

# Use of Sensitivity Analysis to Assess the Effect of Model Uncertainty in Analyzing Accelerated Life Test Data

William Q. Meeker  
Dept. of Statistics  
Iowa State University  
Ames, IA 50011

Luis A. Escobar  
Dept. of Experimental Statistics  
Louisiana State University  
Baton Rouge, LA 70803

August 9, 2001  
17h 51min

## **Abstract**

Accelerated life tests are used to obtain timely information on the durability and reliability of materials. Test units are subjected to higher than usual levels of “stress” and a model is used to estimate life at use conditions. Although it is desirable to use a physically-based model to justify the required extrapolation, in many practical situations, no such model is available or the physical basis for extrapolation is uncertain. In such situations, extrapolation is based on an empirical model. Sensitivity analysis tools then become important to assess the effect of model departures and to allow engineers to make safe design decisions.

This paper presents models, methods, and a description of software tools for performing systematic sensitivity analysis to assess potential model error. These methods are illustrated by studying the results of an experiment that was conducted to determine if the fatigue life of a spring would meet a given specification.

**Key words:** Box-Cox Transformation, Censored Data, Empirical Models, Extrapolation, Fatigue, Maximum Likelihood, Reliability, Weibull Distribution.

# 1 Introduction

## 1.1 Background

Product engineers needed information about the reliability of a spring in order to assess trade offs and to make design decisions for a product. In general, there is a tradeoff between the amount of displacement allowed in the motion of the spring and performance of the product. More displacement leads to higher performance but shorter fatigue life. A large experiment was conducted to determine if the fatigue life of the spring could be improved by using a new processing method and to obtain a quantitative description of the displacement-life tradeoff.

In order to protect proprietary data and information we generated simulated data from the fitted model for the original example, modified the scale of the data, and changed the name of one of the experimental factors. Largely, however, the nature of the application is the same as the original.

## 1.2 The experiment

A sample of 108 springs, divided equally between the new and the old processing method, were tested until failure or 5000 kilocycles (which ever came first). Two other factors, processing temperature and stroke displacement (distance that the spring was compressed in each cycle of the test) were varied in a two by three factorial arrangement with replication so that engineers could develop a regression relationship to describe the effects that these variables have on spring life. The stroke variable was being used as an accelerating variable. Nominal processing temperature and use conditions for these springs are 600°F and a stroke displacement of 20 mils, respectively (a mil is 1/1000 of an inch).

The goal of the experiment was to determine if the new processing method was better than the old method and to see if the spring would have a B10 life, the time by which one expects 10% of the devices to fail, equal to at least 500 megacycles (500,000 kilocycles) at use conditions. This customer specification would imply that no more than 10% percent of the springs would fail before the end of the technological life of the product in applications where the displacement would be 20 mils. If the specification cannot be met, then the product engineers want to know the amount of displacement that would be safe to use (the spring could still be used in the product by limiting displacement, at some loss of product performance).

## 1.3 The data

The data from the spring accelerated life test are given in Table 4. Time to failure is kilocycles. The explanatory variables are Temp in degrees Fahrenheit, Stroke in mils, and the class variable Model



on spring life, including various sensitivity analyses of model assumptions. Section 5 makes some concluding remarks and outlines some possible areas for further related work.

## 1.6 Software

The analyses shown here were done with SPLIDA (SPlus Life Data Analysis), a collection of S-PLUS functions with a graphical interface (GUI), designed for the analysis of reliability data. The most up-to-date version of SPLIDA can be downloaded from [www.public.iastate.edu/~splida](http://www.public.iastate.edu/~splida). Although some of the basic analyses might be possible in some advanced statistical packages like JMP, SAS, or MINITAB, the sensitivity analysis would probably require programming.

## 2 Weibull Distribution and Initial Data Analysis

### 2.1 Individual analyses at each factor-level combination

Analysis of accelerated life test data usually begins by fitting, individually, one or more distributions to each factor-level combination (or, more precisely, at those combinations where there were failures). We will illustrate fitting models based on the Weibull distribution. Other distributions were also investigated (details are not given here), but the Weibull seemed to provide the best fit to the data.

