11-2010

Bayesian Methods for Accelerated Destructive Degradation Test Planning

Ying Shi
Iowa State University

William Q. Meeker
Iowa State University, wqmeeker@iastate.edu

Bayesian Methods for Accelerated Destructive Degradation Test Planning

Abstract

Accelerated Destructive Degradation Tests (ADDTs) provide timely product reliability information in practical applications. This paper describes Bayesian methods for ADDT planning under a class of nonlinear degradation models with one accelerating variable. We use a Bayesian criterion based on the estimation precision of a specified failure-time distribution quantile at use conditions to find optimum test plans. A large-sample approximation for the posterior distribution provides a useful simplification to the planning criterion. The general equivalence theorem (GET) is used to verify the global optimality of the numerically optimized test plans. Optimum plans usually provide insight for constructing compromise plans which tend to be more robust and practically useful. We present a numerical example with a log-location-scale distribution to illustrate the Bayesian test planning methods and to investigate the effects of the prior distribution and sample size on test planning results.

Keywords

compromise plan, general equivalence theorem, large-sample approximation, log-location-scale distribution, optimum plan

Disciplines

Statistics and Probability

Comments

This preprint was published as Ying Shi and W.Q. Meeker, "Bayesian Methods for Accelerated Destructive Degradation Test Planning" *IEEE Transactions on Reliability* (2011): 245-253, doi: 10.1109/TR.2011.2170115.

Bayesian Methods for Accelerated Destructive Degradation Test Planning

Ying Shi
Dept. of Statistics
Iowa State University
Ames, IA 50011
yshi@iastate.edu

William Q. Meeker
Dept. of Statistics
Iowa State University
Ames, IA 50011
wqmeeker@iastate.edu

Abstract

Accelerated Destructive Degradation Tests (ADDTs) provide timely product reliability information in practical applications. This paper describes Bayesian methods for ADDT planning under a class of nonlinear degradation models with one accelerating variable. We use a Bayesian criterion based on the estimation precision of a specified failure-time distribution quantile at use conditions to find optimum test plans. A large-sample approximation for the posterior distribution provides a useful simplification to the planning criterion. The general equivalence theorem (GET) is used to verify the global optimality of the numerically optimized test plans. Optimum plans usually provide insight for constructing compromise plans which tend to be more robust and practically useful. We present a numerical example with a log-location-scale distribution to illustrate the Bayesian test planning methods and to investigate the effects of the prior distribution and sample size on test planning results.

Key Words: Compromise plan; General equivalence theorem; Large-sample approximation; Log-location-scale distribution; Optimum plan; Reliability.

1 Introduction

1.1 Background and Motivation

Many modern high-quality products are expected to perform their functions properly for years or even decades. Traditional life tests result in few or no failures, providing little or no information about reliability. *Accelerated destructive degradation tests* (ADDTs) are sometimes used in manufacturing industries to obtain reliability information more quickly. Usually, variables like temperature, voltage, or pressure can be used as accelerating variables to accelerate the degradation of a material or product. In designing an experiment, decisions must be made before data collection, and data collection is usually restricted by limited resources. Careful test planning is crucial for efficient use of limited resources: test time, test units, and test facilities. The basic idea in designing an experiment is that statistical inference for the quantities of interest can be improved by selecting appropriate test conditions to minimize or otherwise control the variability of the estimator of interest. Generally, ADDT plans specify the test conditions of the accelerating variables, the amount of running time, and the corresponding allocation of test units to each condition. One can find an optimum test plan for a given criterion, such as the estimation precision of a failure-time distribution quantile at use conditions. Optimum test plans provide insight for constructing good practical test plans.

For some applications, specific information about the underlying models or parameters is usually available from past studies or empirical knowledge of the failure mechanism from previous experimentation. When using Bayesian design methods, a prior distribution is used to describe the available information on model parameters. The primary motivation for this paper is to address the need to use such prior information in accelerated destructive degradation test planning. Bayesian methods can be used to formally incorporate prior information into estimation and test planning, providing test plans with better statistical precision (i.e., smaller estimation variance).

1.2 Related Literature

Shi, Escobar, and Meeker (2009) describe methods for ADDT planning using a non-Bayesian approach and outline much of the related literature in this area. Chaloner and Verdinelli (1995) give a nice review of Bayesian design methods. Hamada, Martz, Reese, and Wilson (2001) show methods to find near-optimal Bayesian experimental designs for regression models using genetic algorithms. Clyde, Müller, and Parmigiani (1995) describe Bayesian design methods for a logistic regression model. There is an extensive literature on Bayesian accelerated life tests (ALTs) planning. Work of particular relevance to this paper includes the following. Polson (1993) provides a general decision theory for ALT Bayesian design problems. Zhang and Meeker (2006) present Bayesian methods for planning accelerated life tests (ALTs) to estimate a specific quantile of interest. In this paper, we follow the general Bayesian design framework proposed by Zhang and Meeker (2006), but apply it to ADDT planning.

1.3 Overview

The remainder of this paper is organized as follows. Section 2 presents the ADDT degradation model and provides the degradation distribution and failure-time distribution induced by the model. Section 3 describes the Bayesian planning criterion, prior distribution specification, and general equivalence theorem (GET) used to verify the optimality of test plans. Section 4, based on an application, illustrates the methods of finding Bayesian optimum plans and optimized compromise plans under different situations for the specification of prior information. Section 4 also investigates the effects that changing the amount of prior information and sample size will have on Bayesian test plans. Section 5 gives some conclusions and areas for future research.

2 Degradation Models

2.1 The Model

Shi, Escobar, and Meeker (2009) describe non-Bayesian methods of finding ADDT plans for an important class of destructive degradation models. This paper illustrates the Bayesian ADDT planning methods based on the same degradation model. The ADDT regression model is

$$Y = \beta_0 + \beta_1 \exp(\beta_2 x)\tau + \epsilon \tag{1}$$

where Y is the transformed degradation response, τ and x are known monotone increasing transformations of time t and accelerating variable AccVar, respectively. β_0 is the location parameter of the transformed degradation when $\tau = 0$. The degradation rate with respect to τ at the accelerating variable level x is $\beta_1 \exp(\beta_2 x)$. The sign of β_1 determines whether the degradation is increasing or decreasing over time. A positive value of β_1 corresponds to increasing degradation and a negative value of β_1 corresponds to decreasing degradation. ϵ is a residual deviation that describes unit-to-unit variability with $(\epsilon/\sigma) \sim \Phi(z)$, where $\Phi(z)$ is a completely specified cumulative distribution function (cdf). For example, one could use $\Phi_{\text{nor}}(z)$ for the standardized normal cdf, or $\Phi_{\text{sev}}(z)$ for the standardized smallest extreme value cdf. The model in (1) is nonlinear with respect to the parameters $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \sigma)$, and the elements of $\boldsymbol{\theta}$ are fixed but unknown.

