of the reliability function is given in the appendix.

2 Model and Estimation Methodology

In order to develop a simple model for degradation that is not constrained by a specified parametric form, let us consider the degradation paths plotted in Figure 1. It is interesting to observe that all the sample paths seem to have the same basic shape and they hardly cross one another. It is only the rate of degradation that varies across the units and it appears that all the paths might coincide if they were properly scaled. Motivated by this example, we assume that a strictly increasing function $\eta(\cdot)$ represents the 'baseline' degradation path. We take in to account the heterogeneity present among the individual sample paths through the following model for degradation

$$\mu(t, \theta, \phi) = \eta(\theta t), \quad t \geq 0, \tag{3}$$

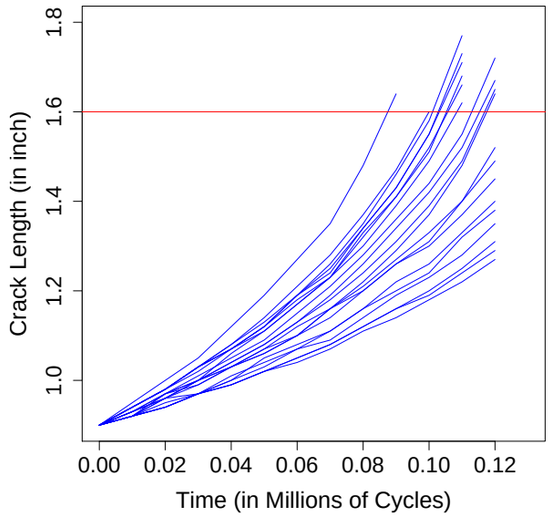


Figure 1: Crack growth propagation path

where θ represents a scalar random effect parameter (commonly referred to as *frailty* in the literature of survival analysis), which accelerates or decelerates the time scale. The random effect parameter θ may have different values for different individuals. These different values of θ are assumed to be random realizations from a distribution $G(\cdot)$, say, whose density $g(\cdot)$ is continuous with compact support $[C, D]$, $C > 0$. Note that the above model is a special case of (2) with $\phi = 1$ and θ is a scalar. The corresponding model for the degradation measurements of K individual units is

$$Y_{ij} = \eta(\theta_i t_{ij}) + \epsilon_{ij}, j = 1, 2, \ldots, n_i, i = 1, 2, \ldots, K, \tag{4}$$

where Y_{ij} is the observed accumulated stress at time t_{ij} contained in the interval $[L, U]$, , and the ϵ_{ij}'s are independently and identically distributed (iid) measurement errors with mean 0 and variance σ^2, the θ_i's are iid with distribution $G(\cdot)$ and the θ_i's and ϵ_{ij}'s are independent. The above model is a special case of shape invariant model (SIM), introduced by Lawton et al. (1972), which is also a special case of the measurement model used by Lu and Meeker (1993):

$$Y_{ij} = \mu(t_{ij}, \theta_i, \phi) + \epsilon_{ij}, j = 1, 2, \ldots, n_i, i = 1, 2, \ldots, K. \tag{5}$$

One can re-write (4), as

$$\begin{aligned} Y_{ij} &= \eta(\theta_i t_{ij}) + \epsilon_{ij} \\ &= \eta_0(\theta_{0i} t_{ij}) + \epsilon_{ij}, j = 1, 2, \ldots, n_i, i = 1, 2, \ldots, K, \end{aligned}$$

where $\eta_0(t) = \eta(t/\alpha)$, and $\theta_{01}, \ldots, \theta_{0K} \overset{iid}{\sim} G_0(\cdot)$, with $G_0(x) = G(x/\alpha)$, α being a fixed and positive number. The pair of functions $\eta(\cdot)$ and $G(\cdot)$ are not identifiable from the given data, since any choice of α gives rise to the same set of measurements. In order to make the model identifiable, without loss of generality, we assume $E(\theta_1) = 1$ (See Kneip and Gasser (1988), for more details). By putting together the model (3) with the general expression (1), we obtain the reliability at time t as

$$R(t) = P\left[\eta(\theta t) < s\right]. \tag{6}$$

In order to estimate the reliability function (6) from the measurements (4), we estimate the baseline degradation path $\eta(\cdot)$ and predict the random scale parameters $\theta_1, \theta_2, \ldots \theta_K$, subject to the constraint $E(\theta_1) = 1$, using the Self-Modeling Nonlinear Regression (SEMOR) algorithm by Kneip and Gasser (1988). Then we estimate the distribution $G(\cdot)$ using the predicted values of $\theta_1, \ldots, \theta_k$ as 'data'. These two estimated functions are then combined to estimate $R(t)$ through (6). For implementation of the SEMOR algorithm, one needs to choose a suitable approximation space for the estimation of $\eta(\cdot)$. We use the space of spline functions, which is a popular choice for estimation in the SEMOR environment (Kneip and Gasser, 1988).

2.1 Monotone Spline

Since the function $\eta(\cdot)$ is assumed to be increasing, we have to estimate it under this constraint. In the special case of (4), with the parameter θ_i being equal to 1 for all individuals, the problem reduces to that of constrained regression with a single explanatory variable. For this problem, Ramsay (1988) used a space of monotonically increasing splines as an approximation space for $\eta(\cdot)$. He defined the 'monotone regression spline' function $\eta(t) = a_0 + \sum_{l=1}^{r} a_l I_l(t)$, where $I_1(\cdot), \ldots, I_r(\cdot)$ are monotonic polynomial basis functions, referred to by Ramsay as Integrated splines or I-splines, and r is the number of basis functions, which is, the dimension of the approximation space. The coefficients $a_1, \ldots, a_r$ should be non-negative for assuring monotonicity. We consider second order

I-spline basis functions (Ramsay, 1988) given by

$$I_l(t) = \begin{cases} 0 & , t < k_l \\ \frac{(t-k_l)^2}{(k_{l+2}-k_l)(k_{l+1}-k_l)} & , k_l \le t < k_{l+1} \\ 1 - \frac{(k_{l+2}-t)^2}{(k_{l+2}-k_l)(k_{l+2}-k_{l+1})} & , k_{l+1} \le t < k_{l+2} \\ 1 & , t \ge k_{l+2}, \end{cases}$$

for $l = 1, \ldots, r$, and an increasing sequence of knot points $k_1 = k_2 < \cdots < k_{r+1} = k_{r+2}$. Ramsay (1988) discussed determination of the number of interior knots and their placement in detail. However, Meyer (2008) noted that the regression splines with shape restriction (such as monotonicity) are often found to be robust to the number of knot points and their placement.

Ramsay (1988) proposed to estimate the coefficients $a_0, \ldots, a_r$ through the method of least squares. Since the data are assumed to arise from the measurement model (4), with $\theta_i = 1$ for all i, this amounts to solving the optimization problem

$$\begin{aligned} \underset{a_0,\ldots,a_r}{\text{minimize}} \quad & \sum_{i=1}^{K} \sum_{j=1}^{n_i} [Y_{ij} - a_0 - \sum_{l=1}^{r} a_l I_l(t_{ij})]^2 \\ \text{subject to} \quad & a_j \ge 0, \ j = 1, \ldots, r. \end{aligned}$$

Ramsay (1988) and Meyer (2008) provided algorithms for finding the constrained least squares solution for the I-splines.

If $\eta(\cdot)$ is a strictly increasing function, then it is desirable that the estimated regression function is strictly increasing. One can easily find such a solution by replacing the constraints by $a_l \ge c_l$, for all $l = 1, \ldots r$, where c_l's are known positive constants. If some prior knowledge about the lower bound of the slope of $\eta(\cdot)$ is available, one can choose suitable values for the c_ls.

2.2 SEMOR Algorithm

In order to estimate $\eta(\cdot)$ and predict $\theta_1, \theta_2 \ldots, \theta_K$ simultaneously, we employ the SEMOR algorithm suggested by Kneip and Gasser (1988). Let us denote $S = \{(a_0, a_1, \ldots, a_r) \colon a_0 \ge 0, a_l \ge \frac{\gamma}{I_l'(k_{l+1})}, l = 1, \ldots, r\}$, where γ is a positive number smaller than $\inf_t \{\eta'(t) : t \in [CL, DU]\}$, which has already been assumed to be positive. Note that the lower bounds of a_l, for $l = 1, \ldots, r$, are chosen in this way such that the slope of the resulting estimate of $\eta(\cdot)$ is greater than or equal to γ.

The steps of the iterative SEMOR algorithm for the estimation of $\eta(\cdot)$ and θ_i, for $i = 1, \ldots, K$, are provided below.

