A Bayesian Semiparametric Analysis of the Reliability and Maintenance of Machine Tools

Jason R. W. Merrick

Department of Statistical Sciences and
Operations Research
Virginia Commonwealth University
Richmond, VA 23284
(*jrmerric@vcu.edu*)

Refik Soyer

Department of Management Science
The George Washington University
Washington, DC 20052
(*soyer@gwu.edu*)

Thomas A. Mazzuchi

Department of Engineering Management & Systems Engineering
The George Washington University, Washington, DC 20052
(*mazzuchi@seas.gwu.edu*)

A Bayesian semiparametric proportional hazards model is presented to describe the failure behavior of machine tools. The semiparametric setup is introduced using a mixture of Dirichlet processes prior. A Bayesian analysis is performed on real machine tool failure data using the semiparametric setup, and development of optimal replacement strategies are discussed. The results of the semiparametric analysis and the replacement policies are compared with those under a parametric model.

KEY WORDS: Bayesian; Markov chain Monte Carlo; Mixtures of Dirichlet processes prior; Proportional hazards model; Semiparametric inference.

1. INTRODUCTION

The useful life of a machine tool is the duration of time that the tool maintains an acceptable quality of performance. As an alternative to replacing tools on observing unacceptable quality of performance, using planned replacements can reduce costs associated with in-service failures of machine tools, thereby resulting in increased productivity due to decreased downtime and scrapping of material and decreased inventory costs due to improved planning. Key to the determination of an optimal replacement strategy is the development of an adequate statistical model for tool life. The development of such a model is complicated because of the fact that the conditions under which a tool operates (called the *operational environment*) vary even for tools on the same shop floor. In addition, commercial engineering materials require variability within specified ranges of chemistry and mechanical properties for their economical production, corresponding to variability in the properties that degrade the tool during its operational life. Variabilities can also be expected in machine tool dynamics and materials-handling performance of the processing systems. Thus, each machine tool may exhibit inherent variabilities in its useful life.

In the early literature, the lack of a universally acceptable physical theory of tool failure led to the use of an empirical model describing the relationship between tool life and operational variables, including cutting speed, feed rate, and depth of cut. Taraman (1974) performed an experiment designed to estimate the parameters of this empirical model. Balakrishnan and DeVries (1985) extended this analysis to allow sequential updating of parameter estimates and inclusion of prior information in the estimation procedure. Mazzuchi and

Soyer (1989) noted that the empirical model proposed by Taraman (1974) accounted for the effect of the machine operating environment but failed to account for aging (or wear out) characteristics of the tool. To account for both aging and the characteristics of the machine operating environment, a proportional hazards model (PHM) was proposed to assess tool life. In specifying the PHM, a Weibull model was assumed for the baseline failure rate to incorporate aging of the tool, and the effect of machining environment, as specified by Taraman (1974), was used to modulate the baseline failure rate.

Because of the variation in individual characteristics of machine tools and the possible omission of relevant variables (describing the operational environment) from the model, a fully parametric PHM for tool life may not be adequate. This may result in suboptimal replacement strategies. Thus a more general model that relaxes some of these parametric assumptions may be more desirable for modeling machine tool life. As pointed out by Gelfand (1999), in a situation where the parametric assumptions may be too restrictive, a semiparametric model can be developed by a nonparametric specification of some portions of the model.

In this article we present a semiparametric analysis for machine tool life that treats the baseline failure rate function in a nonparametric manner while treating the effect of the covariates parametrically, as implied by Taraman (1974), using the PHM approach. A mixture of Dirichlet processes

(MDP) prior, proposed by Antoniak (1974), is assumed for the part of the model that defines the underlying failure distribution. Whereas the foregoing set-up leads to a complex analysis, inference is developed using the framework of Escobar and West (1995) and the efficient sampling algorithms of MacEachern (1994). The MDP approach reduces the restriction of the parametric baseline failure rate and allows assessment of differences between the failure characteristics of tools that cannot be described by the covariates in the model of Taraman (1974). An additional benefit of the MDP approach for the PHM is the ability to compare the parameters of the baseline failure rate corresponding to each individual machine tool, thus addressing the question of whether all necessary covariates are included in the model.

The development of optimal replacement strategies using a nonparametric form for the failure model has been considered from a sampling theory perspective by Frees and Ruppert (1985) and Aras and Whitaker (1991). Neither group considered the effect of covariate information, however. Neither a nonparametric nor semiparametric approach nor the use of covariates has been considered from a Bayesian point of view for developing optimal replacement strategies. The Bayesian decision theoretic framework for replacement policies for machine tools presented here represents an extension of the work of Mazzuchi and Soyer (1995, 1996) to include semiparametric models and covariate effects.

While building on existing analytical techniques proposed in the literature for semiparametric models, the analysis herein has four basic aims:

1. To develop a model describing the failure behavior of machine tools accounting for the operational environment effect as well as aging, thereby capturing a wide class of tool failure behavior

2. To illustrate the use of this model as a diagnostic tool for assessing model adequacy in describing the differences among the machine tools

3. To illustrate the use of this model in predicting the useful life of a machine tool operating in a particular operational environment

4. To illustrate the use of this model in developing optimal replacement strategies for single machine tools and groups of tools.

The PHM for machine tool life is discussed in Section 2, along with the relaxation of parametric assumptions. In Section 3 the semiparametric model is used to analyze the machine tool failure data given by Taraman (1974). The posterior distribution of the parameters of the model are compared to a parametric model, and the predictive ability of the two approaches is compared using both posterior predictive densities and the decision information criterion from Spiegelhalter, Best, Carlin, and van der Linde (2002). The model is used to develop Bayesian semiparametric replacement strategies in Section 4. Conclusions are presented in Section 5.

2. A BAYESIAN SEMIPARAMETRIC PROPORTIONAL HAZARDS MODEL

This section discusses the use of a PHM to analyze machine tool failures under varying operating conditions. Unlike the parametric analysis of Mazzuchi and Soyer (1989), we develop a semiparametric inference using an MDP approach, with inference furnished using the efficient algorithms of MacEachern (1994). This development is, therefore, a nontrivial combination of previously published techniques that allows a full analysis in this particular application. This section outlines the necessary methodology for this combination of techniques.

2.1 A Proportional Hazards Model for Machine Tool Failure

The PHM was proposed by Cox (1972) to incorporate covariate information in a survival or time to failure model. The model has been widely applied in survival and reliability analysis. The PHM is defined using the concept of the failure rate. Let T_i be the life length of machine tool i. Assuming that T_i is continuous, the failure rate function of the distribution of T_i is defined as

$$\lambda_i(t) = \frac{f_i(t)}{R_i(t)},$$

where $f_i(t)$ is the probability density function of T_i and

$$R_i(t) = P(T_i \geq t) = \exp\{-\Lambda_i(t)\}$$

is the reliability of machine tool i at time t, with $\Lambda_i(t) = \int_0^t \lambda_i(s)\, ds$ the cumulative failure rate.

Let Z_i be a vector of p measured covariates describing the operational environments of machine tool i. The covariates available for the machine tool analysis are constant with respect to time and are known before any failure data are observed. Cox (1972) proposed that the distribution of T_i could be made dependent on Z_i via the failure rate by assuming that the failure rate of the ith item is a product of a common base failure rate function and a function of the covariates, explicitly

$$\lambda_i(t; Z_i) = \lambda_0(t)e^{\beta^T Z_i},$$

where β is a vector of p regression parameters and $\lambda_0(t)$ is a baseline failure rate function. The parameters of the model are suppressed in the notation $\lambda_i(t; Z_i)$.

Often a parametric form is assumed for the baseline failure rate. This is equivalent to choosing a common family of distributions for the life lengths of the machine tools. Mazzuchi and Soyer (1989) performed an analysis of the Weibull parametric model for the machine tool failure data. Their analysis used integral approximation techniques to find the marginal posterior distributions of the model parameters. An alternative analysis of the parametric model was developed by Dellaportas and Smith (1993) using Markov chain Monte Carlo techniques, although an analysis of the machine tool failure data using these methods has not yet been published.

Approaches for modeling the baseline cumulative failure rate function include the gamma process proposed by Kalbfleisch (1978) and criticized by Hjort (1990), the extended gamma process presented by Laud, Damien, and Smith (1996), and the beta process presented by Hjort (1990), with a computational model developed by Laud et al. (1998). (For a full review of other semiparametric approaches to inference on regression models, see Gelfand 1999.) In the

next section we present an MDP prior, as defined by Antoniak (1974), for the baseline failure rate of the PHM to analyze the machine tool problem. This prior distribution allows a large family of continuous failure time distributions, thus relaxing the full parametric assumption. Our setup is similar to the semiparametric accelerated failure time model proposed by Kuo and Mallick (1997), but applied to the PHM.