2.2 Degradation Distribution and Failure-Time Distribution

At a given test condition of running time and accelerating variable level, the cdf for the transformed degradation Y is

$$F_Y(y;\tau,x) = \Pr(Y \le y;\tau,x) = \Phi\left[\frac{y-\mu(\tau,x)}{\sigma}\right],$$

where $\mu(\tau,x) = \beta_0 + \beta_1 \exp(\beta_2 x)\tau$.

For applications where performance degrades gradually as usage time increases, failure time, T is defined as the time when the degradation level reaches a specified critical level. This is known as a "soft failure" (see Section 13.4 of Meeker and Escobar 1998). Let y_{f} denote the transformed critical level for the degradation distribution at which failure is assumed to occur.

For a decreasing degradation (i.e., when β_1 is negative), the event that the failure time T is less than t ($T \le t$), is equivalent to the event that the observed transformed degradation Y is less than the transformed critical level y_{f} ($Y \le y_{\mathrm{f}}$) at time t. Then the failure time cdf is

$$\begin{aligned} F_T(t;x) &= \Pr(T \le t) = \Pr(Y \le y_{\mathrm{f}}) = F_Y(y_{\mathrm{f}};\tau,x) \\ &= \Phi\left[\frac{y_{\mathrm{f}}-\mu(\tau,x)}{\sigma}\right] = \Phi\left(\frac{\tau-\nu}{\varsigma}\right), \quad \text{for } t \ge 0 \end{aligned} \tag{2}$$

where

$$\nu = \frac{(\beta_0 - y_{\mathrm{f}})\exp(-\beta_2 x)}{\mid \beta_1 \mid} \quad \text{and} \quad \varsigma = \frac{\sigma \exp(-\beta_2 x)}{\mid \beta_1 \mid}.$$

For increasing degradation (i.e., when β_1 is positive), the derivation of the failure-time cdf is similar.

Let $h_t()$ denote the monotone increasing transformation function for time. That is, $\tau = h_t(t)$. From (2), the p quantile of the failure-time distribution for decreasing degradation is

$$t_p = \begin{cases} h_t^{-1}\left[\nu + \varsigma\Phi^{-1}(p)\right] & \text{if } p \ge \Phi\left(-\nu/\varsigma\right) \\ 0 & \text{otherwise.} \end{cases}$$

2.3 Reparametrization of the Model for Prior Distribution Specification

In our numerical computation for either estimation or test planning, we use an alternative set of "stable" parameters (as defined by Ross 1990). This reparameterization breaks the otherwise strong correlations between some pairs of estimators. It also speeds convergence of the estimation algorithms (both ML and MCMC). Another important advantage of the new parameters is that they have meaningful interpretations. This makes it easier to elicit marginal prior distributions from the engineers working in the area.

Let $\bar{x}$ denote the mean of the accelerating variable and let $\bar{\tau}$ denote the average transformed time. Then the model (1) can be reparameterized as

$$Y = \gamma_0 + \gamma_1 \left\{\exp\left[(x - \bar{x})\,\gamma_2\right]\tau - \bar{\tau}\right\} + \epsilon,$$

where γ_0 is the intercept of the average accelerating line (i.e., degradation line for $\bar{x}$) at $\bar{\tau}$; γ_1 is the slope of the average accelerating line; and γ_2 is the regression coefficient corresponding to the x variable. The vector $\boldsymbol{\varphi} = (\gamma_0, \gamma_1, \gamma_2, \sigma)$ denotes the stable parameters. The relationship between the stable parameters $\boldsymbol{\varphi}$ and the original parameters $\boldsymbol{\theta}$ can be expressed as $\gamma_0 = \beta_0 + \beta_1 \exp(\boldsymbol{\beta}_2 \bar{x})\bar{\tau}$, $\gamma_1 = \beta_1 \exp(\boldsymbol{\beta}_2 \bar{x})$, and $\gamma_2 = \boldsymbol{\beta}_2$.

3 Bayesian ADDT Planning

3.1 Fisher Information Matrix

We denote a test condition in terms of transformed time τ and transformed accelerating variable x by $\boldsymbol{v} = (\tau, x)$. An ADDT plan, denoted by $\boldsymbol{\xi}$, will specify a set of test conditions, $\boldsymbol{v}_i$, and the corresponding proportional allocation π_i of test units at each condition. A test plan $\boldsymbol{\xi}$ with r test conditions is denoted by

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{v}_1, & \pi_1 \\ \boldsymbol{v}_2, & \pi_2 \\ \vdots & \vdots \\ \boldsymbol{v}_r, & \pi_r \end{bmatrix},$$

where $\pi_i > 0$ and $\sum_{i=1}^{r} \pi_i = 1$.

Let $L_i(\boldsymbol{\varphi})$ denote the likelihood of a single observation at test condition $\boldsymbol{v}_i = (\tau_i, x_i)$. Then $L_i(\boldsymbol{\varphi}) = \frac{1}{\sigma}\,\phi\left[(Y_i - \mu_i)/\sigma\right]$, where $\mu_i = \gamma_0 + \gamma_1 \left\{\exp\left[(x_i - \bar{x})\,\gamma_2\right]\tau_i - \bar{\tau}\right\}$ and $\phi(z)$ is a standardized pdf. It can be shown that, under the stable parametrization $\boldsymbol{\varphi}$, the Fisher information for test plan