Step I: Choose initial values of $\hat{\theta}_i^{(0)}$s such that $\frac{1}{K}\sum_{i=1}^{K}\hat{\theta}_i^{(0)} = 1$ and consider the data $(\hat{\theta}_i^{(0)}t_{ij}, Y_{ij})$, $j = 1, \ldots n_i$, and $i = 1, \ldots K$. Now fit a monotone spline of Section 2.1 by solving

$$(\hat{a}_0^{(0)}, \ldots, \hat{a}_r^{(0)}) = \underset{(a_0,\ldots,a_r)\in S}{\operatorname{argmin}} \quad c(a_0, \ldots, a_r, \hat{\theta}_1^{(0)}, \ldots, \hat{\theta}_K^{(0)})$$

where

$$c(a_0, \ldots, a_r, \theta_1, \ldots, \theta_K) = \frac{1}{K}\sum_{i=1}^{K}\frac{1}{n_i}\sum_{j=1}^{n_i}[Y_{ij} - a_0 - \sum_{l=1}^{r} a_l I_l(\theta_i t_{ij})]^2,$$

with the I_l's being the second order I-spline basis functions as given in Section 2.1. Then, get the initial estimate of $\eta(t)$ as $\hat{\eta}^{(0)}(t) = \hat{a}_0^{(0)} + \sum_{l=1}^{r}\hat{a}_l^{(0)}I_l(t)$.

Step II: Determine the $(u+1)$st iterate $\hat{\theta}_i^{(u+1)}$, for $i = 1, \ldots, K$, by solving

$$(\hat{\theta}_1^{(u+1)}, \ldots, \hat{\theta}_K^{(u+1)}) = \underset{\theta_1,\ldots,\theta_K\in[C,D]}{\operatorname{argmin}} \; c(\hat{a}_0^{(u)}, \ldots, \hat{a}_r^{(u)}, \theta_1, \ldots, \theta_K)$$
$$\text{subject to} \quad \frac{1}{K}\sum_{i=1}^{K}\theta_i = 1,$$

where $\hat{a}_l^{(u)}$ is the estimate of a_l from the uth iteration, for $l = 1, \ldots, r$, $u = 0, 1, 2, \ldots$. Then consider the data $(\hat{\theta}_i^{(u+1)}t_{ij}, Y_{ij})$, $j = 1, \ldots n_i$, $i = 1, \ldots K$ and solve

$$(\hat{a}_0^{(u+1)}, \ldots, \hat{a}_r^{(u+1)}) = \underset{(a_0,\ldots,a_r)\in S}{\operatorname{argmin}} \quad c(a_0, \ldots, a_r, \hat{\theta}_1^{(u+1)}, \ldots, \hat{\theta}_K^{(u+1)}),$$

to get the $(u+1)$st iterate $\hat{\eta}^{(u+1)}(t) = \hat{a}_0^{(u+1)} + \sum_{l=1}^{r}\hat{a}_l^{(u+1)}I_l(t)$.

Step III: We stop the iteration process if, for some pre-specified small $\nu > 0$,

$$\frac{|c(\hat{a}_0^{(u)}, \ldots, \hat{a}_r^{(u)}, \hat{\theta}_1^{(u)}, \ldots, \hat{\theta}_K^{(u)}) - c(\hat{a}_0^{(u+1)}, \ldots, \hat{a}_r^{(u+1)}, \hat{\theta}_1^{(u+1)}, \ldots, \hat{\theta}_K^{(u+1)})|}{c(\hat{a}_0^{(u)}, \ldots, \hat{a}_r^{(u)}, \hat{\theta}_1^{(u)}, \ldots, \hat{\theta}_K^{(u)})} < \nu.$$

2.3 Reliability Estimate

We re-write the reliability, function given by equation (6), by making use of $G(\cdot)$ and $\eta(\cdot)$, as

$$R(t) = \begin{cases} 1, & s > \eta(Dt), \\ G\left(\dfrac{\eta^{-1}(s)}{t}\right), & \eta(Ct) \leq s \leq \eta(Dt), \\ 0, & s < \eta(Ct). \end{cases} \tag{7}$$

Note that, since $\eta(\cdot)$ is assumed to be strictly increasing, $\eta^{-1}(\cdot)$ exists. In order to estimate the reliability function, we need to estimate $\eta^{-1}(s)$ and $G(\cdot)$. One can easily obtain $\hat{\eta}^{-1}(s)$ from $\hat{\eta}(\cdot)$ by the method of bisection (Epperson, 2013, p. 90-95). However, for this method to work, one need the condition $\hat{\eta}(CL) \leq s \leq \hat{\eta}(DU)$, which is not guaranteed. In order to obtain a computable solution, we define a modified estimator $\tilde{\eta}(\cdot)$ of $\eta(\cdot)$, as

$$\tilde{\eta}(t) = \begin{cases} \hat{\eta}(t), & t \in (CL, DU), \\ \min\{s, \hat{\eta}(t)\}, & t = CL, \\ \max\{s, \hat{\eta}(t)\}, & t = DU. \end{cases}$$

Since $a_l = \frac{\gamma}{I_l'(k_{l+1})}$, for all $l = 1, \ldots r$, $\hat{\eta}(t)$ is a strictly increasing function (See Section 2.1) and hence $\tilde{\eta}(t)$ is also a strictly increasing function. Then, we estimate $G(x)$ by $\int_0^x \frac{1}{Kh} \sum_{i=1}^{K} W\left(\frac{v-\hat{\theta}_i}{h}\right) dv$, where $W(\cdot)$ is a suitably chosen kernel function with h being the corresponding bandwidth (Silverman, 1986, p. 15). We plug in the estimator of $G(\cdot)$ and $\eta^{-1}(\cdot)$ in the expression (7) and get

$$\hat{R}(t) \quad = \quad \int_0^{\tilde{\eta}^{-1}(s)/t} \frac{1}{Kh} \sum_{i=1}^{K} W\left(\frac{x-\hat{\theta}_i}{h}\right) dx. \tag{8}$$

The integral in equation (8) is calculated by using the R function *integrate*, which utilizes the Fortran functions DQAGE and DQAGIE, available from Netlib, for implementation of the uni-dimensional adaptive 15-point Gauss-Kronrod quadrature formula. The details of the algorithm and its precision are discussed in Piessens et al. (1983).

One can estimate the bias and obtain the standard error of the estimator given by (8) through a model based bootstrap method (see Efron (1982) and Efron and Tibshirani (1986) for details).

The confidence interval of $R(t)$ can also be obtained by using the following steps.

Step I: Calculate the residuals as $e_{ij} = Y_{ij} - \tilde{\eta}(\hat{\theta}_i t_{ij})$, for $i = 1, \ldots, K$, $j = 1, \ldots, n_i$.

Step II: Simulate iid θ_i^*, $i = 1, \ldots, K$, from $\hat{G}(x) = \int_0^x \frac{1}{Kh} \sum_{i=1}^K W(\frac{v-\hat{\theta}_i}{h})dv$ (by adding independent samples from the density $\frac{1}{h}W(\cdot/h)$ and the empirical distribution of the $\hat{\theta}_i$s) and K sample paths $Y_{ij}^* = \tilde{\eta}(\theta_i^* t_{ij}) + e_{ij}^*$, $i = 1, \ldots K$, $j = 1, \ldots n_i$, where e_{ij}^*'s are simulated from the empirical distribution of the observed residuals e_{ij}'s, and t_{ij}'s are the same inspection times used in the original experiment.

Step III: Based on the data (t_{ij}, Y_{ij}^*), for $i = 1, \ldots, K$, $j = 1, \ldots, n_i$, compute $\tilde{\eta^*}(s)$ and $\hat{\theta}_1^*, \ldots, \hat{\theta}_K^*$ using the SEMOR algorithm, as described in Section 2.2, and then obtain the estimate $\hat{R}^*(t) = \int_0^{\tilde{\eta^*}^{-1}(s)/t} \frac{1}{Kh} \sum_{i=1}^K W\left(\frac{x-\hat{\theta}_i^*}{h}\right) dx$, for any given $t > 0$.

Step IV: Replicate Step II and Step III for N_B number of times, where N_B is a large number, say 1000.

Step V: Estimate the bias and obtain the standard error of $\hat{R}(t)$ from the N_B bootstrap estimates of the reliability $\hat{R}^*(t)$'s. Also, the confidence interval of $R(t)$ is obtained from the histogram of the N_B number of bootstrap reliability estimates, for some $t > 0$.