2.2   A Mixture of Dirichlet Processes Prior for the Proportional Hazards Model

Under the MDP approach, the baseline failure rate is assumed to be some continuous function $\lambda_0(t; \theta_i)$, where θ_i is the vector of unknown parameters specific to the ith machine tool. Uncertainty about the θ_i's is described by specifying a prior distribution G. If the form of G is known but the hyperparameters are unknown, then this class of problems is referred to as hierarchical Bayes problems in the sense of Lindley and Smith (1972). If the form of G is unknown, then uncertainty about G must be modeled. One way to model this uncertainty is to follow the development of MacEachern (1994) and West, Muller, and Escobar (1994) and describe uncertainty about G by a Dirichlet process prior denoted by

$$G \sim DP(G_0, M),$$

where G_0 is the baseline prior and M is the strength of belief parameter. (See Ferguson 1973 for a discussion of Dirichlet process priors.)

By specifying a form for $\lambda_0(t_i; \theta_i)$ conditional on θ_i, we specify a conditional parametric model for T_i whose density $f(t_i|\theta_i, \beta, Z_i)$ is given by

$$\lambda_0(t_i; \theta_i)e^{\beta^T Z_i} \exp\{-\Lambda_0(t_i; \theta_i)e^{\beta^T Z_i}\}, \qquad (1)$$

where $\Lambda_0(t_i; \theta_i) = \int_0^{t_i} \lambda_0(s; \theta_i)\, ds$. Specification of the semiparametric PHM is completed by specifying a parametric prior for the covariate effects β, denoted by $\pi(\beta)$, which is independent of the θ_i's. The nonparametric nature of the model arises because the distribution of T_i, unconditional on θ_i, is an unknown mixture of $f(t_i|\theta_i, \beta, Z_i)$ given by

$$f(t_i|G, \beta, Z_i) = \int f(t_i|\theta_i, \beta, Z_i)\, dG(\theta_i),$$

where the distribution of T_i results from mixing with respect to G. These models were termed *Dirichlet process mixed models* by Mukhopadhyay and Gelfand (1997), because G is assumed a priori to be a Dirichlet process.

The semiparametric PHM using a an MDP approach can be summarized as

$$(T_i|\theta_i, \beta, Z_i) \sim f(t_i|\theta_i, \beta, Z_i),$$

$$(\theta_i|G) \sim G,$$

$$(G) \sim DP(G_0, M),$$

$$(\beta) \sim \pi(\beta). \qquad (2)$$

It is also assumed a priori that β and $\Theta = (\theta_1, \ldots, \theta_n)$ are independent of each other.

In addition to its flexibility and ability to capture individual characteristics of the machine tools, the proposed semiparametric PHM also provides an assessment of the completeness

of the set of covariates included in the analysis. In the classical literature, a residual analysis is performed to determine whether differences in the failure characteristics among the machine tools remain after the effect of the covariates has been removed. In the proposed model, differences between the individual machine tools can be assessed by differences between the distributions of the θ_i's.

2.3   Posterior Inference and Prediction

Given failure and covariate data $D = \{T_1 = t_1, \ldots, T_n = t_n, Z_1, \ldots, Z_n\}$ on n machine tools, the likelihood function of G and β given the data D is obtained as

$$L(G, \beta|D) = \prod_{i=1}^{n} \int f(t_i|\theta_i, \beta, Z_i)\, dG(\theta_i).$$

Given an arbitrary prior on β, say $\pi(\beta)$, which is independent of Θ and G, following Kuo and Mallick (1997), the posterior distribution of G given β and D can be obtained as a mixture of Dirichlet processes,

$$(G|\beta, D) \sim \int DP\left(MG_0 + \sum_{j=1}^{n} \delta_{\theta_j}\right) d\Pi(\Theta|\beta, D), \qquad (3)$$

where δ_{θ_j} denotes a point mass distribution concentrated at θ_j and $d\Pi(\Theta|\beta, D)$ is proportional to

$$\prod_{i=1}^{n} f(t_i|\theta_i, \beta, Z_i)\left[MG_0 + \sum_{j=1}^{i-1} \delta_{\theta_j}\right] d\theta_i.$$

It is difficult to sample from the distribution $(G|\beta, D)$ given in (3), because G is effectively an infinite-dimensional parameter (see, e.g., Kuo 1986).

Instead of attempting to perform inference on the mixing distribution G directly, one can perform simple inference using the Markov chain Monte Carlo methods in algorithm 1 of Escobar and West (1995) to obtain a sample from the posterior distribution of Θ and β given the data D. For our problem, the attractive feature of this approach is that computation of $\pi(\Theta, \beta|D)$ based on the Gibbs sampler can be achieved without sampling from the posterior distribution of $(G|\beta, D)$, thus reducing the problem to n dimensions.

Following the derivations of Escobar and West (1995), it can be shown that the full conditional for each θ_i is

$$(\theta_i|\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n, \beta, D)$$
$$\sim q_{i,0}G_b(\theta_i|t_i, \beta, Z_i) + \sum_{j\neq i} q_{i,j}\delta_{\theta_j}(\theta_i),$$

where $\delta_{\theta_j}(\theta_i)$ equals 1 if $\theta_i = \theta_j$ and 0 otherwise. The term $G_b(\theta_i|t_i, \beta, Z_i)$ is the baseline posterior distribution

$$dG_b(\theta_i|t_i, \beta, Z_i) \propto f(t_i|\theta_i, \beta, Z_i)\, dG_0(\theta_i).$$

The terms $q_{i,j}$, for $j \neq i$ represent the positive probability that some of the θ_i's will take the same values due to the discreteness of G (as a result of the Dirichlet process prior). These are given by

$$q_{i,0} \propto M \int f(t_i|\theta_i, \beta, Z_i)\, dG_0(\theta_i)$$

and

$$q_{i,j} \propto f(t_i|\theta_j, \beta, Z_i),$$

where $f(t_i|\theta_j, \beta, Z_i)$ is the density of T_i when $\theta_i = \theta_j$ and

$$q_{i,0} + \sum_{j \neq i} q_{i,j} = 1.$$

An algorithm proposed by MacEachern (1994) exploits this fact to increase the efficiency of the sampling processs by updating the θ_i's in groups or clusters. Using this algorithm, we may draw a sample from the conditional distribution of Θ given β and the data. The other distribution needed to implement the Gibbs sampler for this model is the conditional distribution of β given Θ and the data. Samples from this distribution can be obtained using the methods discussed by Dellaportas and Smith (1993) for the parametric model, because given Θ, a conditionally parametric model is specified by (1).

In implementing the Gibbs sampler, given a current Θ and β, the general steps are as follows:

1. Generate a new Θ conditional on β using the efficient MDP algorithm of MacEachern (1994).

2. Generate a new β conditional on Θ using the methods discussed by Dellaportas and Smith (1993).

An attractive feature of the MDP setup and the proposed algorithm is that posterior predictive densities and reliability functions can be easily evaluated once the posterior sample from Θ and β is available. For example, in predicting T_{n+1}, the posterior predictive density $f(t_{n+1}|D, Z_{n+1})$ is

$$\int_{\beta, \Theta} f(t_{n+1}|\Theta, \beta, Z_{n+1})\pi(\Theta, \beta|D)\, d\Theta\, d\beta,$$

where $f(t_{n+1}|\Theta, \beta, Z_{n+1})$ is

$$\int_{\theta_{n+1}} f(t_{n+1}|\theta_{n+1}, \beta, Z_{n+1})\pi(\theta_{n+1}|\Theta)\, d\theta_{n+1}$$

and $\pi(\theta_{n+1}|\Theta)$ is

$$\frac{M}{M+n}G_0(\theta_{n+1}) + \frac{1}{M+n}\sum_{i=1}^{n}\delta_{\theta_i}(\theta_{n+1}). \qquad (4)$$

Thus the posterior predictive density $f(t_{n+1}|D, Z_{n+1})$ can be written as

$$\int_{\beta, \Theta, \theta_{n+1}} f(t_{n+1}|\theta_{n+1}, \beta, Z_{n+1})\pi(\theta_{n+1}|\Theta)$$

$$\times \pi(\Theta, \beta|D)\, d\theta_{n+1}\, d\Theta\, d\beta, \qquad (5)$$

and using the posterior sample from $\pi(\Theta, \beta|D)$, denoted by $(\theta_{1,l}, \ldots \theta_{n,l}, \beta_l)$ for $l = 1, \ldots, S$, and draws from (4), denoted by $\theta_{n+1,l}$, $f(t_{n+1}|D, Z_{n+1})$ can be approximated as

$$\frac{1}{S}\sum_{l=1}^{S} f(t_{n+1}|\theta_{n+1,l}, \beta_l, Z_{n+1}). \qquad (6)$$