The Weibull distribution cdf is

$$F(t) = \Pr(T \leq t; \eta, \beta) = 1 - \exp \left[ - \left( \frac{t}{\eta} \right)^\beta \right], \quad t > 0. \quad (1)$$

In this parameterization,  $\beta > 0$  is a shape parameter and  $\eta > 0$  is a scale parameter as well as the approximate .632 quantile. The practical value of the Weibull distribution stems from its ability to describe failure distributions with many different commonly occurring shapes. See Section 4.8 of Meeker and Escobar (1998) for more information.

The Weibull distribution is a member of the log-location-scale family. In particular, the logarithm of a Weibull random variable has a smallest extreme value distribution with location parameter  $\mu = \log(\eta)$  and scale parameter  $\sigma = 1/\beta$ . In this form, the Weibull cdf is

$$F(t; \mu, \sigma) = \Phi_{\text{sev}} \left[ \frac{\log(t) - \mu}{\sigma} \right], \quad t > 0 \quad (2)$$

where  $\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$  is the cdf of the standardized ( $\mu = 0, \sigma = 1$ ) smallest extreme value distribution. This parameterization is more convenient for regression modeling of simple relationships between log life and explanatory variables (log-linear models).

Figure 2 is a multiple Weibull probability plot showing individual ML fits for each of the 11 factor-level combinations that had failures. The straight lines on the plots are the individual ML

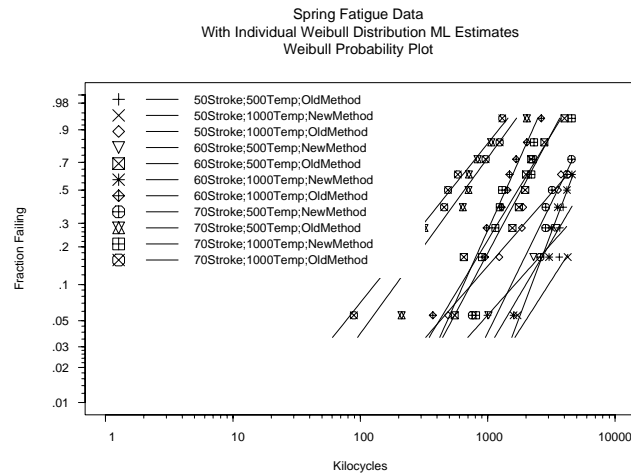


Figure 2: Weibull multiple probability plot of the spring accelerated life test data with individual ML estimates of  $F(t)$  and legend.

estimates of the cdf at each factor-level combination. Figure 3 provides the same information, but without the legend and with the data scaled more appropriately. The ML estimates of the Weibull parameters and corresponding standard errors are given in Table 1 for each condition.

## 2.2 Individual analyses with common Weibull shape parameter

Commonly used regression models and, correspondingly, commonly used accelerated test models assume that the scale parameter  $\sigma$  of the location-scale distribution does not depend on the explanatory variables. Meeker and Escobar (1998) show how this simple structure is implied by simple physical/chemical based models (e.g., a one step chemical reaction causing chemical degradation). On the other hand, it is easy to find counter examples, both in physical theory and in data (e.g., Pascual and Meeker, 1999). In any case, it is important to assess the adequacy of this model assumption. In order to do this, we fit a model that estimates separate distributions for all combinations of the explanatory variables, but with a common Weibull shape parameter. We call this the “floating scale model.” Fitting the floating scale model model is similar to the traditional one-way analysis of variance, but allows for censoring and distributions other than normal (or lognormal).

Figure 4 is a Weibull probability plot, similar to Figure 3, but with parallel lines, reflecting the common Weibull shape parameter model in the floating scale model that was fit to the data. The results of this model fit, including approximate 95% confidence intervals for the model parameters, are given in Table 2. The Intercept parameter estimate corresponds to the location parameter  $\mu$

Table 1: Weibull ML estimates of the individual parameters  $(\eta, \beta)$  at distinct factor-level combinations for the spring accelerated life test data.