$\boldsymbol{\xi}$ is

$$\boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi}) = \mathrm{E}\left[-\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'}\right] = n\sum_{i=1}^{r} \pi_i \mathrm{E}\left[-\frac{\partial^2 \mathcal{L}_i}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'}\right] = \frac{n}{\sigma^2}\sum_{i=1}^{r} \pi_i \mathcal{F}_i,$$

where $\mathcal{L}_i = \log[L_i(\boldsymbol{\varphi})]$, n is the total sample size, and

$$\mathcal{F}_i = \begin{bmatrix} f_{11}\boldsymbol{u}_i\boldsymbol{u}_i' & f_{12}\boldsymbol{u}_i \\ f_{12}\boldsymbol{u}_i' & f_{22} \end{bmatrix}$$

is the scaled Fisher information for a single unit at $\boldsymbol{v}_i$. The vector $\boldsymbol{u}_i$ contains the partial derivatives of the degradation μ_i with respect to the parameters $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2)$ and can be expressed as

$$\boldsymbol{u}_i = \begin{bmatrix} \frac{\partial \mu_i}{\partial \gamma_0} \\ \frac{\partial \mu_i}{\partial \gamma_1} \\ \frac{\partial \mu_i}{\partial \gamma_2} \end{bmatrix} = \begin{bmatrix} 1 \\ \exp\left[\gamma_2(x_i - \bar{x})\right]\tau_i - \bar{\tau} \\ \gamma_1\,(x_i - \bar{x})\,\exp\left[\gamma_2(x_i - \bar{x})\right]\tau_i \end{bmatrix}.$$

The basic elements f_{11}, f_{12}, f_{22} can be computed using the LSINF algorithm (see Escobar and Meeker 1994). For the special case of a normal distribution and no censoring, the basic elements are $f_{11} = 1, f_{12} = 0, f_{22} = 2$.

3.2 Bayesian Planning Criterion

The appropriate criterion for test planning depends on the goal of the experiment. To plan an efficient experiment, one should specify a utility function reflecting the purpose of the experiment and then select a test plan that maximizes the expected utility. In most reliability applications, the objective is to estimate a particular quantile, such as the p quantile, of the failure-time distribution at use conditions, say, t_p. Because $h_t(t_p)$ is a monotone increasing function of t_p, we can use a Bayesian ADDT criterion based on the estimation precision of $h_t(t_p) = \boldsymbol{c}'\boldsymbol{\varphi}$, where $\boldsymbol{c} = \partial h_t(t_p)/\partial \boldsymbol{\varphi}$. For decreasing degradation, the elements of $\boldsymbol{c}$ are

$$\frac{\partial h_t(t_p)}{\partial \gamma_0} = -\frac{1}{\gamma_1 \exp[\gamma_2(x - \bar{x})]}, \qquad \frac{\partial h_t(t_p)}{\partial \gamma_1} = \frac{1}{\gamma_1}\left[\frac{\bar{\tau}}{\exp[\gamma_2(x - \bar{x})]} - h_t(t_p)\right],$$

$$\frac{\partial h_t(t_p)}{\partial \gamma_2} = -(x - \bar{x})h_t(t_p), \qquad \frac{\partial h_t(t_p)}{\partial \sigma} = -\frac{\Phi^{-1}(p)}{\gamma_1 \exp[\gamma_2(x - \bar{x})]}.$$

where $(\bar{x}, \bar{\tau})$ are as defined in Section 2.3. For increasing degradation, the elements of $\boldsymbol{c}$ are the same as those for decreasing degradation except that

$$\frac{\partial h_t(t_p)}{\partial \sigma} = -\frac{\Phi^{-1}(1-p)}{\gamma_1 \exp[\gamma_2(x-\bar{x})]} \ .$$

We define the utility function as minus the posterior variance of $h_t(t_p)$. An optimum ADDT plan maximizes this utility function. Since the posterior variance depends on the unobserved data, a marginal expectation of the posterior variance over all possible data can be used as an objective function under a Bayesian test planning criterion. Generally, approximations must be used for the posterior variance because the exact expected utility is, in general, a complicated integral that has no closed form and that is computationally intractable. When sample sizes are reasonably large, the posterior variance can be expressed as a simple function of information from the prior distribution and the data. Let $\boldsymbol{S}$ denote the variance-covariance matrix of the prior distribution for $\boldsymbol{\varphi}$. Then $\boldsymbol{S}^{-1}$ is the prior precision matrix for $\boldsymbol{\varphi}$. Let $p(\boldsymbol{\varphi})$ denote the joint prior distribution for $\boldsymbol{\varphi}$. Following the general approach in Zhang and Meeker (2006), for large-sample approximations, the Bayesian test planning criterion is to find a test plan $\boldsymbol{\xi}$ that maximizes the objective function

$$\Psi(\boldsymbol{\xi}) = -\int \boldsymbol{c}'[\boldsymbol{S}^{-1} + \boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})]^{-1}\boldsymbol{c}\, p(\boldsymbol{\varphi})\, d\boldsymbol{\varphi}. \qquad (3)$$

A similar approximation was used in Clyde, Müller, and Parmigiani (1995).

To estimate t_p precisely, the confidence interval for t_p should be as narrow as possible. Denote the ML estimate of t_p by $\widehat{t}_p$. An approximate $100(1-\alpha)\%$ confidence interval for $\log(t_p)$ is

$$\log(\widehat{t}_p) \pm z_{(1-\alpha/2)}\sqrt{\widehat{\text{Var}}\left[\log(\widehat{t}_p)\right]} = \log(\widehat{t}_p) \pm \log(\widehat{R}).$$

Exponentiation yields an approximate confidence interval for t_p,

$$[\widehat{t}_p/\widehat{R},\ \widehat{t}_p\widehat{R}]$$

where

$$\widehat{R} = \exp\left[z_{(1-\alpha/2)}\sqrt{\widehat{\text{Var}}\left[\log(\widehat{t}_p)\right]}\ \right]. \qquad (4)$$

For Bayesian test planning, the estimated variance $\widehat{\mathrm{Var}}\left[\log(\widehat{t}_p)\right]$ in (4) is replaced by the large-sample approximation of the expected posterior variance of $\log(t_p)$. Similar to deriving the objective function in (3), the large-sample approximation of the expected posterior variance of $\log(t_p)$ can be expressed as

$$\int \frac{1}{t_p^2}\left(\frac{\partial t_p}{\partial h_t(t_p)}\right)^2 \boldsymbol{c}'[\boldsymbol{S}^{-1} + \boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})]^{-1}\boldsymbol{c}\, p(\boldsymbol{\varphi})\, d\boldsymbol{\varphi}.$$

This gives

$$R = \exp\left[z_{(1-\alpha/2)}\sqrt{\int \frac{1}{t_p^2}\left(\frac{\partial t_p}{\partial h_t(t_p)}\right)^2 \boldsymbol{c}'[\boldsymbol{S}^{-1} + \boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})]^{-1}\boldsymbol{c}\, p(\boldsymbol{\varphi})\, d\boldsymbol{\varphi}}\,\right].$$

We call R the "precision factor." The upper (lower) endpoint of the confidence interval for t_p is approximately $100(R-1)\%$ larger (smaller) than the ML estimate $\widehat{t}_p$. Minimizing the R precision factor is equivalent to maximizing the objective function in (3). Because R is easier to interpret as a measure of precision for a positive parameter t_p, we can use it for the comparisons among different Bayesian ADDT plans.