Because the results under the MDP setup may be sensitive to the choice of M, we use a further extension by incorporating M into the Gibbs sampler. Our approach follows that of Escobar and West (1995), assuming a priori that M follows an arbitrary prior $\pi(M)$. In their development, Escobar and West defined K as the number of unique values of $\theta_1, \ldots \theta_n$, also referred to as the number of cliques by MacEachern (1998). Conditioned on K, the full conditional distribution of

M is independent of all other parameters with density proportional to

$$M^{K-1}(M+n)B(M+1, n)\pi(M),$$

where $B(\cdot)$ is the standard beta function. Escobar and West (1995) offered a simple two-step process for sampling from this distribution if $\pi(M)$ is assumed to be a gamma distribution. The approach involves using a data-augmentation step at each iteration of the Gibbs sampler. Thus K is also recorded in the Gibbs sample, because the distribution of the number of cliques is of interest in the reliability analysis in Section 3. We note that at each iteration of the Gibbs sampler, once M is drawn, the rest of the quantities are sampled as discussed in steps 1 and 2.

3. ANALYSIS OF MACHINE TOOL FAILURE DATA USING PARAMETRIC AND SEMIPARAMETRIC INFERENCE

The data used in this analysis, given in Table 1, were first presented by Taraman (1974). They consist of the failure times of 24 machine tools and their corresponding cutting speed, feed rate, and depth of cut. Each experimental run used a workpiece material of SAE 1018 cold-rolled steel, 4 inches in diameter and 2 feet long. The 24 machine tools used for the cutting were tungsten carbide disposable inserts mounted in a tool holder. A 7.5-horsepower engine lathe equipped with a three-jaw universal chuck and a live center mounted in the tailstock was used to perform the cutting operation. The cutting operations were performed without using cutting fluids.

In this section we use both parametric and semiparametric inference methods to assess the effect of the covariates on the useful life of these machine tools. The parametric analysis mirrors the development of Mazzuchi and Soyer (1989) by specifying $\lambda_i(t; z_i) = \lambda_0(t)\exp\{\beta_1 \ln Z_{i,1} + \beta_2 \ln Z_{i,2} + \beta_3 \ln Z_{i,3}\}$,

Table 1. The Machine Tool Failure Data

Machine tool	Speed (fpm)	Feed (ipr)	Depth of cut (inches)	Tool life (min.)
1	340	.00630	.02100	70.0
2	570	.00630	.02100	29.0
3	340	.01410	.02100	60.0
4	570	.01416	.02100	28.0
5	340	.00630	.02100	64.0
6	570	.00630	.04000	32.0
7	340	.01416	.04000	44.0
8	570	.01416	.04000	24.0
9	440	.00905	.02900	35.0
10	440	.00905	.02900	31.0
11	440	.00905	.02900	38.0
12	440	.00905	.02900	35.0
13	305	.00905	.02900	52.0
14	635	.00905	.02900	23.0
15	440	.00472	.02900	40.0
16	440	.01732	.02900	28.0
17	440	.00905	.01350	46.0
18	440	.00905	.04550	33.0
19	305	.00905	.02900	46.0
20	635	.00905	.02900	27.0
21	440	.00472	.02900	37.0
22	440	.01732	.02900	34.0
23	440	.00905	.01350	41.0
24	440	.00905	.04550	28.0

where $Z_{i,1}$ is cutting speed, $Z_{i,2}$ is feed rate, $Z_{i,3}$ is depth of cut, and $\lambda_0(t) = \alpha\gamma t^{\gamma-1}$. For computational efficiency, we use the techniques developed by Dellaportas and Smith (1993) for inference in the parametric model. The semiparametric inference follows the methods developed in Section 2. The two approaches are compared in several ways, using the posterior distributions of the model parameters, posterior distributions of the predicted reliabilities, and posterior predictive densities, as discussed by Gelfand (1996), and the deviance information criteria (DIC), as described by Spiegelhalter et al. (2002).

3.1 Comparison of the Posterior Distributions of the Model Parameters

For comparison of the semiparametric inference method to the parametric method proposed by Mazzuchi and Soyer (1989), the conditional baseline failure rate in (2) is assumed to be a Weibull density with scale parameter α_i and shape parameter γ. Thus, under the notation of Section 2, $\theta_i = (\alpha_i, \gamma)$, with $n = 24$. Note that each item is assumed to wear at the same rate; thus γ is common to each item, whereas the scale parameter α_i is allowed to vary from one machine tool to the next. The conditional parametric density in (2) is then $f(t_i|\alpha_i, \gamma, \beta, Z_i^*) = \alpha_i\gamma t_i^{\gamma-1} \exp\{\beta^T Z_i^*\} \exp\{-\alpha_i t_i^{\gamma} \exp\{\beta^T Z_i^*\}\}$, where $Z_i^* = (\ln Z_{i,1}, \ln Z_{i,2}, \ln Z_{i,3})^T$. For the prior best guess of the mixing distribution of the α_i's, G_0 in (2), a gamma distribution is chosen.

The prior assumptions of Mazzuchi and Soyer (1989) had low variances and seemingly specific values for the means of each parameter. However, no motivation was given for these prior assumptions. Thus our prior distributions are noninformative with large variances assumed on each parameter. A priori, γ, β_1, β_2, and β_3 are independent of each other and $\alpha_1, \ldots, \alpha_n$. A normal prior, with mean 0 and variance 20, was assumed for each of the covariate effect parameters, β_1, β_2, and β_3, where the covariate values had been scaled so that they were of the same order of magnitude. This reflects our lack of knowledge of whether the covariates would increase or decrease failure time. The prior distribution of the shape parameter, γ, was assumed to be a normal distribution truncated at 0 with mean 1 and variance 10 reflecting no strong prior belief concerning the failure rate behavior (whether it is increasing or decreasing). The best-guess prior distribution G_0 for the scale parameters was assumed to be a gamma distribution with mean 1 and variance 10. A priori, M is assumed to follow an uninformative gamma distribution with mean 24 and standard deviation 100.

A single-chain Gibbs sampler was run to obtain 2,500 samples with a warmup of 5,000 and a lag of 25 between successive samples. Boxplots of the marginal posterior distributions of the natural log of the scale parameters for the tools in the data obtained under the semiparametric model are shown in Figure 1. A second Gibbs sampler was run for the parametric model using the methods of Dellaportas and Smith (1993). The same prior distributions were assumed, except that the prior distribution of the scale parameter was assumed to be the best-guess prior, G_0, in the semiparametric model. The boxplot of the distribution obtained under the parametric model is also shown in Figure 1, denoted by a P on the x-axis.
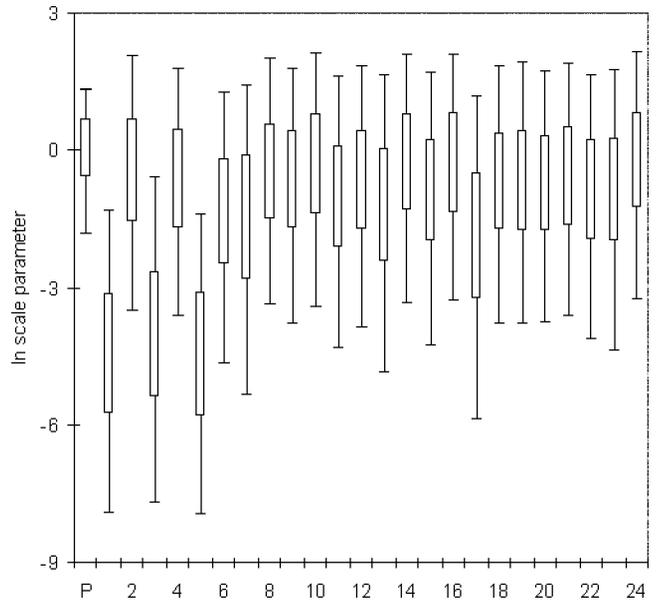
Figure 1. Boxplots of the Marginal Posterior Distributions of the Log Scale Parameters of the Machine Tools Under the Parametric and Semiparametric Models.