Spring Fatigue Data

Maximum likelihood estimation results:

Response units: Kilocycles

Weibull Distribution

	Stroke, Temp, Method	Log likelihood	eta	se_eta	beta	se_beta
1	50Stroke;500Temp;NewMethod	NA	NA	NA	NA	NA
2	50Stroke;500Temp;OldMethod	-39.85	6639.6	2122.7	1.868	0.8845
3	50Stroke;1000Temp;NewMethod	-21.59	11104.8	7802.6	1.723	1.1772
4	50Stroke;1000Temp;OldMethod	-56.45	4430.4	1446.9	1.265	0.4536
5	60Stroke;500Temp;NewMethod	-31.17	10101.2	6480.1	1.236	0.6723
6	60Stroke;500Temp;OldMethod	-74.39	2193.1	372.5	2.067	0.5406
7	60Stroke;1000Temp;NewMethod	-55.10	4781.6	677.9	2.920	1.0591
8	60Stroke;1000Temp;OldMethod	-70.45	1606.2	228.4	2.466	0.6460
9	70Stroke;500Temp;NewMethod	-63.58	4169.5	700.5	2.250	0.7468
10	70Stroke;500Temp;OldMethod	-67.43	835.9	194.1	1.521	0.3799
11	70Stroke;1000Temp;NewMethod	-74.49	2108.8	408.3	1.831	0.4420
12	70Stroke;1000Temp;OldMethod	-66.08	665.7	169.7	1.378	0.3730

Total log likelihood= -620.6

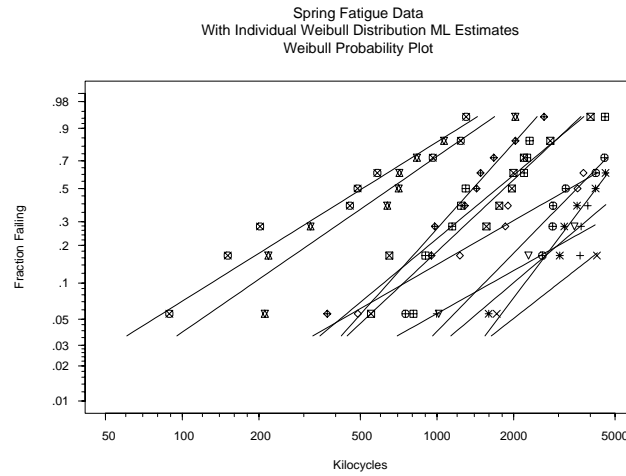


Figure 3: Weibull multiple probability plot of the spring accelerated life test data with individual ML estimates of  $F(t)$  using data-defined Kilocycles axis and no legend.

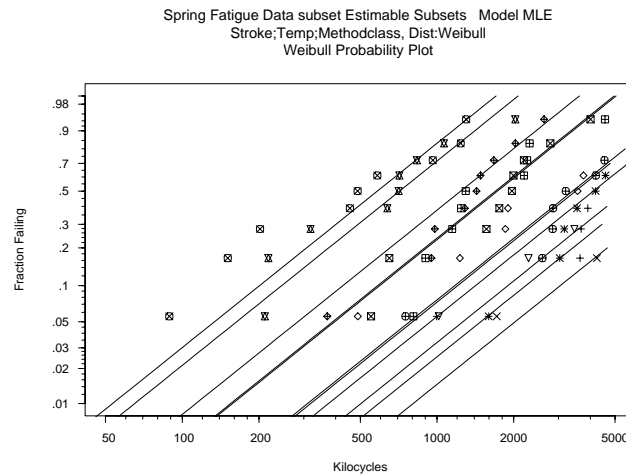


Figure 4: Weibull multiple probability plot of the spring accelerated life test data with individual ML estimates of  $F(t)$ , assuming a common Weibull shape parameter.