3.3 The Prior Distribution

Prior distributions for the parameters can be obtained from engineering judgement, previous experiments and past data. In Bayesian experimental design, it is often necessary to specify two different prior distributions:

- The prior distribution to be used to design the experiment,
- The prior distribution to be used in the inference.

Some papers, for example, Tsutakawa (1972) and Etzione and Kadane (1993), have discussed the need to use different prior distributions for the design and for the inference. One motivation for this need is that the risk of those conducting the experiment is different from that those who are concerned with the accuracy of the inference from the experiment. This idea can be seen from the objective function (3), in which the precision matrix $\boldsymbol{S}^{-1}$ quantifies the prior information for the inference, and $p(\boldsymbol{\varphi})$ represents the prior distribution for the design. Generally, the prior distribution to be used to design the experiment must be informative for all parameters. An experimenter may, however, prefer to use a non-informative prior for the inference by having $\boldsymbol{S}^{-1}$ be identically zero, as was done in Chaloner and Larntz (1989).

We will explore several different combinations of prior distributions in doing Bayesian ADDT planning. In particular, we will use a point-mass prior $p_0(\boldsymbol{\varphi})$, an informative prior for all parameters $p_1(\boldsymbol{\varphi})$, an informative prior for partial parameters $p_2(\boldsymbol{\varphi})$, and a non-informative prior $p_3(\boldsymbol{\varphi})$. Table 1 summarizes different cases that we will use for test planning in terms of the specification of the prior distribution for the design $p(\boldsymbol{\varphi})$ and for the inference $\boldsymbol{S}^{-1}$ separately. As mentioned above, the non-informative prior $p_3(\boldsymbol{\varphi})$ for the inference is implemented by setting $\boldsymbol{S}^{-1} = \mathbf{0}$. Note that informative prior distributions are used for the design in all cases. This is because that test planning criteria are highly sensitive to the particular form of any diffuse prior for the design. Some information about the model parameters is required in order to obtain sensible test planning results.

Table 1: Prior distribution specification.

Case	Design $p(\boldsymbol{\varphi})$	Inference $\boldsymbol{S}^{-1}$
A	point-mass $p_0(\boldsymbol{\varphi})$	non-informative $p_3(\boldsymbol{\varphi})$
B	informative for all parameters $p_1(\boldsymbol{\varphi})$	non-informative $p_3(\boldsymbol{\varphi})$
C	informative for all parameters $p_1(\boldsymbol{\varphi})$	informative for partial parameters $p_2(\boldsymbol{\varphi})$
D	informative for all parameters $p_1(\boldsymbol{\varphi})$	informative for all parameters $p_1(\boldsymbol{\varphi})$

3.4 General Equivalence Theorem

Whittle's (1973) general equivalence theorem (GET) can be used to check the optimality of test plans. The outputs for an application of the GET can also suggest that an optimum plan is unique or not or suggest why a given plan is not optimum. We will use the GET to check the optimality of the test plans that we find.

The directional derivative, Λ, of Ψ at $\boldsymbol{\xi}$ and in the direction of an alternative plan $\boldsymbol{\eta}$ is defined as

$$\Lambda(\boldsymbol{\xi}, \boldsymbol{\eta}) = \lim_{\delta \to 0^+} \frac{\Psi[(1-\delta)\boldsymbol{\xi} + \delta\boldsymbol{\eta}] - \Psi(\boldsymbol{\xi})}{\delta}.$$

In Bayesian ADDT planning, the derivative function of (3) at $\boldsymbol{\xi}$ can be derived as

$$\Lambda(\boldsymbol{\xi}, \boldsymbol{\eta}) = \int \boldsymbol{c}'\boldsymbol{V}(\boldsymbol{\varphi}, \boldsymbol{\xi})\boldsymbol{V}(\boldsymbol{\varphi}, \boldsymbol{\eta})^{-1}\boldsymbol{V}(\boldsymbol{\varphi}, \boldsymbol{\xi})\boldsymbol{c}\, p(\boldsymbol{\varphi})\, d\boldsymbol{\varphi} + \Psi(\boldsymbol{\xi}), \tag{5}$$

where $\boldsymbol{V}(\boldsymbol{\varphi}, \boldsymbol{\xi}) = [\boldsymbol{S}^{-1} + \boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})]^{-1}$. Let $\boldsymbol{\xi}_{\boldsymbol{v}}$ be a singular test plan that puts all units at a single test condition $\boldsymbol{v}$. Suppose that a given ADDT plan $\boldsymbol{\xi}$ has r test conditions, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$. Then this plan $\boldsymbol{\xi}$ is an optimum plan for the Bayesian criterion if it satisfies $\Lambda(\boldsymbol{\xi}, \boldsymbol{\xi}_{\boldsymbol{v}_1}) = \Lambda(\boldsymbol{\xi}, \boldsymbol{\xi}_{\boldsymbol{v}_2}) = \ldots = \Lambda(\boldsymbol{\xi}, \boldsymbol{\xi}_{\boldsymbol{v}_r}) = 0$ and $\Lambda(\boldsymbol{\xi}, \boldsymbol{\xi}_{\boldsymbol{v}}) \leq 0$ for any other singular plan $\boldsymbol{\xi}_{\boldsymbol{v}}$ in the experimental region.

4 Numerical Example

In this section, we use the adhesive bond B application in Shi, Escobar, and Meeker (2009) to illustrate the Bayesian ADDT planning methods.