Figure 1 shows marked differences between the posterior distributions of the scale parameters. The posterior distribution of the scale parameter under the parametric model has a smaller variance and obviously cannot represent the variability among the tools demonstrated by the individual scale parameters under the MDP setup. In the semiparametric model, the individual scale parameters express the differences between the machine tools that are not explained by the covariates. We can examine such differences among the machine tools through the marginal posterior distribution of K, the number of groups of α_i's in the MDP setup, shown in Figure 2. Figure 2 indicates that there is little support for 1 group, as in
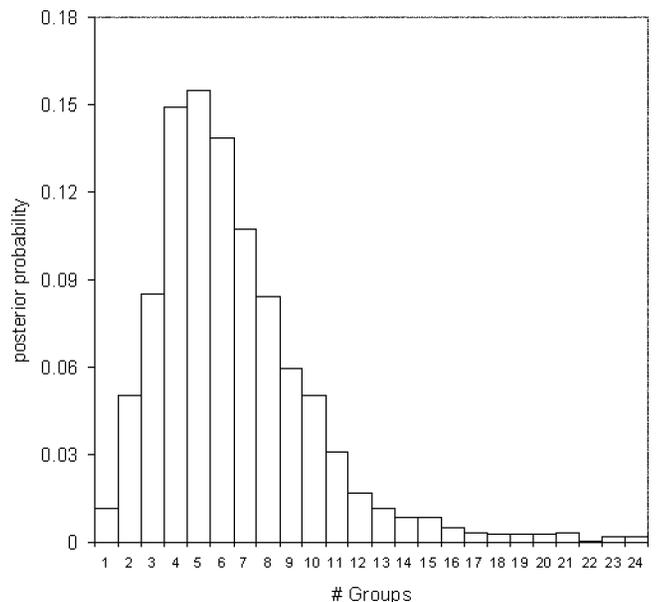

Figure 2. Marginal Posterior Distribution of the Number of Groups of α_i's in the Semiparametric Model.

the simple parametric model, or for 24 groups, a parametric model with overdispersion. In fact, there is a .78 posterior probability that there are between 3 and 9 groups in the data, with the expected number of groups estimated at 6.6. More formally, the Bayes factor can be used to test the appropriateness of the two models. This can be achieved using the harmonic mean approximation of Newton and Raftery (1994). Using this approximation, the Bayes factor in favor of the semiparametric model is $10^{7.61}$. This is decisive evidence in the scale of Jeffreys (1961). (See Kass and Raftery 1995 for further discussion of Bayes factors and their approximation.)

Comparison of the posterior distributions of the covariate effect parameters, β_1, β_2, and β_3, and the shape parameter, γ, is shown in Figures 3–6 graphically and given in Table 2 numerically. In each figure, the dotted line represents the posterior density obtained from the parametric model, and the solid line represents the density from the semiparametric model. In the table, the means of each marginal posterior distribution are given, with standard deviations in brackets. The posterior distribution of the four parameters are clearly different under the two models. Bayes factors in favor of the hypothesis $\beta_k \neq 0$, $k = 1, 2, 3$, under the parametric and semiparametric models are given in Table 3. Under the parametric model, only cutting velocity has a pronounced effect. Evidence in favor of nonzero effects is inconclusive for feed rate and depth of cut (Jeffreys 1961). However, under the semiparametric model, using Jeffreys's scale there is still decisive evidence that cutting velocity has an effect, but there is also substantial evidence that feed rate and depth of cut affect the lifetime of machine tools. The evidence of positive values for the covariate effect parameters implies a decrease in the lifetime of the tool with increased cutting velocity, feed rate, and depth of cut. In both models, the mass of the posterior distribution of the shape parameter is in the region $\gamma > 1$, indicating aging, with a more pronounced aging effect indicated under the semiparametric model.
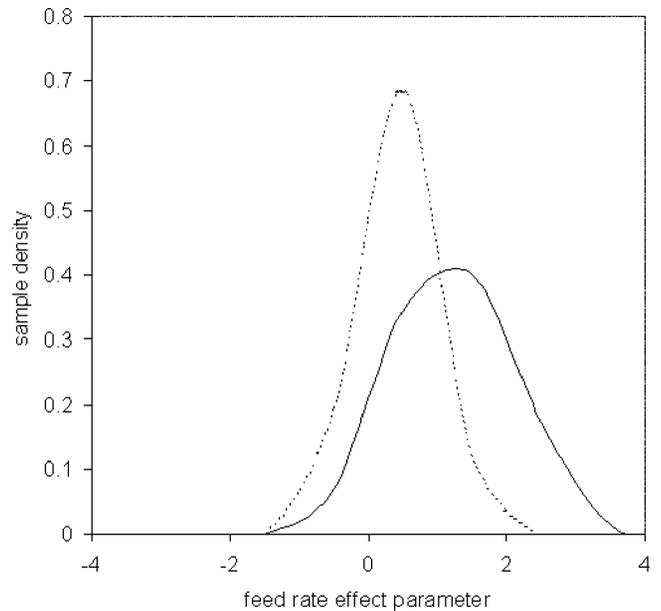

Figure 4. Estimated Posterior Densities of β_2 for the Parametric ($\cdots$) and Semiparametric (—) Models.

Thus the inference under the semiparametric model differs markedly from that obtained using a parametric model. Because the semiparametric model allows for differences among the machine tools not explained by the three covariates, the effect of the covariates included in the analysis can be estimated without these differences affecting the analysis. However, this does not necessarily imply that the predictive ability will be superior.

3.2 Comparison of the Predicted Reliability

The distribution of the predictive reliability of a given machine tool can be found using the sample approximations given in (6). As discussed in Section 2, calculating the
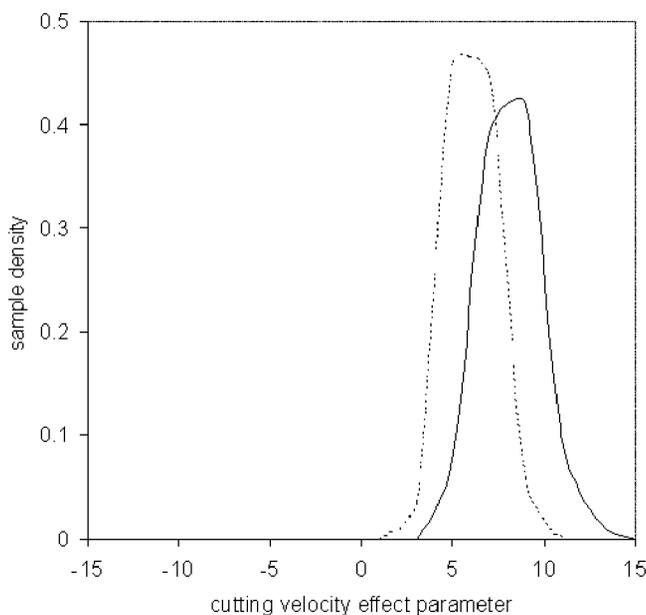

Figure 3. Estimated Posterior Densities of β_1 for the Parametric ($\cdots$) and Semiparametric (—) Models.
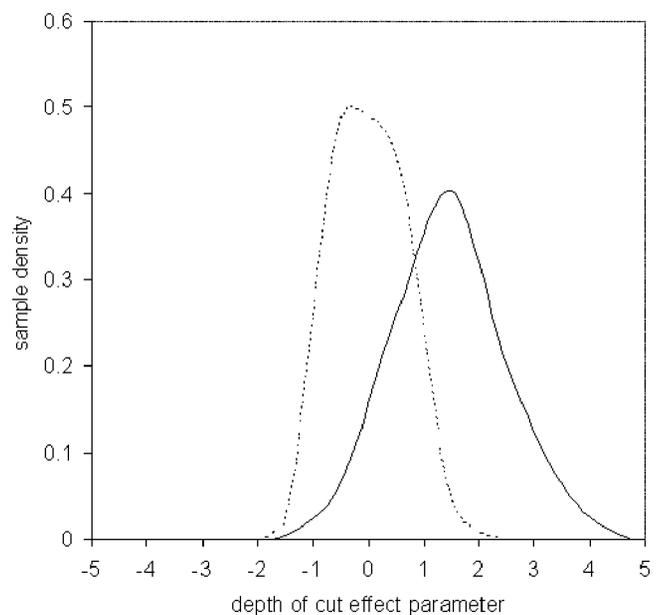

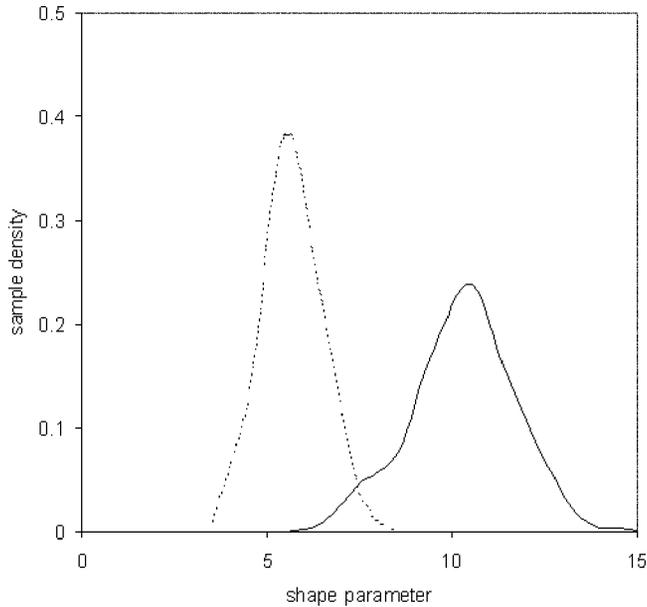Figure 5. Estimated Posterior Densities of β_3 for the Parametric ($\cdots$) and Semiparametric (—) Models.