Table 2: Weibull ML estimate of the floating scale model for the spring accelerated life test data

Spring Fatigue Data

Maximum likelihood estimation results:

Response units: Kilocycles

Weibull Distribution

Log likelihood at maximum point: -623.9

Parameter	Approx Conf. Interval			
	MLE	Std.Err.	95% Lower	95% Upper
Intercept	8.9828	0.33058	8.3349	9.6307
70Stroke;500Temp;NewMethod	-0.6426	0.39322	-1.4133	0.1281
50Stroke;500Temp;OldMethod	-0.1609	0.43220	-1.0080	0.6862
60Stroke; 500Temp;OldMethod	-1.3203	0.38343	-2.0718	-0.5688
70Stroke; 500Temp;OldMethod	-2.2057	0.38629	-2.9628	-1.4486
50Stroke;1000Temp;NewMethod	0.3113	0.51700	-0.7020	1.3246
60Stroke;1000Temp;NewMethod	-0.4542	0.40150	-1.2412	0.3327
70Stroke;1000Temp;NewMethod	-1.3383	0.38483	-2.0925	-0.5840
50Stroke;1000Temp;OldMethod	-0.6151	0.40279	-1.4046	0.1743
60Stroke;1000Temp;OldMethod	-1.6525	0.38263	-2.4025	-0.9026
70Stroke;1000Temp;OldMethod	-2.4059	0.38573	-3.1619	-1.6499
sigma	0.5656	0.05429	0.4686	0.6827
weibull.beta	1.7680	0.16970	1.4649	2.1340

at the baseline condition “60Stroke; 500Temp; NewMethod,” which we denote by  $\mu_{\text{base}}$ . The other regression coefficients estimate  $\mu_i - \mu_{\text{base}}$  where the index  $i$  corresponds to each of the other factor-level combinations (except for “50Stroke; 500Temp; NewMethod” where there were no failures). Table 2 also provides ML estimates of the common Weibull shape parameter  $\beta$  and  $\sigma = 1/\beta$ .

### 2.3 Test for Weibull shape parameter homogeneity

Comparing Figures 3 and 4 show some differences among the shape parameter estimates obtained from the individual ML fits and the ML estimate obtained from the floating scale model. A formal test can be used to see if the observed differences can be explained by natural variability under the floating scale model. From Tables 1 and 2, the total log likelihood values from the corresponding models are  $-620.6$  and  $-623.9$ , respectively. The log likelihood ratio statistic for the comparison is

$Q = 2 \times (-620.6 - (-623.9)) = 6.6$ . In large samples, under the null hypothesis that the Weibull shape parameter is the same in all groups, the distribution of the log likelihood ratio statistic (under standard regularity conditions that are met here) has a chisquare distribution with degrees of freedom equal to the difference in the number of parameters in the full and the reduced models. The difference in the number of parameters estimated in the two models (again ignoring the “50Stroke; 500Temp; NewMethod” combination where there were no failures) is  $22 - 12 = 10$ . The approximate  $p$ -value for the test comparing these two different models is  $\Pr(\chi_{10}^2 > 6.6) = 0.237$ , indicating that the differences in slopes can be explained by chance alone.

## 2.4 Residual analysis

Figure 5 is a Weibull probability plot of the residuals for the floating scale model. This plot provides clear assessment of the Weibull distributional assumption, pooling all of the available data after they have been standardized to a common scale. Assessment of the adequacy of the distributional assumption at this stage of the modeling process has the advantage that the assessment can be done without the influence and possible bias introduced by the structure of a regression model.

The knee in the lower tail of the distribution seen in Figure 5 is the kind of deviation that can be expected from chance alone. This can be demonstrated by using Monte Carlo simulations like those described in Section 6.6.1 of Meeker and Escobar (1998). Thus it appears that there is not strong evidence against the assumption that the fatigue life distribution can be described by a Weibull distribution.