4.1 The Application

Adhesive bond B was to be evaluated for use in manufacturing an inkjet cartridge. The adhesive is used to bond a protective coating to protect the printhead of the inkjet cartridge. When the adhesive becomes weak, there can be delamination and ink can attack the electronics in the printhead, causing failure. The degradation response is the strength (in Newtons) of the adhesive bond over time. There was special interest in estimating the time at which 1% of the product would have a strength below 40 Newtons when operating at the use condition of 25 °C (i.e., the 0.01 quantile of the failure-time distribution). For this application, the accelerating variable is temperature, and the degradation strength model is as given in (1) with

$$
\begin{aligned}
Y &= \log(\text{Strength in Newtons}) \\
\tau &= \sqrt{\text{Time in Weeks}} \\
x &= -\frac{11605}{\text{Temperature in }^{\circ}\text{C} + 273.15} \\
(\epsilon/\sigma) &\sim \Phi_{\text{nor}}(z).
\end{aligned}
$$

Figure 1 provides a visualization of the degradation distributions as a function of time at 25 °C for specific values of the parameters $\beta_0, \beta_1, \beta_2$, and σ. The strength axis is a logarithmic axis and the time axis is a square root axis, corresponding to model assumptions that imply linear degradation in these scales. The horizontal line at 40 Newtons is the failure-definition degradation level for this application. At each point in time with a vertical line, the shaded area below the horizontal line is the failure probability at that time.

The original ADDT plan for this application used 88 test units. As a baseline, 8 units with no aging were measured at the start of the experiment. A total of 80 additional units were aged and

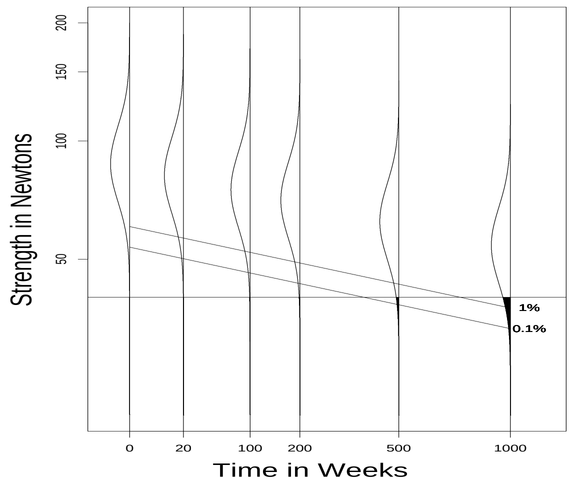

Figure 1: Adhesive Bond B degradation distributions as a function of time at 25 °C.

measured according to the temperature and time schedule in Table 2.

Table 2: Original ADDT plan. The — indicates that at time 0, the level of temperature has no effect on the model.

Temperature °C	Weeks 0	2	4	6	12	16	Totals
—	8						8
50		8	0	8	8	7	31
60		6	0	6	6	6	24
70		6	6	4	9	0	25
Totals	8	20	6	18	23	13	88

For the reparameterization to stable parameters, the centroid of the accelerating variable $\bar{x}$ and the average transformed time $\bar{\tau}$ are obtained based on this original test plan. In particular, we use $\bar{x} = \sum \pi_i x_i$ and $\bar{\tau} = \sum \pi_i \tau_i$, where π_i is the proportional allocation at test condition (τ_i, x_i). The numerical values for this centroid are $\bar{x} = -34.833$ and $\bar{\tau} = 2.455$.

4.2 Specification of the Prior Distribution

For this application, parameter γ_2 can be interpreted as an effective activation energy. Given previous experience with a failure mechanism, engineers often have useful prior information for this parameter. Prior distributions for the other three parameters γ_0, γ_1, and σ could be obtained from previous experiments and past data (e.g., Escobar et al. 2003). This application has decreasing degradation so the degradation slope γ_1 is negative. The effective activation energy γ_2 and standard deviation of the residuals σ are positive parameters. Hence, the uncertainty for the four parameters $(\exp(\gamma_0), -\gamma_1, \gamma_2, \sigma)$ can be described by an independent multivariate lognormal distribution with specified 0.01 and 0.99 quantiles (any particular quantiles could be used, but we find these particular values to be useful in eliciting prior information from engineers). Due to the prior specification issues mentioned in Section 3.3, we will illustrate Bayesian ADDT planning by specifying four possible prior distributions [i.e., a point-mass $p_0(\boldsymbol{\varphi})$, informative for all parameters $p_1(\boldsymbol{\varphi})$, informative for parameter γ_2 only $p_2(\boldsymbol{\varphi})$, and a non-informative prior $p_3(\boldsymbol{\varphi})$ (implemented by setting $\boldsymbol{S}^{-1} = \mathbf{0}$)].

Point-mass Prior Distribution $p_0(\boldsymbol{\varphi})$: Shi, Escobar, and Meeker (2009) describe non-Bayesian methods of finding ADDT plans for the same adhesive bond B application. The locally optimum test plans developed in that paper require the specification of planning values for the model parameters. The planning values of the parameters used there will be used to specify a point-mass prior distribution for the parameters. This will allow us to compare, directly, the non-Bayesian and Bayesian test plans. A point-mass prior distribution is assumed to be highly informative centered around the planning values. Thus the point-mass prior distribution can be approximately specified by normal distributions with the mean at the planning values and a small standard deviation. For this application, the approximate point-mass prior distribution is $\gamma_0 \sim N(3.97, 0.002), \log(-\gamma_1) \sim N(1.59, 0.002), \log(\gamma_2) \sim N(-0.45, 0.002)$, and $\log(\sigma) \sim N(-1.84, 0.002)$.

Informative Prior Distribution $p_1(\boldsymbol{\varphi})$: Table 3 summarizes the independent multivariate lognormal distribution for parameters $(\exp(\gamma_0), -\gamma_1, \gamma_2, \sigma)$ and the corresponding log-location scale hyperparameters for the informative prior information.

Table 3: Informative prior distributions specified by marginal lognormal distributions.

	Prior specification		Hyperparameter	
Parameter	0.01 quantile	0.99 quantile	mean	standard deviation
$\exp(\gamma_0)$	51	54	3.96	0.013
$-\gamma_1$	0.15	0.25	−1.64	0.110
γ_2	0.55	0.75	−0.44	0.067
σ	0.1	0.2	−1.96	0.149

Partial Informative Prior Distribution $p_2(\boldsymbol{\varphi})$: For this application, engineers often have access to highly informative prior information for the effective activation energy (i.e., parameter γ_2

in our model). Often there is strong information about this parameter, based on previous experience and knowledge of the physics or chemistry of the failure mechanism (indeed, in some applications of accelerated testing, the effective activation energy is assumed to be known). For other parameters, often, the prior information is limited. Table 4 summarizes the marginal lognormal distribution for the parameters $(\exp(\gamma_0), -\gamma_1, \gamma_2, \sigma)$ and their corresponding log-location scale hyperparameters for the partial informative prior information.