Figure 6. Estimated Posterior Densities of γ for the Parametric ($\cdots$) and Semiparametric (—) Models.

Table 3. Bayes Factors in Favor of Nonzero Covariate Effects, That is, $\beta_k \neq 0$, Under the Parametric and Semiparametric (MDP) Models

Hypothesis	Parametric	MDP
$\beta_1 \neq 0$	$10^{6.75}$	$10^{3.35}$
$\beta_2 \neq 0$	10^{18}	$10^{.67}$
$\beta_3 \neq 0$	10^{15}	$10^{.58}$

distribution under the semiparametric approach requires that a new scale parameter, denoted by $\alpha_{n+1,l}$, be sampled for the new tool from the posterior distribution given the data from n tools. Given this sample, a sample from the posterior predictive reliability at time t for a tool with covariate values Z_{n+1}^* is

$$\exp\{-\alpha_{n+1,l}t^{\gamma_l}e^{\beta_l^T Z_{n+1}^*}\},$$

where $(\alpha_{1,l}, \ldots, \alpha_{n,l}, \gamma_l, \beta_l, M_l)$, $l = 1, \ldots, S$, are samples from the posterior distribution obtained using the semiparametric model and the dataset D. The distribution of the scale parameter for a new tool, $\alpha_{n+1,l}$, obtained from the posterior distribution of G given the data D, is multinomial, (4), with probability $\frac{M_l}{M_l+n}$, the draw is from the best-guess prior distribution G_0; with probability $\frac{1}{M_l+n}$ for each parameter α_i ($i = 1, \ldots, n$), the draw equals the sample from that parameter, $\alpha_{i,l}$, at the lth draw.

The posterior predictive distribution of the reliability of a given machine tool at defined mission times are shown as boxplots in Figure 7 for the parametric model and in Figure 8 for the semiparametric model. The new machine tool is predicted to operate with a cutting velocity of 440 fpm, a feed rate of .00472 ipr, and a .029-inch depth of cut. The variance in the predicted mission time reliabilities is greater under the semiparametric model. Figure 1 shows that in most cases, the

variance in the posterior distributions of the scale parameters under the semiparametric model is higher than that exhibited by the single scale parameter of the parametric model; this is also true for each of the other parameters of the model. Thus the variance of the mission time reliabilities is higher under the semiparametric model.

Figure 9 shows the medians of the posterior predictive distributions of the reliabilities of the same machine tool at different mission times, with the dotted line indicating the parametric model and the solid line indicating the semiparametric model. The medians are used due to the skewness of the posterior predictive distributions. It can be seen that despite the higher variability in the posterior distributions, the semiparametric model predicts a lower variability in machine tool reliability than the parametric model. The question now becomes which of these two predictions is better.

3.3 Comparison of the Predictive Ability

The parametric and semiparametric models can be compared using posterior predictive densities, as discussed by Gelfand (1996), and the DIC, as described by Spiegelhalter et al. (2002). To perform a comparison of the posterior predictive densities, we split our set of observed failure times, D, into two sets, one (D_1) for updating and the other (D_2) for prediction. The posterior predictive densities of the second

Table 2. The Posterior Means (standard deviations) of the Covariate Effect and Shape Parameters for the Parametric and Semiparametric (MDP) Models

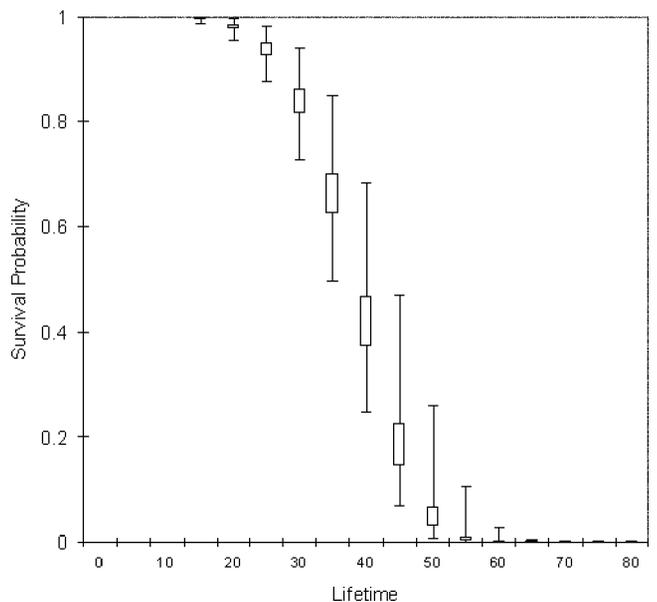|	Parametric	MDP
β_1	6.14 (1.16)	6.85 (3.60)
β_2	.41 (.48)	1.01 (.87)
β_3	−.01 (.56)	1.18 (1.05)
γ	5.72 (.78)	8.60 (4.30)

Figure 7. Boxplots of the Posterior Distribution of the Survival Probability at Fixed Lifetimes Under the Parametric Model.
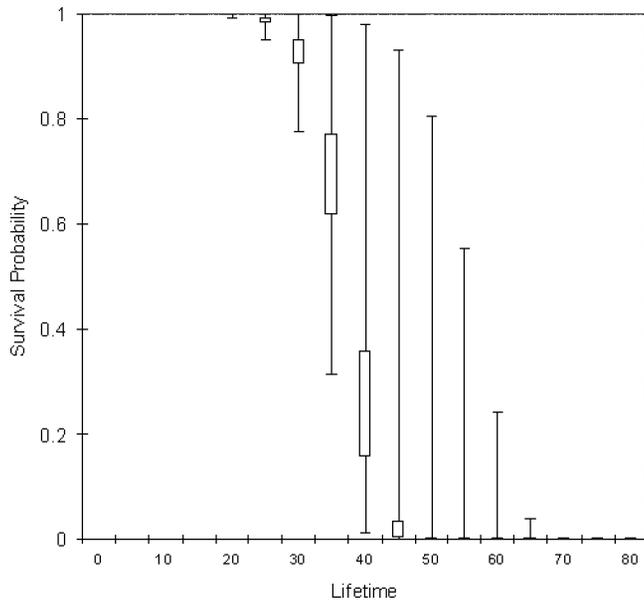
Figure 8. Boxplots of the Posterior Distribution of the Survival Probability at Fixed Lifetimes Under the Semiparametric Model.

dataset are obtained given the first and compared for the two approaches. The posterior distributions of the parameters of the two models, denoted by Φ_1 and Φ_2, are obtained given the dataset D_1. The posterior predictive density of the data in D_2 is then computed under each model. The posterior predictive density of D_2 given D_1 are compared under each model, given by

$$P^i(D_2|D_1) = \int P(D_2|\Phi_i)P(\Phi_i|D_1)\,d\Phi_i,$$

for $i = 1, 2$.

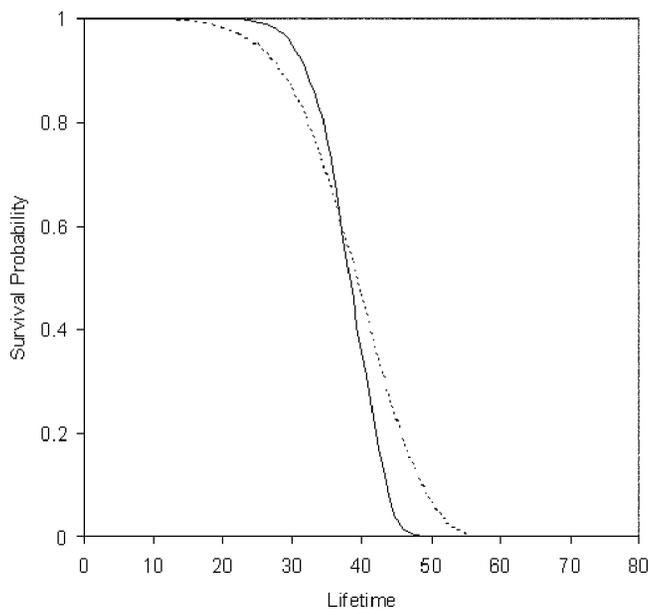The posterior analyses given in the foregoing examples for the parametric and semiparametric models were repeated including only 20 of the machine tools from the complete set of 24. A comparison of the log of the posterior predictive density functions for the life lengths of the remaining four tools allows a comparison of the predictive ability of each model. Once posterior inference has been performed using the Gibbs sampler, predictive inference can be performed using the posterior sample. The posterior predictive density function, given in (5), can then be calculated at the values of the failure times in the dataset D_2.