3 Simulation Study

For the purpose of investigating the small sample performance of the proposed estimator, we generate samples from the model (4) with the baseline degradation path as $\eta(t) = \exp(t)$. The corresponding random effect scale parameters $\theta_1, \theta_2, \ldots, \theta_K$ are drawn from the location shifted Beta distribution with the parameter values 1.2 and 1.5, and the shift parameter 0.5, so that $\theta_1 - 0.5 \sim Beta(1.2, 1.5)$. Note that the θ_i's take values in the interval $(0.5, 1.5)$. The iid random errors are generated from the normal distribution with mean 0 and standard deviation 0.1. We assume that failure occurs as soon as the amount of degradation equals or exceeds $s = 5$ units. We consider K sample paths and measurements of degradation are taken at n_I equispaced time points over the range $[0.2, 1.95]$, with measurements terminating prematurely in case the degradation reaches the threshold of failure before the last measurement time. The choice of the threshold s and the observation window is such that, nearly half of the units do not fail at the end of the test, which is in line with the data in the real example discussed in the next section. The number of interior knots $r-2$, needed for the implementation of the SEMOR algorithm, is taken as $(K \times n_I)/50$

and these are spaced uniformly over the available range of the data at any stage of iteration. For convergence, the threshold used on the relative error of the criterion was $\nu = 10^{-15}$. We choose $W(\cdot)$ as the Gaussian kernel and obtain the corresponding bandwidth by using Silverman's rule of thumb (Silverman, 1986, p. 48). In order to estimate the bias, the variance (VAR) and the mean square error (MSE) of the SEMOR estimate of the degradation path and the proposed estimator of reliability, we repeat the simulation procedure 1000 times, and summarize the results in Table 1 and 2, respectively, for three different combinations of K and n_I. As expected, the mean square error of the estimator of baseline degradation path decreases as the number of units under test increases. A similar pattern is observed as we increase the number of inspection points. The bias, the variance and the mean square error of the estimator of reliability decreases as the number of units under test increases, and also with increase in number of inspection points. The improvement of the mean square error is more pronounced when K is increased.

Table 1: Small sample performance of the SEMOR estimator of the baseline degradation path

		$K=10, n_I=10$			$K=20, n_I=10$			$K=20, n_I=20$		
Parameter	True	Bias×100	VAR×100	MSE×100	Bias×100	VAR×100	MSE×100	Bias×100	VAR×100	MSE×100
$\eta(0.30)$	1.35	-0.12	0.10	0.10	0.10	0.06	0.06	-0.10	0.06	0.06
$\eta(0.45)$	1.57	-0.01	0.07	0.07	-0.01	0.06	0.06	-0.01	0.04	0.04
$\eta(0.60)$	1.82	0.12	0.09	0.09	-0.05	0.06	0.06	-0.03	0.04	0.04
$\eta(0.75)$	2.12	-0.22	0.06	0.06	0.08	0.07	0.07	0.05	0.04	0.04
$\eta(0.90)$	2.46	-0.24	0.08	0.08	-0.11	0.06	0.06	0.04	0.06	0.06
$\eta(1.05)$	2.86	0.22	0.11	0.11	-0.06	0.09	0.09	0.08	0.07	0.07
$\eta(1.20)$	3.32	0.29	0.12	0.12	-0.05	0.10	0.10	0.16	0.10	0.10
$\eta(1.35)$	3.86	-0.82	0.16	0.17	-0.06	0.17	0.17	0.07	0.14	0.14
$\eta(1.50)$	4.48	-0.93	0.25	0.26	-0.06	0.23	0.23	0.37	0.19	0.20
$\eta(1.65)$	5.21	0.15	0.37	0.37	-0.48	0.36	0.36	0.35	0.31	0.31

Table 2: Small sample performance of the reliability estimator

		$K=10, n_I=10$			$K=20, n_I=10$			$K=20, n_I=20$		
Parameter	True	Bias×100	VAR×100	MSE×100	Bias×100	VAR×100	MSE×100	Bias×100	VAR×100	MSE×100
$R(1.120)$	0.98	-3.27	0.19	0.30	-2.58	0.09	0.16	-2.49	0.08	0.14
$R(1.164)$	0.95	-2.80	0.32	0.40	-2.15	0.17	0.21	-2.03	0.15	0.19
$R(1.226)$	0.90	-2.04	0.53	0.57	-1.55	0.30	0.32	-1.38	0.26	0.28
$R(1.284)$	0.85	-1.46	0.75	0.77	-1.14	0.43	0.44	-0.93	0.37	0.38
$R(1.342)$	0.80	-1.05	0.97	0.98	-0.87	0.55	0.56	-0.63	0.47	0.48
$R(1.400)$	0.75	-0.76	1.19	1.19	-0.67	0.65	0.66	-0.42	0.57	0.57
$R(1.460)$	0.70	-0.59	1.40	1.38	-0.53	0.74	0.74	-0.27	0.65	0.65
$R(1.588)$	0.60	-0.57	1.71	1.71	-0.33	0.82	0.83	-0.13	0.77	0.77
$R(1.732)$	0.50	-0.70	1.84	1.84	-0.23	0.82	0.83	-0.10	0.80	0.80
$R(2.097)$	0.30	-0.29	1.26	1.26	0.21	0.62	0.62	0.18	0.61	0.61

In order to compare the performance of the semi-parametric estimator with the parametric estimator proposed by Lu and Meeker (1993), we generate degradation measures at 10 equispaced time points for $K = 20$ under the following choices of models given by (4) and (5).

(i) $\eta(\theta t) = \exp(\theta t)$, $t \in [0.2, 1.95]$, $\theta - 0.5 \sim Beta(1.2, 1.5)$,

(ii) $\eta(\theta t) = \exp\left[1.8\{1 + \exp\left(-\frac{\theta t - 10}{2}\right)\}^{-1}\right] - 0.012$, $t \in [0.1, 1.4]$ $\theta \sim U[8, 12]$,

(iii) $\mu(t, \theta, \phi) = \phi(1 - \phi^{\theta^{(2)}}\theta^{(1)}\theta^{(2)}t)^{-\frac{1}{\theta^{(2)}}}$, $t \in [0.1, 1.2]$, $\theta = (\theta^{(1)}, \theta^{(2)})'$ follows the bivariate normal distribution with mean $(10, -2)'$, and covariance matrix $\left(\begin{smallmatrix} 0.5 & -0.1 \\ -0.1 & 0.05 \end{smallmatrix}\right)$, and ϕ is known to be 1.

In each of the above three simulation models, the iid errors are from $N(0, 0.1)$ and the threshold $s = 5$. In Figure 2, we plot the degradation paths (i)-(iii) against time, keeping the respective random effect parameters fixed at their expected values. (That is, $\eta(E(\theta)t)$ and $\mu(t, E(\theta), \phi)$ vs. time). The degradation paths are chosen such that the shapes differ from one another (convex, S-shaped, and concave for (i)-(iii), receptively). Also, note that the amount of degradation at $t = 0$ is unity for all the three choices (i)-(iii). We repeat the simulation procedure 1000 times and compute

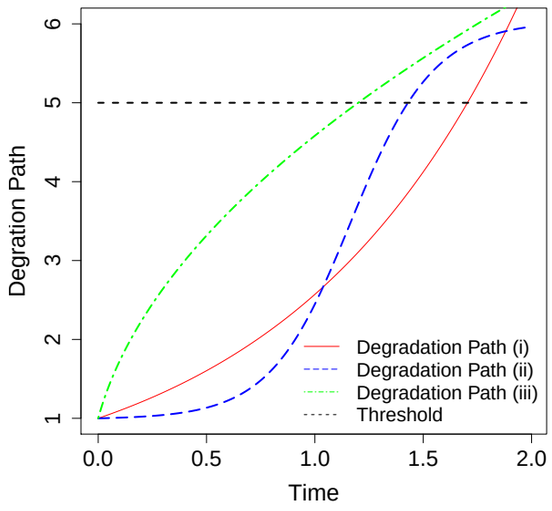


Figure 2: Baseline degradation path

the bias, the variance and the MSE based on the 1000 estimators of reliability. The degradation path considered in (iii) is a special case of the well-known Paris equation, which was used to model

fatigue-crack growth data by Lu and Meeker (1993). For the implementation of the parametric estimator proposed by Lu and Meeker (1993), one needs to assume a suitable parametric form of the degradation path. For this purpose, we use the three degradation paths mentioned in (i)-(iii) and obtain the corresponding reliability estimates. The results of the simulation study under the model (i) are summarized in Figure 3. In this figure, we show the bias, the variance and the MSE of the semi-parametric estimator as well as the parametric estimator with presumed degradation paths (i)-(iii). Since the data are generated from the model (i), it is not surprising to see that the parametric estimator with presumed degradation path (i) has better performance than that with (ii) with respect to MSE. It is interesting to observe that the parametric estimator with presumed degradation path (iii) has comparable performance with that for presumed degradation path (i) with regard to all three of bias, variance and MSE. This is due to the fact that the function $\mu(t, \theta, 1)$ in the Paris equation, with $\theta^{(1)} > 0, \theta^{(2)} > 0$, is convex in nature and it can well approximate the exponential growth curve within the observation window under consideration. The semi-parametric estimator, that does not assume any particular form of the degradation path, has comparable MSE with the appropriate parametric estimator. Similar results of the simulation study under the model (ii) are summarized in Figure 4. As expected, the parametric estimator with the presumed degradation path (ii) has better performance compared to other parametric estimators. However, the semi-parametric estimator has comparable MSE with the appropriate parametric estimator. We have similar results for the simulation study under the model (iii) (See Figure 5). We have also obtained similar results from simulation studies under different other choices of parametric models, the results of which are not reported here for the sake of brevity.