Table 4: Partially informative prior distributions specified by marginal lognormal distributions.

	Prior specification		Hyperparameter	
Parameter	0.01 quantile	0.99 quantile	mean	standard deviation
$\exp(\gamma_0)$	40	74	4	0.13
$-\gamma_1$	0.05	0.5	-1.85	0.495
γ_2	0.55	0.75	-0.44	0.067
σ	0.05	0.3	-2.1	0.385

4.3 Bayesian Optimum Test Plans

There are usually practical constraints for test planning. For the adhesive bond B application, the constraints are:

- 88 test units available for the sample size,
- 70 °C is the maximum temperature,
- 16 weeks is maximum available time for testing.

After the specification of the prior distributions, we can explore Bayesian test planning for the different cases listed in Table 1. Optimum test plans obtained under each situation in the following part are all continuous test plans (a continuous test plan is one that has non-integer allocations because optimization was done without integer constraints on the number of units allocated to the test conditions).

Case A: In nonlinear models, the non-Bayesian optimum plans are expected to be special cases of Bayesian optimum plans which correspond to point-mass prior distributions for the design and non-informative prior distributions for the inference. Shi, Escobar, and Meeker (2009) present a

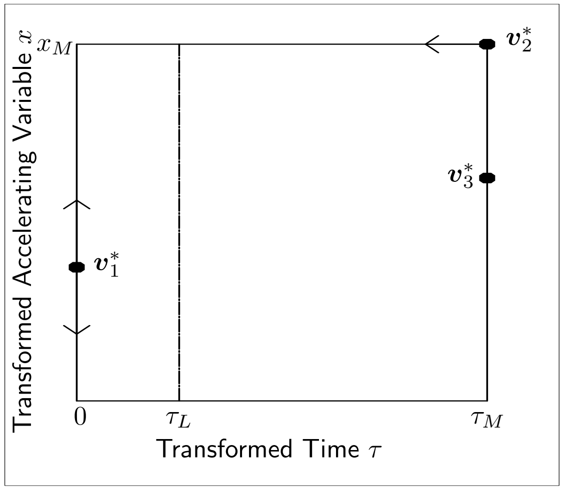

Figure 2: Optimum plan structure.

non-Bayesian optimum ADDT plan structure in terms of transformed accelerating variable x and transformed time τ, as shown in Figure 2.

For Bayesian test planning with a point-mass prior distribution, we explore optimum plans following the same structure as Figure 2. x_M and τ_M are the maximum transformed accelerating variable and transformed time, respectively. Because (1) has three parameters, there are expected to be three test conditions for an optimum plan. The three test conditions include a baseline condition $\boldsymbol{v}_1^*$, a highest stress test condition $\boldsymbol{v}_2^*$ at x_M and τ_M, and a right boundary test condition $\boldsymbol{v}_3^*$ at τ_M. The variable x^* for the condition $\boldsymbol{v}_3^*$ and two proportional allocations π_1^* , π_2^* are optimized to maximize the objective function in (3). For this case, the resulting Bayesian optimum plan $\boldsymbol{\xi}^*$ is shown in Table 5. As expected, this optimum plan from Bayesian test planning with a point-mass prior for the design and a non-informative prior for the inference is the same as the one obtained from non-Bayesian test planning methods in Shi, Escobar and Meeker (2009). As from (3), when the prior information for the inference $\boldsymbol{S}^{-1}$ is $\mathbf{0}$, the absolute value of the objective function is inversely proportional to the sample size n which is reflected in $\boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})$. Under the sample size of $n = 88$, the objective function of this optimum plan, $\Psi(\boldsymbol{\xi}^*)$, is -20.43, and the R precision factor is 1.907. The optimality of this Bayesian test plan can be verified using the GET, as described in Section 3.4. The plot of the directional derivatives $\Lambda(\boldsymbol{\xi}^*, \boldsymbol{\xi}_{\boldsymbol{v}})$ is the same as the one obtained from the non-Bayesian method in Shi, Escobar and Meeker (2009), and is omitted here to avoid redundancy.

Table 5: An optimum ADDT plan $\boldsymbol{\xi}^*$ corresponding to Bayesian methods with a point-mass prior distribution for the design and a non-informative prior distribution for the inference. The — indicates that at time 0, the level of temperature has no effect on the model.

Optimum Test Condition	Weeks	Temperature °C	Proportional Allocations
$\boldsymbol{v}_1^*$	0	—	0.203
$\boldsymbol{v}_2^*$	16	70	0.162
$\boldsymbol{v}_3^*$	16	54.765	0.635

Case B: Following the same steps used in case A, a Bayesian optimum plan $\boldsymbol{\xi}^*$ for case B is shown in Table 6. This optimum test plan is close to the one obtained in case A. It verifies the conclusion from Chaloner and Larntz (1989), that for prior distributions with support over a small region, Bayesian optimum plans are similar to non-Bayesian optimum plans. Again for this case, the absolute value of the objective function is inversely proportional to the sample size. Under the sample size of 88, the objective function of this optimum plan, $\Psi(\boldsymbol{\xi}^*)$, is -24.07, and the R precision factor is 1.881.

Table 6: A Bayesian optimum ADDT plan $\boldsymbol{\xi}^*$ with an informative prior for the design and a non-informative prior for the inference. The — indicates that at time 0, the level of temperature has no effect on the model.

Optimum Test Condition	Weeks	Temperature °C	Proportional Allocations
$\boldsymbol{v}_1^*$	0	—	0.213
$\boldsymbol{v}_2^*$	16	70	0.162
$\boldsymbol{v}_3^*$	16	55.331	0.625

Cases C and D: For these two cases, we incorporate some prior information for inference by specifying an informative $\boldsymbol{S}^{-1}$. From the objective function (3), we can see that the sample size n reflected in $\boldsymbol{I}_{\boldsymbol{\varphi}}(\boldsymbol{\xi})$, relative to the amount of prior information, plays a role in the posterior distribution and test planning. When the sample size is large, the posterior distribution will tend to be driven by the data and will not be sensitive to the prior distribution for the inference. In contrast, when the sample size is small, the prior distribution will have more effect on both the posterior distribution and the design. Hence, we investigate Bayesian optimum plans under two different sample sizes: a small sample size $n = 88$ and a large one $n = 300$. As before, we explore optimum plans following the structure in Figure 2. Tables 7 and 8 show test conditions, the values of the objective function, and the R precision factors for the optimum plans of two cases under sample sizes 88 and 300, respectively. Note that the first two columns in both tables are common for the two cases. The optimality of these test plans can be verified using the GET in the same way as was done above. Figure 3 shows a plot of the directional derivatives $\Lambda(\boldsymbol{\xi}^*, \boldsymbol{\xi}_{\boldsymbol{v}})$ for case D with $n = 88$. The shapes of the directional derivatives plots for the other cases are similar except for magnitude changes in the values.