We note that the posterior predictive density comparison is dependent on how we partition the failure data. A better indication of the predictive ability of the two models can be obtained by considering all such partitions, but this is computationally infeasible. Thus in our analysis 100 partitions were selected at random, and the posterior predictive densities were calculated under each model. Figure 10 shows boxplots of the posterior predictive densites obtained for the 100 random partitions under the parametric and semiparametric (labeled MDP) models. It can be seen that overall, the semiparametric model gives better predictions.

The second comparison uses the DIC proposed by Spiegelhalter et al. (2002). The DIC supplies an estimate for the effective number of parameters, which is of particular interest in the MDP model. Let the set of available data be denoted by D. The Bayesian deviance is defined as

$$DV(\Phi) = -2\ln p(D|\Phi) + 2\ln f(D),$$

where $f(D)$ is some fully specified standardizing term that is a function of the data alone and thus does not affect model comparison. The model "fit" is then represented by $\overline{DV} = E_{\Phi|D}[DV]$, and the effective number of parameters is represented by $p_{DV} = E_{\Phi|D}[DV] - DV(E_{\Phi|D}[\Phi]) = \overline{DV} - DV(\overline{\Phi})$. The DIC is then defined as the sum

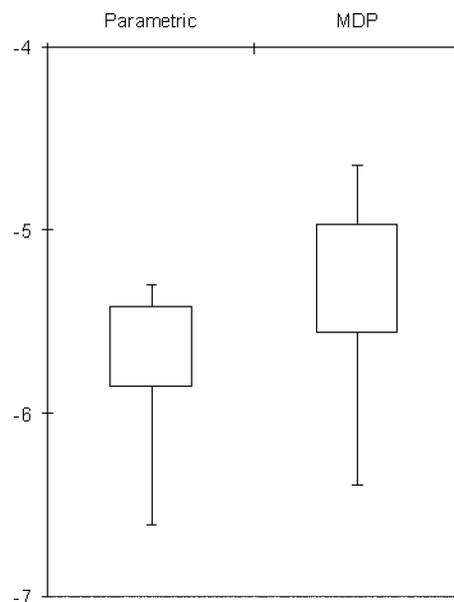$$DIC = \overline{DV} + p_{DV},$$

Figure 9. Medians of the Posterior Distribution of the Survival Probability at Fixed Times for the Parametric (· · ·) and Semiparametric (—) Models.

Figure 10. Boxplots of the Posterior Predictive Densities for 100 Randomly Selected Partitions for the Parametric and Semiparametric (MDP) Models.

and thus is a measure of model "fit" penalized by a measure of the effective number of parameters.

The model fit indicated by $\overline{DV}$ is 158.84 for the parametric model and 132.44 for the semiparametric model, results strongly in favor of the semiparametric approach. The semiparametric model is penalized in the final DIC for the higher number of parameters indicated by p_{DV}, 11.76 versus 4.41 for the parametric model. Thus the DIC is 163.25 for the parametric model and 144.20 for the semiparametric model. This again strongly favors the semiparametric model. Note that the effective number of parameters for the parametric model is very close to the true number of parameters, five. However, for the semiparametric model, the effective number of parameters is about 12, although the model includes 3 covariate effect parameters, a shape parameter, and potentially 24 scale parameters. The estimated value of p_{DV} would indicate eight groups of similar machine tools in the data for which the scale parameters are clustered. This result is consistent with the posterior distribution of the number of groups in Section 3.1.

In conclusion, each predictive comparison criterion shows strong evidence of the superior predictive ability of the semiparametric model over the parametric model.

4. SEMIPARAMETRIC OPTIMAL REPLACEMENT STRATEGIES FOR MACHINE TOOLS

As in the original article by Taraman (1974) and the later work by Balakrishnan and DeVries (1985), the aim of machine tool life modeling is to aid decisions concerning the operation of machine tools. An important decision in the operation of machine tools is when to replace them. In this section we demonstrate how the MDP setup can lead to markedly different recommendations when compared with the parametric model.

Two of the most commonly used replacement strategies for systems/items subject to aging/wear are age and block replacement (see Cox 1962 for an introduction). Under the age-replacement protocol, a planned replacement is made at age t_A if the item survives until then, or an in-service replacement is made whenever the item fails. This protocol can be applied to individual machine tools. Under the block-replacement protocol, a group of units are replaced at time points $t_B, 2t_B, \ldots$, irrespective of their ages, and an in-service replacement or repair is made whenever failures occur. The machine tools are nonrepairable and thus operate under good-as-new (GN) replacement. Under the GN assumption, a failed machine tool is replaced with a new one. This section offers a semiparametric approach to optimal replacement decisions that also incorporates the effect of covariate information. We first address age replacement, then block replacement.

4.1 Replacing Individual Machine Tools

As shown by Mazzuchi and Soyer (1995), under the age-replacement protocol, the cost per unit time for the ith machine tool, with life length T_i, is given by

$$C(t_A, T_i) = \frac{c_F}{T_i} I(T_i < t_A) + \frac{c_P}{t_A} I(T_i \geq t_A),$$

where c_P is the cost of a planned replacement, c_F is the cost of an in-service replacement such that $c_F > c_P$, and I is the indicator function that takes the value 1 (0) if its argument is true (false). The optimal strategy t_A^* is then determined by minimizing $E[C(t_A, T_i)]$, the expected cost per unit time, when a failure model is specified for T_i.

Conditional on the parameters β and Θ and the covariate vector Z_i, t_A^* is chosen such that

$$E[C(t_A, T_i)|\beta, \Theta, Z_i] = \left\{ \int_0^{t_A} \frac{c_F}{t_i} f(t_i|\beta, \Theta, Z_i)\, dt_i \right.$$
$$\left. + \int_{t_A}^{\infty} \frac{c_P}{t_A} f(t_i|\beta, \Theta, Z_i)\, dt_i \right\} \quad (7)$$

is minimized. The first term represents the cost per unit time of in-service failures, and the second term represents the cost per unit time of planned replacements. We note that given Θ, T_i does not depend on M in the semiparametric model.

A Bayesian optimal age-replacement interval is determined by minimizing the expectation of (7) with respect to the posterior distribution of distribution of β and Θ; that is, t_A^* is obtained by minimizing

$$E[C(t_A, T_i)] = \int E[C(t_A, T_i)|\beta, \Theta, Z_i]\pi(\beta, \Theta|D)\, d\beta\, d\Theta,$$

where D represents the observed failure and covariate data. The optimal age-replacement interval can also be determined before any testing of the tools, and in that case $D \equiv D_0$, the prior information. Following the semiparametric setup, the failure model for the ith machine tool conditional on θ_i is described by a PHM with density function

$$f(t_i|\theta_i, \beta, Z_i) = \lambda_0(t_i; \theta_i)e^{\beta^T Z_i} \exp\left\{ -\int_0^{t_i} \lambda_0(s; \theta_i)e^{\beta^T Z_i}\, ds \right\}.$$

Given test data D, as discussed in Section 2, the Gibbs sampler generates a posterior sample from the posterior density $\pi(\Theta, \beta, M|D)$. Let $(\theta_{1,l}, \ldots, \theta_{n,l}, \beta_l, M_l)$ for $l = 1, \ldots, S$ denote the sample; then $E[C(t_A, T_i)]$ can be approximated by

$$\frac{1}{S}\sum_{l=1}^{S} \left[\int_0^{t_A} \frac{c_F}{t_i} f(t_i|\theta_{i,l}, \beta_l, Z_i)\, dt_i + \frac{c_P}{t_A} R(t_A|\theta_{i,l}, \beta_l, Z_i) \right], \quad (8)$$

and the optimal interval is approximated by minimizing (8) with respect to t_A.