In order to study the effects of misspecification of the random effects structure, we have generated data from the following models, which are generalizations of (i), (ii) and (iii), respectively.

(I) $\eta(\theta t) = \lambda + \exp(\theta t)$, $t \in [0.2, 1.95]$, $\theta - 0.5 \sim Beta(1.2, 1.5)$, $\lambda \sim U[0.5, 1.5]$,

(II) $\eta(\theta t) = \lambda + \exp\left[1.8\{1 + \exp\left(-\frac{\theta t - 10}{2}\right)\}^{-1}\right] - 0.012$, $t \in [0.1, 1.4]$ $\theta \sim U(8, 12)$, $\lambda \sim U[0.5, 1.5]$,

(III) $\mu(t, \theta, \phi) = \lambda + \phi(1 - \phi^{\theta^{(2)}}\theta^{(1)}\theta^{(2)}t)^{-\frac{1}{\theta^{(2)}}}$, $t \in [0.1, 1.2]$, $\theta = (\theta^{(1)}, \theta^{(2)})'$ follows the bivariate normal distribution with mean $(10, -2)'$, and covariance matrix $\left(\begin{smallmatrix} 0.5 & -0.1 \\ -0.1 & 0.05 \end{smallmatrix}\right)$, $\lambda \sim U[0.5, 1.5]$ and ϕ is known to be 1.

Note that the only change appearing in the above models is an additional location parameter, which is random. The mean of this parameter is chosen as 1, which is a fifth of the threshold

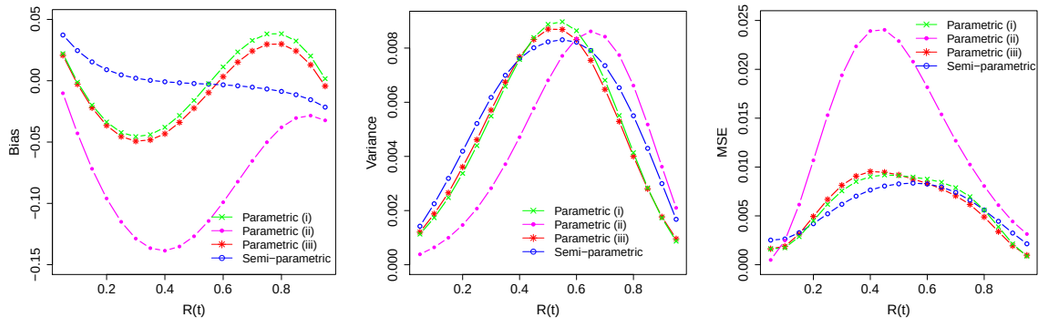


Figure 3: Comparison of bias, variance and MSE of the parametric and semi-parametric estimators under the model (i)

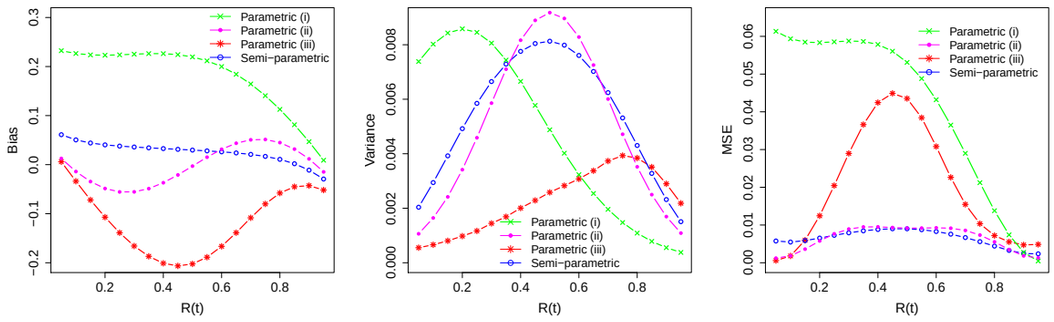


Figure 4: Comparison of bias, variance and MSE of the parametric and semi-parametric estimators under the model (ii)

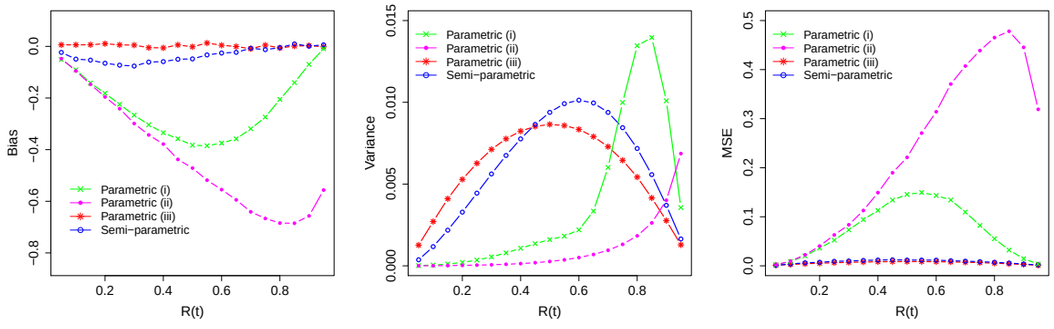


Figure 5: Comparison of bias, variance and MSE of the parametric and semi-parametric estimators under the model (iii)

for failure. We simulate data from (I), (II) and (III), and analyze them with the proposed semi-parametric method along with the parametric method by Lu and Meeker (1993) with presumed degradation paths (i), (ii) and (iii), respectively. In Figure 6, we compare the MSE of the reliability estimates from the aforementioned misspecified models (I)-(III) with the corresponding MSE for the appropriate data generation models (i)-(iii), respectively. The semi-parametric method is found to be less affected by model misspecification than the parametric method meant for the original model without random location change, when the data are generated from models I and II. The two methods are somewhat similarly affected when the data are generated from model III.

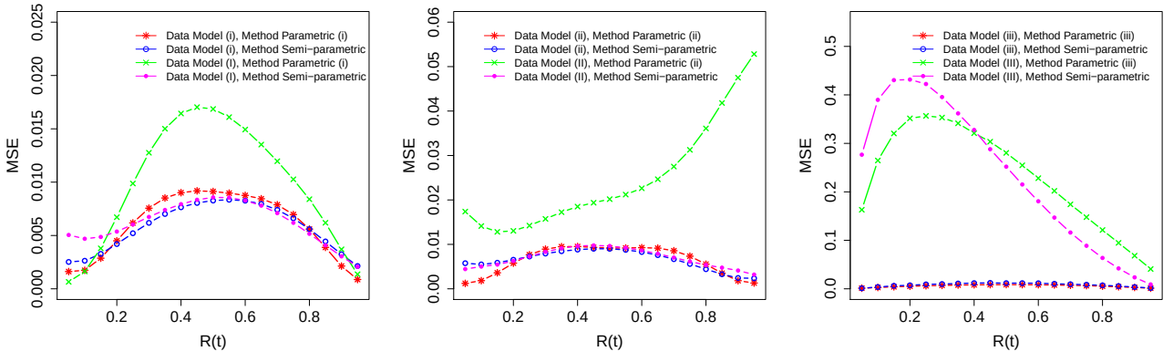


Figure 6: Comparison MSE of the misspecified and correctly specified model

In summary, the simulation results indicate that the proposed semi-parametric estimator performs comparably with the most appropriate parametric estimator under various simulation models, while the parametric estimators may perform poorly when the presumed model is inappropriate.