Table 7: Bayesian Optimum ADDT plans $\boldsymbol{\xi}^*$ for cases C and D under sample size of $n = 88$. The — indicates that at time 0, the level of temperature has no effect on the model.

All cases		Case C				Case D			
Prior for the Design Prior for the Inference		Informative Prior $p_1(\boldsymbol{\varphi})$ Partial Informative Prior $p_2(\boldsymbol{\varphi})$				Informative Prior $p_1(\boldsymbol{\varphi})$ Informative Prior $p_1(\boldsymbol{\varphi})$			
Conditions	Weeks	°C	π^*	$\Psi(\boldsymbol{\xi}^*)$	R	°C	π^*	$\Psi(\boldsymbol{\xi}^*)$	R
$\boldsymbol{v}_1^*$	0	—	0.159			—	0.185		
$\boldsymbol{v}_2^*$	16	70	0.143	−15.66	1.664	70	0.200	−14.07	1.617
$\boldsymbol{v}_3^*$	16	55.061	0.698			55.299	0.615		

Table 8: Bayesian Optimum ADDT plans $\boldsymbol{\xi}^*$ for cases C and D under sample size of $n = 300$. The — indicates that at time 0, the level of temperature has no effect on the model.

All cases		Case C				Case D			
Prior for the Design Prior for the Inference		Informative Prior $p_1(\boldsymbol{\varphi})$ Partial Informative Prior $p_2(\boldsymbol{\varphi})$				Informative Prior $p_1(\boldsymbol{\varphi})$ Informative Prior $p_1(\boldsymbol{\varphi})$			
Conditions	Weeks	°C	π^*	$\Psi(\boldsymbol{\xi}^*)$	R	°C	π^*	$\Psi(\boldsymbol{\xi}^*)$	R
$\boldsymbol{v}_1^*$	0	—	0.199			—	0.204		
$\boldsymbol{v}_2^*$	16	70	0.156	−6.00	1.371	70	0.174	−5.79	1.362
$\boldsymbol{v}_3^*$	16	55.009	0.645			55.315	0.622		

From Tables 7 and 8, we can see that, with a large sample size, the prior distribution for the inference will not strongly influence Bayesian test planning, as compared to test plans using a small sample size. As the sample size becomes larger and larger, the prior information becomes less influential in the inference, and the resulting optimum plans should approach a plan for which the available prior information is to be used in test planning but not for the inference (i.e., the test plan obtained in Case B).

For all of the cases mentioned above, alternative optimum plans exist. As seen from Figure 3, the directional derivatives $\Lambda(\boldsymbol{\xi}^*, \boldsymbol{\xi}_{\boldsymbol{v}})$ are all zero when the alternative singular plan $\boldsymbol{\xi}_{\boldsymbol{v}}$ puts all the

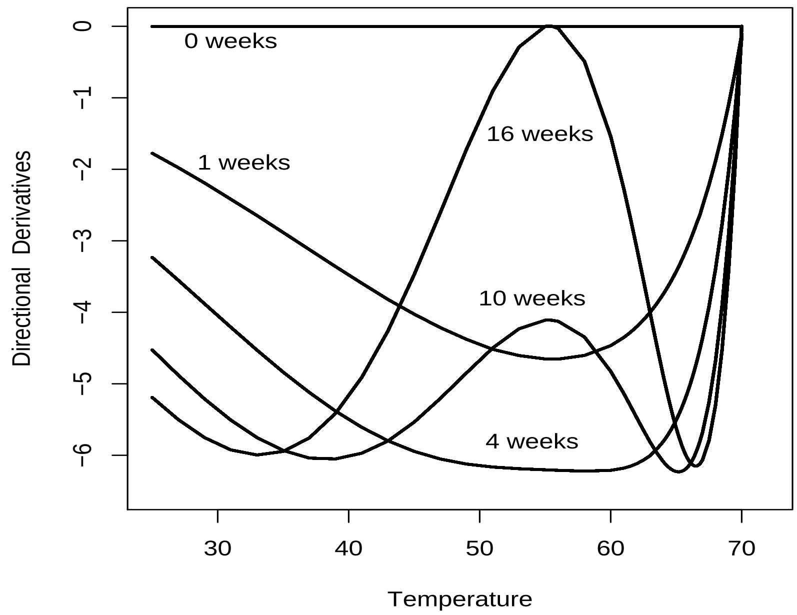

Figure 3: Directional derivatives $\Lambda(\boldsymbol{\xi}^*, \boldsymbol{\xi}_{\boldsymbol{v}})$ of the optimum plan $\boldsymbol{\xi}^*$ as a function of temperature and time for case D under a sample size $n = 88$.

test units at a test condition with 70 °C as the temperature level. This indicates the existence of alternative optimum test plans that can be obtained by moving the test condition $\boldsymbol{v}_2^*$ along the upper temperature boundary in Figure 2. Using the plan specification notation in Section 3.1, an alternative optimum plan $\boldsymbol{\xi}^a$ can be expressed in terms of π_1^*, π_2^*, x^* of the initial optimum plan $\boldsymbol{\xi}^*$, which is given as

$$\boldsymbol{\xi}^a = \begin{bmatrix} \boldsymbol{v}_1 = (0, -), & \pi_1 = \pi_1^* + \pi_2^* - \pi_2^* \dfrac{\tau_M}{\tau_a} \\ \boldsymbol{v}_2 = (\tau_a, x_M), & \pi_2 = \pi_2^* \dfrac{\tau_M}{\tau_a} \\ \boldsymbol{v}_3 = (\tau_M, x^*), & \pi_3 = \pi_3^* \end{bmatrix}, \tag{6}$$

where $\tau_L \leq \tau_a \leq \tau_M$, and $\tau_L = \tau_M \pi_2^*/(\pi_1^* + \pi_2^*)$. The optimality of alternative plans can be proved in a way that is similar to the proof given in Appendix B.2 of Shi, Esboar and Meeker (2009), and is omitted here.