To demonstrate this approach, assume that the cost of a planned replacement is $100 and the cost of an in-service replacement is $200. Consider the first machine tool in Table 1, with a cutting speed of 340 fpm, a feed rate of .00630 ipr, and a depth of cut of .021 inch. The posterior distribution of the scale parameter for this tool was significantly shifted to lower values compared with that obtained under the parametric model and the scale parameters of the other tools. This shows that some individual variation not explained by the three covariates can result in higher predicted machine tool reliability under the semiparametric model.

The costs per unit time for various age-replacement periods under the parametric model and semiparametric approach are shown in Figure 11. The optimal age-replacement interval under the parametric and semiparametric models is 32 and 42, respectively, reflecting the difference between the two models. It can also be seen in Figure 11 that the optimal expected cost
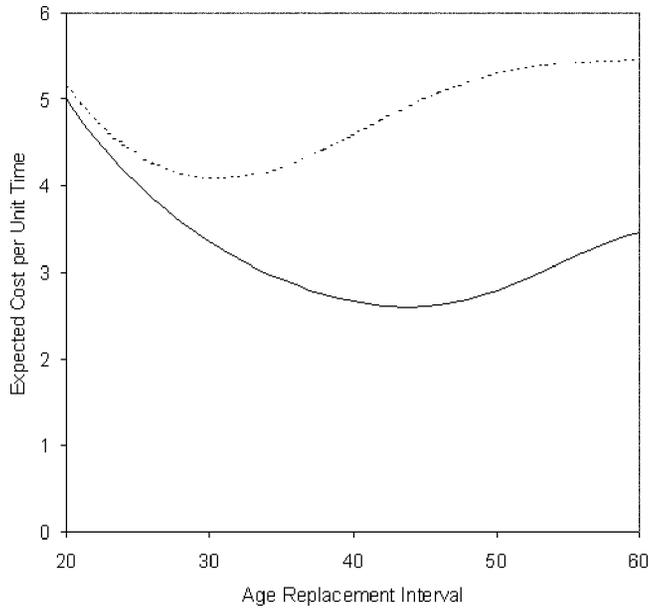
Figure 11. Expected Cost per Unit Time for Various Age-Replacement Periods for the Parametric (···) and Semiparametric (—) Models.

per unit time of the age-replacement protocol is predicted to be lower under the semiparametric model compared with the parametric model, again because the machine tool is predicted to be more reliable. However, for other covariate combinations this situation may be reversed. For this dataset, the semiparametric model had superior posterior predictive densities and DIC results (see Sec. 3.3; thus the decision maker should trust its optimal maintenance predictions over the parametric model.)

4.2 Replacing a Group of Machine Tools

Whereas the age-replacement policies are adequate for maintenance of a single machine tool, when several machine tools are to be maintained it may be more cost-effective to replace them as a group (block) rather than individually. Under the block-replacement protocol, Mazzuchi and Soyer (1996) calculated the cost per unit time for the ith machine tool to be

$$C(t_B, N_i(t_B)) = \frac{c_P + c_F N_i(t_B)}{t_B},$$

where $N_i(t)$ represents the number of in-service failures for the ith unit that occur in an interval of length t_B, c_P is the cost of a planned replacement, and c_F is the cost of an in-service replacement or repair. Assuming that m units will be replaced at time t_B, the total cost per unit time is given by

$$C(t_B, N(t_B)) = \sum_{i=1}^{m} C(t_B, N_i(t_B)),$$

where $N(t_B) = (N_1(t_B), \ldots, N_m(t_B))$.

Similar to the age-replacement protocol, the optimal block-replacement interval, t_B^*, is determined by minimizing $E[C(t_B, N(t_B))]$ when models are specified for the $N_i(t_B)$'s. Under the GN assumption, $N_i(t_B)$ is described by a renewal process, with the failure time distribution of the ith machine tool denoted

by $F_i(t_i|\beta, \Theta, Z_i) = P(T_i \le t_i|\beta, \Theta, Z_i)$. The expected cost per unit time, conditional on β and Θ, is given by

$$E[C(t_B, N(t_B))|\beta, \Theta, Z] = \frac{mc_P + c_F \sum_{i=1}^{m} E[N_i(t_B)|\beta, \Theta, Z_i]}{t_B},$$

where $E[N_i(t_B)|\beta, \Theta, Z_i]$ is the renewal function for $F_i(t_i|\beta, \Theta, Z_i)$ under the GN assumption. Thus a Bayesian-optimal block-replacement interval is determined by minimizing

$$E[C(t_B, N(t_B))]$$
$$= \int E[C(t_B, N(t_B))|\beta, \Theta, Z]\pi(\beta, \Theta|D)\, d\beta\, d\Theta \quad (9)$$

with respect to t_B, where $Z = (Z_1, \ldots, Z_m)$. As in Section 4.1, given Θ, the distribution of T_i is independent of M.

Under the semiparametric framework of Section 2, the failure model for the ith renewal process conditional on θ_i is described by a PHM with density function $f_i(t_i|\theta_i, \beta, Z_i)$. Determination of the optimal block-replacement interval t_B^* requires the evaluation of $E[C(t_B, N(t_B))]$ given by (9), which involves the conditional renewal functions $M_i(t_B|\beta, \Theta, Z_i) = E[N_i(t_B)|\beta, \Theta, Z_i]$. However, each renewal process is assumed to have a new base parameter θ_i' sampled from the posterior predictive distribution in (4), as discussed in Section 2.2. If a posterior sample $(\theta_{i,l}', \beta_l, M_l)$, for $l = 1, \ldots, S$ is available from the distribution $\pi(\theta_i', \beta, M|D)$, then a Bayesian estimate for the renewal function can be obtained via

$$M_i(t_B|D, Z_i) \approx \frac{1}{S}\sum_{l=1}^{S} M_i(t_B|\theta_{i,l}', \beta_l, Z_i),$$

if $M_i(t_B|\theta_{i,l}', \beta_l, Z_i)$ can be evaluated. For most well-known parametric forms, the renewal function is not available analytically. One way to evaluate the renewal function is to use an approximation such as the one offered by Smeitnik and Dekker (1990). An alternative approach is to simulate the renewal process and approximate the conditional renewal function $M_i(t_B|\theta_i', \beta, Z_i)$ (see, e.g., Ross 1989). Note that given θ_i', β, and Z_i, for the given form of F_i, the renewal process $N_i(t_B|\theta_i', \beta, Z_i)$ can be easily simulated by generating failure times $T_i^1, T_i^2, \ldots$ until their sum exceeds t_B. A sample from $N_i(t_B|\theta_i', \beta, Z_i)$ is then

$$N_i^{(j)} = \sum_k I(T_i^k \le t_B),$$

where I is the indicator function. For given values of θ_i', β, and Z_i, multiple samples, say K, are generated from $N_i(t_B|\theta_i', \beta, Z_i)$, and the conditional renewal function can be approximated as

$$\widehat{M}_i(t_B|\theta_i', \beta, Z_i) = \frac{1}{K}\sum_{j=1}^{K} N_i^{(j)}. \quad (10)$$

Thus to evaluate (9) for each sample point $(\theta_{i,l}', \beta_l)$ from $\pi(\theta_i', \beta|D)$, K realizations of the renewal process $N_i(t_B|\theta_{i,l}', \beta_l, Z_i)$ are generated, and the conditional renewal function (10) is approximated. This is repeated for all m machine tools. Then $E[C(t_B, N(t_B)]$ can be approximated as

$$\frac{1}{S}\sum_{l=1}^{S} \frac{mc_P + c_F \sum_{i=1}^{m} \widehat{M}_i(t_B|\theta_{i,l}', \beta_l, Z_i)}{t_B}, \quad (11)$$

and the optimal block-replacement interval t_B^* is estimated by minimizing (11) with respect to t_B.

To demonstrate the semiparametric approach for determining the optimal block-replacement period for a group of machine tools, the cost of a planned replacement is assumed to be \$100 and the cost of an in-service replacement is \$200. The posterior samples obtained in Section 3 are assumed to represent our posterior beliefs about the parameters of the lifetime distribution. The covariate values of the first four machine tools listed in Table 1 are chosen to illustrate block replacement. The first and third tools, shown in Figure 1, have lower values for their scale parameters than the second and fourth tools. The optimal value is found to be 24 under the parametric model and 23 under the semiparametric model. This difference between the optimal intervals predicted by the two approaches is much less than that obtained for the age-replacement intervals of individual tools, because the differences in the reliabilities of the tools average out in the group. Thus in this case there is no difference between the policies under the two models.