4 Data Analysis

In this section, we analyze the fatigue-crack-growth data from Hudak et al. (1978). This data set has been provided and analyzed by Lu and Meeker (1993, Table 1). Metal specimens with initial crack length of 0.9 inch were considered and the crack lengths were observed at 13 equispaced time intervals up to 0.12 million cycles or until crack size reached the critical length of 1.6 inches (See Figure 1). Note that, in this experiment, the actual time of failure is recorded even if it is outside the observation window (Lu and Meeker, 1993, Table 2). Such additional data, which would normally not be available in practice, permit computation of the empirical survival function. From a

preliminary analysis of the crack length data on the basis of model (4), we observe that the residuals exhibit larger spread at larger values of time. Therefore, we take the logarithm of crack lengths to fit model (4). We obtain reasonably good fit by estimating the baseline degradation path and predicting the scale parameters through the SEMOR algorithm with two uniformly spaced interior knots. In Figure 7, we plot the estimated baseline degradation path $\tilde{\eta}(\cdot)$ along with the scatter plot of the observed data adjusted with the predicted scale parameters (Y_{ij} vs. $\hat{\theta}_i t_{ij}$). The plot shows a satisfactory fit with no visible anomaly. Next, we treat the predicted scale parameters

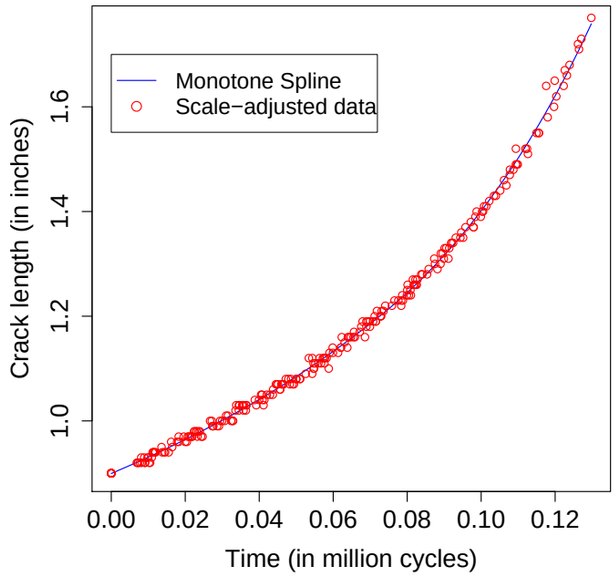


Figure 7: Baseline degradation path

as observed realizations from $G(\cdot)$, and compute the kernel density estimator of the density $g(\cdot)$. For this purpose, we choose the Gaussian kernel and select the corresponding bandwidth by using Silverman's rule of thumb. The comparison of the proposed semi-parametric estimate of the reliability function $R(t)$, given by (8), the parametric estimate obtained by Lu and Meeker (1993), and the empirical survival function from the exact failure time data (not available in practice) are provided in Figure 8. For computation of the parametric estimator, Lu and Meeker (1993) presume that the log-degradation path follows the model $\mu(t, \theta, \phi) = \log(\phi) - \frac{1}{\theta^{(2)}} \log(1 - \phi^{\theta^{(2)}} \theta^{(1)} \theta^{(2)} t)$, with $\phi = 0.9$ and $\theta = (\theta^{(1)}, \theta^{(2)})$ follows bivariate Normal distribution. The parametric estimator deviates

substantially from the empirical survival function near the right tail, while the proposed estimator has more deviation near the left tail. The performance of the parametric and semi-parametric estimators of the reliability function are measured in terms of Integrated absolute error (IAE) (Devroye and Gyrfi, 1985) and Integrated squared error (ISE) (Silverman, 1986, p. 48), defined as $\int_{t_{(1)}}^{t_{(k)}} |\tilde{R}(t) - \hat{R}(t)|dt$, and $\int_{t_{(1)}}^{t_{(k)}} \{\tilde{R}(t) - \hat{R}(t)\}^2 dt$, respectively, where $\tilde{R}(t)$ is the empirical survival function and $t_{(1)}$ and $t_{(k)}$ are the smallest and the largest order statistics, respectively, based on the exact failure time data (Lu and Meeker, 1993, Table 2). The superior performance of the proposed estimator is evident from the result that the IAE×100 (ISE×10000) values for the semi-parametric and parametric estimates are 0.22 (1.14) and 0.35 (2.08), respectively.

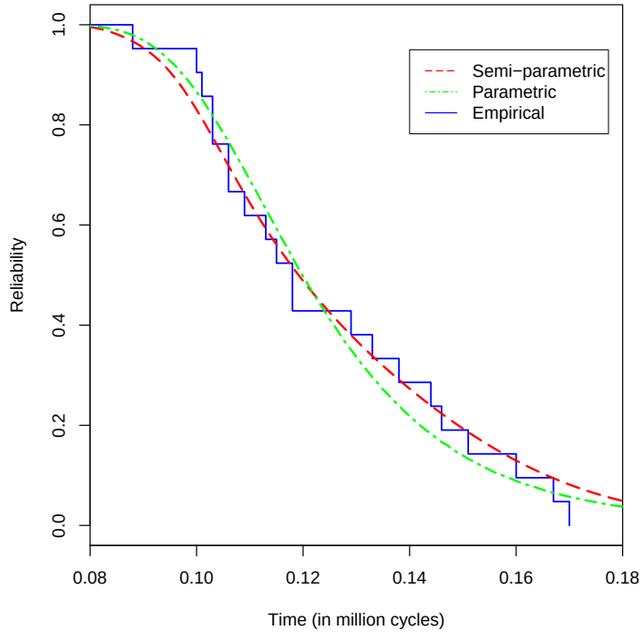


Figure 8: Estimated reliability function

We perform the model based bootstrap method, as explained at the end of Section 2.3, for the purpose of estimating the bias (BBias), standard error (BSE) and point wise confidence interval (BCI) of the reliability with $s = 1.6$ inches, $K = 21$, and $N_B = 1000$. The simulated paths that cross the threshold $s = 1.6$ inches are truncated at the time point after crossing. The results are compared with asymptotic standard error (ASE) and asymptotic confidence interval (ACI) using normality of

the K-M curve in Table 3. One can notice that the sign and magnitude of the bootstrap estimate of bias by and large represents the gap between the semi-parametric estimate and the empirical survival function. In other words, correction of this estimated bias would push the semi-parametric estimator further towards the empirical survival function. Note that the ASE and ACI can not be obtained for $t \geq 0.17$.

Table 3: Results of the fatigue-crack-growth data analysis

Parameter	BBias×100	BSE×100	ASE×100	95%BCI	95%ACI
$R(0.09)$	-2.13	4.02	4.65	(0.836,0.995)	(0.876,1)
$R(0.10)$	-2.38	6.62	6.41	(0.663,0.934)	(0.799,1)
$R(0.11)$	0.01	8.54	10.60	(0.475,0.809)	(0.445,0.793)
$R(0.12)$	0.49	9.19	10.80	(0.307,0.679)	(0.251,0.606)
$R(0.13)$	0.13	8.91	10.60	(0.190,0.561)	(0.207,0.555)
$R(0.14)$	0.14	8.08	9.86	(0.125,0.442)	(0.124,0448)
$R(0.15)$	0.67	7.03	8.57	(0.069,0.349)	(0.050,0.331)
$R(0.16)$	1.41	5.93	6.41	(0.031,0.276)	(0,0.201)
$R(0.17)$	1.99	4.90	-	(0.012,0.214)	-

5 Concluding Remarks

Modeling through degradation data has been a significant approach to evaluation of the reliability and safety of some critical systems. Degradation data can be modeled in many ways depending upon the nature of the failure mechanism. Ample work on parametric degradation modeling can be found in the literature (Gorijan et al., 2009; Nikulin et al., 2010). However, in most real life scenarios, the nature of the degradation path is unknown, and a pre-determined parametric model for it may be inadequate. In this paper, we have made an attempt to estimate the degradation path without assuming any parametric form for it, while accounting for individual heterogeneity through a multiplicative random parameter. This model is motivated by a real life data set.

However, the scope of the model goes far beyond this particular data set. A case in point is the data on degradation of transconductance in semiconductors, used by Lu et al. (1997) along with the model

$$Y_{ij} \approx \Theta_{0i} + \Theta_{1i} log(t_{ij}).$$

Figure 1 of Lu et al. (1997) show sample degradation paths in log-scale. These paths happen to be parallel to one another, that is, Θ_{1i}'s should be identical. This simplification reduces the above model to a special case of (4), with $\theta_i = \exp(\Theta_{0i}/\Theta_{11})$ and $\eta(t) = \log(t^{\Theta_{11}})$.

The scope of the model (4) may be extended further by letting the Y'_{ij}s to be monotonically transformed versions of the measured degradation, rather than the degradation itself. This modification would allow the use of a general scale of degradation to be decided by the user. For example, one can use the percentiles of degradation, as in Lu et al. (1997). All the theoretical results of this paper readily apply to this modified model.

One can consider estimation of reliability from the generalized version of model (4) considered by Lawton et al. (1972), where three additional random parameters account for a shift in the time variable and shift and scale change in the degradation function. Some of the parameters in this generalized version of model (4), may be fixed effect parameters. The SEMOR algorithm may be used in these cases also, and the expression (8) may be adjusted to the case of a random vector parameter $\boldsymbol{\theta}$. The issue of consistency of the estimator in this case is not straight-forward. This problem may be considered in future.

Since we estimate reliability at a fixed point of time, the estimator (8) is properly defined even when the strength $s(t)$ is a non-constant but known function of time. The same method of estimation and proof of consistency of $\hat{R}(t)$ given in Subsection 2.3 and in the appendix also apply for this general case. The asymptotic distribution of $\hat{R}(t)$ is difficult to obtain. One can verify asymptotic normality of the proposed estimator through simulation experiments. However, this is a challenging task in terms of the associated computational burden.