4.4 Bayesian Optimized Compromise Test Plans

Optimum plans have some disadvantages. For example, they tend to be highly sensitive to model specification errors and such plans provide little or no information about departures from the acceleration model. For this reason, it has been suggested (e.g., Chapter 6 of Nelson 1990, and Chapter 20 of Meeker and Escobar 1998) to construct compromise test plans that tend to be more robust and practical. An optimum plan can usually provide useful insight for obtaining good compromise test plans.

For the adhesive bond B application, Shi, Escobar, and Meeker (2009) propose an optimized compromise plan for non-Bayesian test planning. The idea there can also be used to find Bayesian optimized compromise plans. For the compromise plan, we allocate some test units at the baseline conditions and an equal proportion of units at each of nine additional equally-spaced test conditions. The nine equally-spaced test conditions use three fixed time levels at 12, 14, and 16 weeks and a fixed highest temperature level at 70 °C. The lowest temperature level is chosen to maximize the objective function $\Psi(\boldsymbol{\xi})$ in (3). The middle temperature level is the mean of the other two temperature levels. For case D, the informative prior distribution $p_1(\boldsymbol{\varphi})$ is used for both the design and for inference. With sample size of $n = 88$, after rounding in the allocations, the compromise plan has 7 units at the baseline and 9 units at each of the other nine test conditions. The optimized lowest temperature level is 53.2 °C and the middle temperature level is 61.6 °C. This Bayesian optimized compromise plan is presented in Table 9. The objective function $\Psi(\boldsymbol{\xi})$ for this compromise plan is -16.46. And the R precision factor for this plan is 1.676, compared with 1.617 for the corresponding optimum plan under case D with sample size $n = 88$, suggesting there is little loss in the estimation precision. For other cases listed in Table 1, similar Bayesian optimized compromise plans can be found in the same way but the details are not given here.

5 Conclusions and Areas for Future Research

Planning ADDTs with prior information is useful for making reliability inferences in practical applications. In this paper, we present Bayesian test planning methods for ADDT problems under an important class of nonlinear regression models when prior information is available on the model parameters. We use a Bayesian criterion based on the estimation precision of a failure-time distribution quantile at use conditions. A large-sample approximation provides a useful simplification for

Table 9: Bayesian optimized compromise ADDT plan for case D under sample size of $n = 88$. The — indicates that at time 0, the level of temperature has no effect on the model.

Temperature	Weeks				Totals
°C	0	12	14	16	
—	7				7
53.2		9	9	9	27
61.6		9	9	9	27
70		9	9	9	27
Totals	7	27	27	27	88

the posterior distribution. The GET is an important tool to verify that the numerically optimized plans are globally or near-globally optimum. We also examine the effects of changing the amount of prior information and sample size on doing Bayesian test planning.

The Bayesian methods illustrated in this paper can be extended to the ADDT planning problems with more complicated degradation models, such as models with multiple accelerating variables (e.g., temperature, humidity) or nonlinear relationships between degradation and time. For some products, there may be more than one failure mechanism. This can cause degradation observations to be right-censored, as described in Escobar et al. (2003). The statistical competing risk model (see David and Moeschberger 1978) can be used as the degradation model for such applications. Bayesian test planning methods could be used for ADDT problems with such competing risk models. In addition, Monte Carlo simulation methods could complement the results obtained from the large-sample approximation approach used in this paper. Such simulations are particularly useful for providing visualization of sampling variability resulting from different test plans, as illustrated in Chapter 20 of Meeker and Escober (1998).

References

Chaloner, K., and Larntz, K. (1989), "Optimal Bayesian Designs Applied to Logistic Regression Experiments," *Journal of Statistical Planning and Inference*, **21**, 191–208.

Clyde, M., Müller, P., and Parmigiani, G. (1995), "Optimal Design for Heart Defibrillators," in *Case Studies in Bayesian Statistics 2*, eds. C. Gatsonis, J. Hodges, R. E. Kass, and N. Singpurwalla, New York: Springer-Verlag, pp. 278–292.

Chaloner, K., and Verdinelli, I. (1995), “Bayesian Experimental Design: A Review,” *Statistical Science*, **10**, 273–304.

David, H. A., and Moeschberger, M. L. (1978), *The Theory of Competing Risks*, London: Griffin.

Escobar, L. A., Meeker, W. Q., Kugler, D. L., and Kramer, L. L. (2003), “Accelerated Destructive Degradation Tests: Data, Models, and Analysis,” Chapter 21 in *Mathematical and Statistical Methods in Reliability*, B. H. Lindqvist and K. A. Doksum, Editors, River Edge, NJ: World Scientific Publishing Company.

Etzione, R., and Kadane, J. B. (1993), “Optimal experimental design for another’s analysis,” *Journal of the American Statistical Association*, **88**, 1404–1411.

Hamada, M., Martz, H. F., Reese, C. S., and Wilson, A. G. (2001), ”Finding Near-Optimal Bayesian Experimental Designs via Genetic Algorithms,” *The American Statistician*, **55**, 175–181.

Meeker, W. Q., and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, New York: John Wiley & Sons.

Nelson, W. (1990), *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, New York: John Wiley & Sons.

Polson, N. G. (1993), “A Bayesian Perspective on the Design of Accelerated Life Tests,” in A. P. Basu: *Advances in Reliability*, New York: Elsevier, pp. 321–330.

Ross, Gavin J. S. (1990), *Nonlinear Estimation*, New York: Springer-Verlag.

Shi, Y., Escobar, L. A., and Meeker, W. Q. (2009), “Accelerated Destructive Degradation Test Planning,” *Technometrics*, **51**, 1–13.

Tsutakawa, R. K. (1972), “Design of an Experiment for Bioassay,” *Journal of the American Statistical Association*, **67**, 584–590.

Whittle, P. (1973), “Some General Points in the Theory of Optimal Experimental Design,” *Journal of the Royal Statistical Society*, Ser. B, **35**, 123–130.

Zhang, Y., and Meeker, W. Q. (2006), “Bayesian Methods for Planning Accelerated Life Tests,” *Technometrics*, **48**, 49–60.