Consider the first, third, and fifth machine tools as a block, however. In this case, the three tools have significantly lower scale parameter values than the other tools (see Fig. 1). The cost per unit time for various block-replacement periods for each approach is shown in Figure 12. The predicted optimal replacement intervals are now 27 for the parametric model and 33 for the semiparametric model—a larger difference, with the expected cost again lower for the semiparametric model due to the lower scale parameter values. Thus the semiparametric model will yield different decisions in cases where differences are picked up by the scale parameters for each tool, and again we favor these because of the results in Section 3.3.
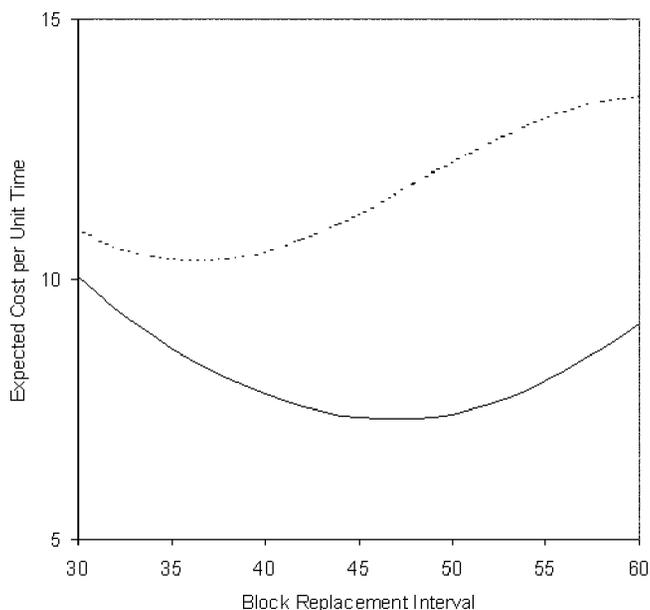
5. CONCLUSIONS

In this article we have presented a semiparametric model developed for the analysis of machine tool failure data. The analysis was motivated by the important economic and product quality implications surrounding replacement strategies for such tools and the lack of a uniformly accepted model for machine tool life prediction. We demonstrated the flexibility of the semiparametric model using actual machine tool failure data from Taraman (1974). Analysis of the data illustrated distinct differences in the findings obtained using the semiparametric and parametric models. Specifically, the semiparametric model illustrated the inadequacy of the simple set of covariates in describing the failure characteristics of the machine tools and proved to be superior in predictive capability.

We also compared the two models in the development of age- and block-replacement strategies for machine tools. The difference in the two models was apparent, with the semiparametric results favored because of its superior predictive ability. The use of semiparametric models and covariate information in developing Bayesian replacement strategies represents a novel addition to the literature.

The PHM was chosen for this study following Taraman (1974), Balakrishnan and DeVries (1985), and Mazzuchi and Soyer (1989). An interesting future study would combine this covariate effect model with the accelerated failure time model (Kuo and Mallick 1997), examining different covariate effects on machine tool reliability. Another important area of study is the use of these models in real time analysis and prediction of machine tool failure using secondary variables such as acoustic emission and vibration. Such an analysis would require the use of time-dependent covariates or dynamic environment modeling techniques different from those used here.

Figure 12. Expected Cost per Unit Time for Various Block-Replacement Periods for the Parametric (· · ·) and Semiparametric (—) Models.

REFERENCES

Antoniak, C. E. (1974), "Mixtures of Dirichlet Processes With Applications to Bayesian Non-Parametric Problems," *The Annals of Statistics*, 2, 1152–1174.

Aras, G., and Whitaker, R. L. (1991), "Sequential Non-Parametric Estimation of an Optimal Age Replacement Policy," Technical Report 158, University of California, Santa Barbara.

Balakrishnan, P., and DeVries, M. F. (1985), "Sequential Estimation of Machinability Parameters for Adaptive Optimatization of Machinability Data Base Systems," *Journal of Engineering for Industry*, 27, 159–166.

Cox, D. R. (1962), *Renewal Theory*, London: Methuen.

—— (1972), "Regression Models and Life Tables," *Journal of the Royal Statistical Society*, Ser. B, 34, 187–220.

Dellaportas, P., and Smith, A. F. M. (1993), "Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling," *Applied Statistics*, 42, 443–459.

Escobar, D. M., and West, M. (1995), "Bayesian Density Estimation and Inference Using Mixtures," *Journal of the American Statistical Association*, 90, 577–588.

Ferguson, T. S. (1973), "A Bayesian Analysis of Some Non-Parametric Problems," *The Annals of Statistics*, 1, 705–711.

Frees, E. W., and Ruppert, D. (1985), "Sequential Non-Parametric Age Replacement Policies," *The Annals of Statistics*, 13, 650–652.

Gelfand, A. E. (1996), "Model Determination Using Sampling-Based Methods," in *Markov Chain Monte Carlo in Practice*, eds. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, London: Chapman and Hall, pp. 145–161.

—— (1999), "Approaches for Bayesian Semi-Parametric Regression," in *Asymptotics, Nonparametrics, and Time Series: A Tribute to Madan Lal Puri*, eds. M. L. Puri and S. Ghosh, New York: Marcel Dekker.

Hjort, N. L. (1990), "Non-Parametric Bayes Estimates Based on Beta Processes in Models for Life History Data," *The Annals of Statistics*, 18, 1259–1294.

Jeffreys, H. (1961), *Theory of Probability* (3rd ed.), Oxford, U.K.: Oxford University Press.

Kalbfleisch, J. D. (1978), "Non-Parametric Bayes Analysis of Survival Time Data," *Journal of the Royal Statistical Society* B, 40, 214–221.

Kass, R. E., and Raftery, A. E. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773–795.

Kuo, L. (1986), "Computation of Mixtures of Dirichlet Processes," *SIAM Journal of Scientific and Statistical Computing*, 7, 60–71.

Kuo, L., and Mallick, B. K. (1997), "Bayesian Semi-Parametric Inference for the Accelerated Failure Time Model," *Canadian Journal of Statistics*, 25, 457–472.

Laud, P. W., Damien, P., and Smith, A. F. M. (1996), "Monte Carlo Methods for Approximating a Posterior Hazard Rate Process," *Statistics and Computing*, 6, 77–83.

—— (1998), "Bayesian Non-Parametric and Covariate Analysis of Failure Time Data," in *Practical Nonparametric and Semi-Parametric Bayesian Statistics*, eds. D. Dey, P. Muller, and D. Sinha, New York: Springer-Verlag, pp. 213–226.

Lindley, D. V., and Smith, A. F. M. (1972), "Bayes Estimates for the Linear Model" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 34, 1–41.

MacEachern, S. N. (1994), "Estimating Normal Means With a Conjugate-Style Dirichlet Process Prior," *Communications in Statistics, Part B—Simulation and Computation*, 23, 727–741.

—— (1998), "Computational Methods for Mixture of Dirichlet Process Models," in *Practical Nonparametric and Semi-Parametric Bayesian Statistics*, eds. D. Dey, P. Muller, and D. Sinha, New York: Springer, pp. 23–43.

Mazzuchi, T. A., and Soyer, R. (1989), "Assessment of Machine Tool Reliability Using a Proportional Hazards Model," *Naval Research Logistics*, 36, 765–777.

—— (1995), "A Bayesian Perspective on Some Replacement Strategies," *Reliability Engineering and System Safety*, 51, 295–303.

—— (1996), "Adaptive Bayesian Replacement Strategies," in *Bayesian Statistics 5*, eds. J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. Smith, Oxford, UK: Oxford University Press, pp. 667–671.

Mukhopadhyay, S., and Gelfand, A. E. (1997), "Dirichlet Process Mixed Generalized Linear Models," *Journal of the American Statistical Association*, 92, 633–639.

Newton, M. A., and Raftery, A. E. (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 56, 3–48.

Ross, S. M. (1989), "Estimating the Mean Number of Renewals by Simulation," *Probability in the Engineering and Informational Sciences*, 3, 319–321.

Smeitnik, E., and Dekker, R. (1990), "A Simple Approximation to the Renewal Function," *IEEE Transactions in Reliability*, 39, 71–75.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2002), "Bayesian Measures of Model Complexity and Fit," To appear in the *Journal of the Royal Statistical Society*, Ser. B, 64, 583–639.

Taraman, K. (1974), "Multi Machining Output—Multi-Independent Variable Turning Research by Response Surface Methodology," *International Journal of Production Research*, 12, 233–245.

West, M., Muller, P., and Escobar, M. D. (1994), "Hierarchical Priors and Mixture Models With Application in Regression and Density Estimation," in *Aspects of Uncertainty: A Tribute to D. V. Lindley*, eds. A. F. M. Smith and P. Freeman, Wiley, London, pp. 363–368.