Note that, in some cases, the degradation path may not necessarily be monotonic. In such cases, one can use ordinary splines instead of monotonic splines and estimate the baseline degradation path $\eta(\cdot)$ using the penalized least square technique (Hardle et al., 2004, p. 102). Since ordinary regression splines are sensitive to the number of knots and their placements, one needs to pay attention to these issues for the implementation of the SEMOR algorithm.

Appendix: Consistency of the Estimator

Let $m = \min\{K, n_1, \ldots, n_K\}$ and r_m be the dimension of the approximation space for fixed m. We write $\hat{R}(t)$ (8) as

$$\hat{R}_m(t) = \int_0^{\hat{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{x - \hat{\theta}_{i,m}}{h_m}\right) dx,$$

while using m as subscript of the relevant quantities in order to indicate their explicit dependence on m. In order to prove that $\hat{R}_m(t)$ is a consistent estimator of $R(t)$ as $m \to \infty$, we first state the convergence properties of the estimators $\hat{\eta}(\cdot)$ and $\hat{\theta}_i$, for $i = 1, \ldots, K_m$ obtained by using the SEMOR algorithm. The following conditions would be needed for the results to hold.

(C1) The error terms ϵ_{ij}, $j = 1, \ldots, n_i$, $i = 1, \ldots, K_m$ are independent and identically distributed random variables with zero mean and finite fourth order moment.

(C2) The inspection times are such that, for each subinterval $J \subset [L, U]$, there exists $n_0 \in \mathbb{N}$ and a $\delta > 0$ so that $\frac{1}{n_i} \sum_{j=1}^{n_i} 1(t_{ij} \in J) > \delta$, whenever $n_i \geq n_0$, for $i = 1, 2, \ldots$.

Theorem A.1 (Kneip and Gasser, 1988)**.** *Suppose $r_m \to \infty$ as $m \to \infty$ and the condition C1-C2 hold. Then, with probability 1, (i) $\lim_{m\to\infty} \hat{\eta}_m(\theta t) = \eta(\theta t)$ uniformly for all $(\theta, t) \in [C, D] \times [L, U]$, (ii) for each $i = 1, 2, \ldots$, $\lim_{m\to\infty} \hat{\theta}_{i,m} = \theta_i$, and (iii) for every $\xi > 0$, $\lim_{m\to\infty} \frac{1}{K_m} \sum_{i=1}^{K_m} \mathbb{1}(|\hat{\theta}_{i,m} - \theta_i| \geq \xi) = 0$.*

Before proceeding further, we prove a corollary of part (*iii*) of the above theorem, which will be needed in the eventual proof of consistency.

Corollary A.1. *Suppose $r_m \to \infty$ as $m \to \infty$ and conditions C1-C2 hold. Then, with probability 1, $\lim_{m\to\infty} \frac{1}{K_m} \sum_{i=1}^{K_m} |\hat{\theta}_{i,m} - \theta_i| = 0$.*

Proof. Fix $\xi > 0$ and write $A_m = \{i \in \{1, \ldots, K_m\} \colon |\hat{\theta}_{i,m} - \theta_i| \geq \frac{\xi}{2}\}$. Let $B = \{\omega : \lim_{m\to\infty} \frac{1}{K_m} \sum_{i=1}^{K_m} \mathbb{1}(i \in A_m) = 0\}$. Note that $P(B) = 1$, by Theorem A.1 (*iii*). Now, for any

fixed $\omega_0 \in B$, $|\hat{\theta}_{i,m} - \theta_i| \leq D - C$, for all $i = 1, 2, \ldots K_m$. Then,

$$\begin{aligned}
\frac{1}{K_m}\sum_{i=1}^{K_m}|\hat{\theta}_{i,m} - \theta_i| &= \frac{1}{K_m}\sum_{i\in A_m}|\hat{\theta}_{i,m} - \theta_i| + \frac{1}{K_m}\sum_{i\in A_m^c}|\hat{\theta}_{i,m} - \theta_i| \\
&\leq \frac{D-C}{K_m}\sum_{i=1}^{K_m}\mathbb{1}(i \in A_m) + \frac{1}{K_m}\sum_{i=1}^{K_m}\mathbb{1}(i \in A_m^c)\frac{\xi}{2} \\
&\leq \frac{D-C}{K_m}\sum_{i=1}^{K_m}\mathbb{1}(i \in A_m) + \frac{\xi}{2}. \qquad (9)
\end{aligned}$$

Since $\lim_{m\to\infty}\frac{1}{K_m}\sum_{i=1}^{K_m}\mathbb{1}(i \in A_m) = 0$, there exists $m_0 \in \mathbb{N}$ such that $\lim_{m\to\infty}\frac{D-C}{K_m}\sum_{i=1}^{K_m}\mathbb{1}(i \in A_m) \leq \frac{\xi}{2}$, for all $m \geq m_0$. Now, from (9), we get $\frac{1}{K_m}\sum_{i=1}^{K_m}|\hat{\theta}_{i,m} - \theta_i| \leq \xi$, for all $m \geq m_0$. Hence, the proof. □

Next, we show that $\tilde{\eta}_m^{-1}(s)$, inverse of the estimator defined in Section 2.3, is a consistent estimator of $\eta^{-1}(s)$. For this purpose, we first state the following theorem on the convergence of sequences of inverse functions and then prove the required result.

Theorem A.2 (Barvinek et al., 1991). *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real injections, D_n being the domain of f_n. If the sequence converges uniformly to a function f_0 on $[c, d] \subseteq \bigcap_{n=1}^{\infty} D_n$, where f_0 is a continuous injection on $[c, d]$, and if $[u, v]$ is an interval contained in $\bigcap_{k=0}^{\infty} f_k([c, d])$, then f_n^{-1} converges to f_0^{-1} uniformly on $[u, v]$.*

We now derive the related probabilistic result needed for our purpose.

Theorem A.3. *Suppose $r_m \to \infty$ as $m \to \infty$ and the conditions C1-C2 hold. Then, with probability 1, $\lim_{m\to\infty}\tilde{\eta}_m^{-1}(s) = \eta^{-1}(s)$, for all $s \in \eta((CL, DU))$.*

Proof. Fix $s \in \eta((CL, DU))$. Let $B = \{\omega : \lim_{m\to\infty}\hat{\eta}_m(t) = \eta(t)$ uniformly for all $t \in [CL, DU]\}$. Note that $P(B) = 1$, by Theorem A.1 (i). Since $\eta(\cdot)$ is a strictly increasing and continuous function, there exists a unique $t_0 \in (CL, DU)$ such that $\eta(t_0) = s$. Note that $\eta(t_0 - \frac{t_0-CL}{2}) < s$, and $\eta(t_0 + \frac{DU-t_0}{2}) > s$. Fix $\epsilon_0 \in (0, \min\{s - \eta(t_0 - \frac{t_0-CL}{2}), \eta(t_0 + \frac{DU-t_0}{2}) - s\})$. Now, for any fixed $\omega_0 \in B$, by Theorem A.1 (i), there exists m_0 such that, for all $m \geq m_0$, and for all $t \in [CL, DU]$, $|\hat{\eta}_m(t) - \eta(t)| < \epsilon_0$. Therefore, $\hat{\eta}_m(t_0 - \frac{t_0-CL}{2}) < s$ and $\hat{\eta}_m(t_0 + \frac{DU-t_0}{2}) > s$, for all $m \geq m_0$. Since $\hat{\eta}_m(\cdot)$ is a continuous function, for all $m \geq m_0$, by the Intermediate Value Theorem, there exits

$\tau_m \in (t_0 - \frac{t_0 - CL}{2}, t_0 + \frac{DU - t_0}{2}) \subset [CL, DU]$, such that $\hat{\eta}_m(\tau_m) = s$. Hence, $\tilde{\eta}_m(t) = \hat{\eta}_m(t)$ for all $t \in [CL, DU]$, for all $m \geq m_o$ (See Subsection 2.3).

Now we show that $\lim_{m\to\infty} \tilde{\eta}_m(t) = \eta(t)$ uniformly for all $t \in [CL, DU]$ for the chosen ω_0. Fix $\epsilon_1 > 0$. Since $\omega_0 \in B$, there exists m_1 such that, for all $m \geq m_1$ and for all $t \in [CL, DU]$, we have $|\hat{\eta}_m(t) - \eta(t)| < \epsilon_1$, using Theorem A.1(i). Write $m^* = \max\{m_0, m_1\}$ and observe that, for all $m \geq m^*$ and for all $t \in [CL, DU]$, we have $|\tilde{\eta}_m(t) - \eta(t)| = |\hat{\eta}_m(t) - \eta(t)| < \epsilon_1$. Therefore, for the chosen $\omega_0 \in B$, we have $\lim_{m\to\infty} \tilde{\eta}_m(t) = \eta(t)$ uniformly for all $t \in [CL, DU]$.

Let $B^* = \{\omega : \lim_{m\to\infty} \tilde{\eta}_m(t) = \eta(t) \text{ uniformly for all } t \in [CL, DU]\}$. We have shown that $B \subset B^*$ and hence $P(B^*) = 1$. Now fix $\omega_o^* \in B^*$. Note that $\eta(t)$ is a strictly increasing function by assumption (See Section 2). It has also been noted in Subsection 2.3 that $\tilde{\eta}_m(t)$ is a strictly increasing function for all $m = 1, 2, \ldots$. If $s \in \bigcap_{m=1}^{\infty} \tilde{\eta}_m([CL, DU]) \bigcap \eta([CL, DU])$, then it follows from Theorem A.2, that $\lim_{m\to\infty} \tilde{\eta}_m^{-1}(s) = \eta^{-1}(s)$. Hence the proof. □

In order to prove the consistency of the estimator of the reliability function $R(t)$, we present some relevant asymptotic results on the kernel density estimator under the following condition.

(C3) The kernel function $W(\cdot)$ satisfies the properties: (i) $W(\cdot)$ is uniformly continuous and of bounded variation, (ii) $W(x) \to 0$ as $|x| \to \infty$, (iii) $W(\cdot)$ is integrable and $\int W(x)dx = 1$, (iv) $\int |x \log |x||^{\frac{1}{2}} |dW(x)| < \infty$.

Theorem A.4 (Silverman, 1978)**.** *Suppose $X_1, X_2, \ldots, X_n$ are samples from a distribution with uniformly continuous density $\psi(\cdot)$ and the kernel function $W(\cdot)$ satisfies the condition C3. Suppose $h_n \to 0$ and $(nh_n)^{-1} \log n \to 0$ as $n \to \infty$. Then, with probability 1, $\sup |\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) - \psi(x)| \to 0$ as $n \to \infty$.*

Lemma A.1. *Suppose $X_1, X_2, \ldots, X_n$ are samples from a distribution with uniformly continuous and bounded density $\psi(\cdot)$ and the kernel function $W(\cdot)$ satisfies the condition C3. Then $\{\omega : \lim_{n\to\infty} \sup |\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) - \psi(x)| \to 0\} \subseteq \{\omega : \frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) \leq M$, for some $M > 0, n = 1, 2, \ldots\}$.*

Proof. Fix $\omega_0 \in \{\omega : \lim_{n\to\infty} \sup |\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) - \psi(x)| \to 0\}$ and take $\epsilon = 1$. There exists $n^* \in \mathbb{N}$ such that $\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) \leq \psi(x) + 1$, for all $n \geq n^*$. Since $\psi(\cdot)$ is assumed to be bounded, there exists a constant $M_1 > 0$ such that $\psi(x) \leq M_1$ for all x, and hence $\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) \leq M_1 + 1$, for all $n \geq n^*$. By assumption C3(i), there exists a constant $M_2 > 0$, such that $W(x) \leq M_2$ for all x. Therefore, $\frac{1}{nh_n} \sum_{i=1}^{n} W(\frac{x - X_i}{h_n}) \leq \frac{M_2}{h_n}$, for $n = 1, \ldots, n^* - 1$. Let

$M = \max\{M_1 + 1, \frac{M_2}{h_1}, \ldots, \frac{M_2}{h_{n^*-1}}\}$. Then, $\frac{1}{nh_n}\sum_{i=1}^{n} W(\frac{x-X_i}{h_n}) \leq M$ for all $n = 1, 2, \ldots$. Hence, $\omega_0 \in \{\omega : \frac{1}{nh_n}\sum_{i=1}^{n} W(\frac{x-X_i}{h_n}) \leq M, \text{ for some } M > 0, n = 1, 2, \ldots\}$. □

We would need Corollary A.1, Theorems A.1 and A.4 and Lemma A.1 in order to prove the consistency of the proposed estimator of the reliability function. We need to assume a certain rate associated with the convergence established in Corollary A.1. We state below this assumption and a further condition on the kernel function.

(C4) $\frac{1}{K_m}\sum_{i=1}^{K_m} |\hat{\theta}_{i,m} - \theta_i| = o(\frac{1}{a_m})$ with probability 1, for some positive sequence $\{a_m\}_m$, where $a_m \to \infty$ as $m \to \infty$.

(C5) $W'(x)$ exists for all x and is bounded.

We now present the main result of this section.

Theorem A.5. *Suppose $r_m \to \infty$ as $m \to \infty$, and conditions C1-C5 hold. Then, for all $t > 0$ and for all $s \in \eta((CL, DU))$, there exists a bandwidth sequence $\{h_m\}$ such that $\lim_{m\to\infty} \hat{R}_m(t) = R(t)$ with probability 1.*

Proof. Let $B_1 = \{\omega : \frac{1}{K_m}\sum_{i=1}^{K_m} \left|\hat{\theta}_{i,m} - \theta_i\right| = o(\frac{1}{a_m})\}$. Note that $P(B_1) = 1$ by Corollary A.1 and assumption (C4). For any fixed $\omega_0 \in B_1$, let us consider a sequence $\{h_m^*\}$ such that it converges to 0 and $\lim_{m\to\infty} \frac{\log K_m}{K_m h_m^*} = 0$. Now we define $h_m = \max\{h_m^*, \frac{1}{\sqrt{a_m}}\}$, for all $m = 1, 2, \ldots$, such that $\{h_m\}$ converges to 0. We also have $0 \leq \lim_{m\to\infty} \frac{\log K_m}{K_m h_m} \leq \lim_{m\to\infty} \frac{\log K_m}{K_m h_m^*} = 0$, and hence $\lim_{m\to\infty} \frac{\log K_m}{K_m h_m} = 0$. Similarly, $0 \leq \lim_{m\to\infty} \frac{1}{K_m h_m^2}\sum_{i=1}^{K_m} \left|\hat{\theta}_{i,m} - \theta_i\right| \leq \lim_{m\to\infty} \frac{a_m}{K_m}\sum_{i=1}^{K_m} \left|\hat{\theta}_{i,m} - \theta_i\right| = 0$, and hence $\lim_{m\to\infty} \frac{1}{K_m h_m^2}\sum_{i=1}^{K_m} \left|\hat{\theta}_{i,m} - \theta_i\right| = 0$.

Let $B_1^* = \{\omega : \lim_{m\to\infty} \frac{1}{K_m h_m^2}\sum_{i=1}^{K_m} \left|\hat{\theta}_{i,m} - \theta_i\right| = 0\}$. The above argument shows that $B_1^* \supseteq B_1$ and, hence, $P(B_1^*) = 1$. Also, let $B_2 = \{\omega : \lim_{m\to\infty} \sup |\frac{1}{K_m h_m}\sum_{i=1}^{K_m} W(\frac{v-\theta_i}{h_m}) - g(v)| \to 0\}$ and $B_3 = \{\omega : \lim_{m\to\infty} \tilde{\eta}_m^{-1}(s) = \eta^{-1}(s)\}$, where $g(\cdot)$ is the density corresponding to the common distribution $G(\cdot)$ of the θ_i's. Note that $P(B_2) = 1$ and $P(B_3) = 1$, by Theorem A.4 and Theorem A.3, respectively. It follows that $P(B_1^* \cap B_2 \cap B_3) = 1$.

For a chosen $t > 0$ such that $\eta(Ct) \leq s \leq \eta(Dt)$, we have

$$\begin{aligned}
|\hat{R}_m(t) - R(t)| &= \left| \int_0^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \hat{\theta}_{i,m}}{h_m}\right) dv - G\left(\frac{\eta^{-1}(s)}{t}\right) \right| \\
&\leq \int_0^{\tilde{\eta}_m^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \hat{\theta}_{i,m}}{h_m}\right) - \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \theta_i}{h_m}\right) \right| dv \\
&\quad + \left| \int_{\eta^{-1}(s)/t}^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \theta_i}{h_m}\right) dv \right| \\
&\quad + \int_0^{\eta^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \theta_i}{h_m}\right) - g(v) \right| dv. \qquad (10)
\end{aligned}$$

We would now chose $\omega_1 \in B_1^* \cap B_2 \cap B_3$ and show that each of the three terms on the right hand side of (10) tends to zero.

Consider the first term and note that $\tilde{\eta}_m^{-1}(s) \in [CL, DU]$. Therefore

$$\begin{aligned}
&\int_0^{\tilde{\eta}_m^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \hat{\theta}_{i,m}}{h_m}\right) - \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v - \theta_i}{h_m}\right) \right| dv \\
\leq &\int_0^{DU/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} \left| W\left(\frac{v - \hat{\theta}_{i,m}}{h_m}\right) - W\left(\frac{v - \theta_i}{h_m}\right) \right| dv. \qquad (11)
\end{aligned}$$

By applying the Mean Value Theorem (Rudin, 1976, Theorem 5.10, p. 108) on the right hand side of (11), we obtain

$$\begin{aligned}
&\int_0^{DU/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} \left| W\left(\frac{v - \hat{\theta}_{i,m}}{h_m}\right) - W\left(\frac{v - \theta_i}{h_m}\right) \right| dv \\
= &\int_0^{DU/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} \frac{1}{h_m} \left| W'\left(\frac{v - \rho_i}{h_m}\right) \right| \left| \hat{\theta}_{i,m} - \theta_i \right| dv, \qquad (12)
\end{aligned}$$

where ρ_i lies between θ_i and $\hat{\theta}_{i,m}$, for $i = 1, \ldots K$. It follows from (11) and (12) that

$$\begin{aligned}
&\int_0^{\tilde{\eta}_m^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\hat{\theta}_{i,m}}{h_m}\right) - \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) \right| dv \\
\leq & \int_0^{DU/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} \frac{1}{h_m} \left| W'\left(\frac{v-\rho_i}{h_m}\right) \right| \left| \hat{\theta}_{i,m} - \theta_i \right| dv \\
\leq & \int_0^{DU/t} \frac{M_1}{K_m h_m^2} \sum_{i=1}^{K_m} \left| \hat{\theta}_{i,m} - \theta_i \right| dv = \frac{DUM_1}{tK_m h_m^2} \sum_{i=1}^{K_m} \left| \hat{\theta}_{i,m} - \theta_i \right|, \qquad (13)
\end{aligned}$$

where $M_1 > 0$ is a constant such that $W'(x) \leq M_1$ for all x. Since $\omega_1 \in B_1^*$, we obtain after taking limit $m \to \infty$ on both sides of (13),

$$\begin{aligned}
&\lim_{m\to\infty} \int_0^{\tilde{\eta}_m^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\hat{\theta}_{i,m}}{h_m}\right) - \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) \right| dv \\
\leq & \lim_{m\to\infty} \frac{DUM_1}{tK_m h_m^2} \sum_{i=1}^{K_m} \left| \hat{\theta}_{i,m} - \theta_i \right| = 0.
\end{aligned}$$

We next consider the second term on the right hand side of (10). Since $\omega_1 \in B_2$ we have by Lemma A.1, $\frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) \leq M_2$ for some $M_2 > 0$. Therefore,

$$\left| \int_{\eta^{-1}(s)/t}^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) dv \right| \leq \left| \int_{\eta^{-1}(s)/t}^{\tilde{\eta}_m^{-1}(s)/t} M_2 dv \right| \qquad (14)$$

By taking limit $m \to \infty$ on both sides of (14) and using the fact that $\omega_1 \in B_3$, we obtain

$$\begin{aligned}
\lim_{m\to\infty} \left| \int_{\eta^{-1}(s)/t}^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) dv \right| &\leq \lim_{m\to\infty} \left| \frac{M_2}{t} (\tilde{\eta}_m^{-1}(s) - \eta^{-1}(s)) \right| \\
&= 0.
\end{aligned}$$

Finally, we consider the third term on the right hand side of (10). By using Theorem 7.16 of Rudin

(Rudin, 1976, p. 151) and the fact that $\omega_1 \in B_2$, we get

$$\begin{aligned} & \lim_{m\to\infty} \int_0^{\eta^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) - g(v) \right| dv \\ = & \int_0^{\eta^{-1}(s)/t} \lim_{m\to\infty} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) - g(v) \right| dv = 0. \end{aligned} \tag{15}$$

We have thus established that, whenever t is such that $\eta(ct) \leq s \leq \eta(Dt)$, the left hand side of (10) converges to 0 with probability 1.

Now, for a fixed $t > 0$ such that $s > \eta(Dt)$, we have

$$\begin{aligned} |\hat{R}_m(t) - R(t)| &= \left| \int_0^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\hat{\theta}_{i,m}}{h_m}\right) dv - 1 \right| \\ &= \left| \int_0^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\hat{\theta}_{i,m}}{h_m}\right) dv - \int_0^{\eta^{-1}(s)/t} g(v)dv \right| \\ &\leq \int_0^{\tilde{\eta}_m^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\hat{\theta}_{i,m}}{h_m}\right) - \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) \right| dv \\ &\quad + \left| \int_{\eta^{-1}(s)/t}^{\tilde{\eta}_m^{-1}(s)/t} \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) dv \right| \\ &\quad + \int_0^{\eta^{-1}(s)/t} \left| \frac{1}{K_m h_m} \sum_{i=1}^{K_m} W\left(\frac{v-\theta_i}{h_m}\right) - g(v) \right| dv, \end{aligned} \tag{16}$$

where $g(v) = 0$, for all $v \notin [C, D]$. We can treat the three terms on the right hand side of (16) in the same way as those of (10), and show that each term converges to 0 as $m \to \infty$.

Proof of convergence of $|\hat{R}_m(t) - R(t)|$ in the case $s < \eta(Ct)$ follows similar arguments. □

Acknowledgement

The authors are thankful to Dr. Subir Bhandari, Dr. Anup Dewanji and Dr. Mary C. Meyer for many helpful comments and suggestions.

References

Barvinek, E., Daler, I., and Francu, J. (1991). Convergence of sequence of inverse functions. *Archivum Mathematicum*, 27:1013–1033.

Bhuyan, P. and Dewanji, A. (2015). Reliability computation under dynamic stress strength modeling with cumulative stress and strength degradation. *Communications in Statistics - Simulation and Computation, DOI:10.1080/03610918.2015.1057288*.

Bhuyan, P. and Dewanji, A. (2017). Estimation of system reliability for dynamic stress-strength modeling with cumulative stress and strength degradation. *Statistics: A Journal of Theoretical and Applied Statistics, DOI: 10.1080/02331888.2016.1277224*.

Devroye, L. and Gyrfi, L. (1985). *Nonparametric Density Estimation: the L1 view*. John Wiely & Sons, Inc.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans (Regional Confidence Series in Applied Mathematics, No. 38)*. Society of Industrial and Applied Mathematics.

Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1:54–77.

Epperson, J. F. (2013). *An Introduction to Numerical Mthods And Analysis*. John Wiely & Sons, Inc.

Gorijan, N., Ma, L., Mittinty, M., Yarlagadda, P., and Sun, Y. (2009). A review on degradation models in reliability analysis. *Proceedings of the 4th World Congress on Engineering Asset Management*.

Hardle, W., Muller, M., Sperlinch, S., and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer-Verlag.

Hudak, S. J., s, A. S., Bucci, R. J., and Malcolm, R. C. (1978). Development of standard methods of testing and analyzing fatigue crack growth rate data. *Technical Report AFML-TR-78-40: Westinghouse R & D Center, Westinghouse Electric Corporation*.

Kneip, A. and Gasser, T. (1988). Convergence and consistency results for self-modeling nonlinear regression. *The Annals of Statistics*, 16:82–112.

Lawless, J. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10:213–227.

Lawton, W. H., Sylvestre, E. A., and Maggio, M. S. (1972). Self modeling nonlinear regression. *Technometrics*, 14:513–532.

Lu, C. J. and Meeker, W. Q. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35:161–174.

Lu, J.-C., Park, J., and Yang, Q. (1997). Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics*, 39:391–400.

Meyer, M. C. (2008). Inference using shape-restricted regression splines. *Annals of Applied Statistics*, 2:425–441.

Nakagawa, T. (2007). *Shock and Damage Models in Reliability Theory*. SpringerVerlag London Ltd.

Nikulin, M. S., Limnios, N., Balakrishnan, N., Kahle, W., and Huber-Carol, C. (2010). *Advances in Degradation Modeling*. Birkhauser Boston.

Park, C. and Padgett, W. J. (2005). Accelerated degradation models for failure based on geometric brownian motion and gamma processes. *Lifetime Data Analysis*, 11:511–527.

Piessens, R., Doncker-Kapenga, E. D., Uberhuber, C. W., and Kahaner, D. K. (1983). *Quadpack a Subroutine Package for Automatic Integration*. Springer-Verlag.

Ramsay, J. O. (1988). Monotonic regression splines in action. *Statistical Sciences*, 3:425–441.

Rudin, W. (1976). *Principles of Mathematical Analysis*. MacGraw Hill Education Private Limited.

Silverman, B. W. (1978). Weak and strong consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, 6:177–184.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

Sobczyk, K. and Trbicki, J. (1989). Modelling of random fatigue by cumulative jump processes. *Engineering Fracture Mechanics*, 34:477–493.