A lognormal analysis (details not included here) also provided a reasonable fit to the data, but it is not as good as the Weibull fit.

## 3 Response Surface Model Analysis

This section describes the construction and evaluation of a model that relates spring fatigue life to the factors in the experiment.

### 3.1 Acceleration/response surface model

The response surface model suggested by the engineers running the test for the spring fatigue life was

$$\mu = \beta_0 + \beta_1 \log(\text{Stroke}) + \beta_2 \text{Temp} + \beta_3 \text{Method} \quad (3)$$

where Method = 0 for the new method and Method = 1 for the old method. The log transformation for stroke was chosen on the basis of previous experience and tradition. There was no previous

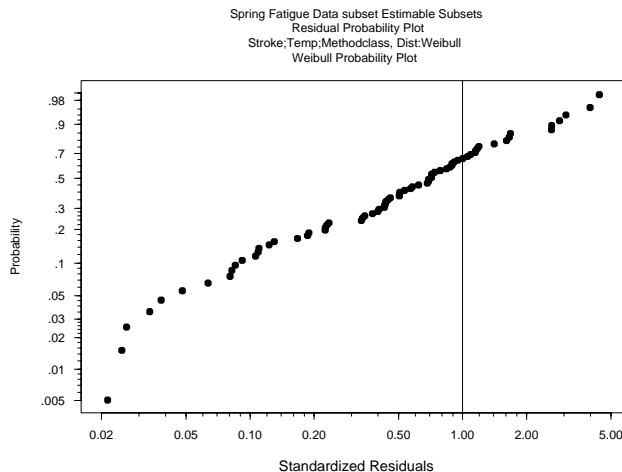


Figure 5: Weibull probability plot of the spring accelerated life test data residuals under the floating scale model.

experience relating the processing temperature to life so no transformation was used.

The results from fitting this model are given in Figure 6 and summarized in Table 3. This figure is similar to Figure 4 except that now the relationship between the lines is controlled by the model in (3). The response surface model allows computation of an estimate of  $F(t)$  at condition “50Stroke; 500Temp; NewMethod”, the right-most line on Figure 6, even though there were no failures at those conditions.

### 3.2 Test for departure from the acceleration/response surface model

To assess the adequacy of the response surface model, we can use Figure 6 to compare the non-parametric estimates (the plotted points) with the fitted  $F(t)$  lines, at each of the combinations of experimental factors. Because of the additional constraint of the model connecting the location of the distributions as a function of the explanatory variables, there will be more deviations between the nonparametric estimates and the fitted  $F(t)$  lines in Figure 6 than in Figure 4.

To test whether such deviations are statistically important, as opposed to being explainable by the natural variability in the data, under the model in (3), we can again do a likelihood ratio test, this time comparing the results in Tables 2 and 3. The total log likelihood values from the corresponding models are  $-623.9$  and  $-625.8$ , respectively. The log likelihood ratio statistic for the comparison is  $Q = 2 \times (-623.9 - (-625.8)) = 3.8$ . The difference in the number of parameters estimated in the two models (again ignoring the “50Stroke; 500Temp; NewMethod” combination where there were

Table 3: ML estimates of the linear response surface model for the spring accelerated life test data.

## Spring Fatigue Data

Maximum likelihood estimation results:

Response units: Kilocycles

Weibull Distribution

Relationship(s)

1 Stroke: Log

2 Temp: Linear

3 Method: Class

Log likelihood at maximum point: -625.8

Parameter	Approx Conf. Interval			
	MLE	Std.Err.	95% Lower	95% Upper
Intercept	32.026955	2.4843573	27.157704	36.8962056
Stroke	-5.509575	0.5872085	-6.660482	-4.3586674
Temp	-0.000883	0.0002709	-0.001414	-0.0003521
MethodOld	-1.272388	0.1475136	-1.561510	-0.9832672
sigma	0.569491	0.0539010	0.473067	0.6855692
weibull.beta	1.755953	0.1661969	1.458642	2.1138649

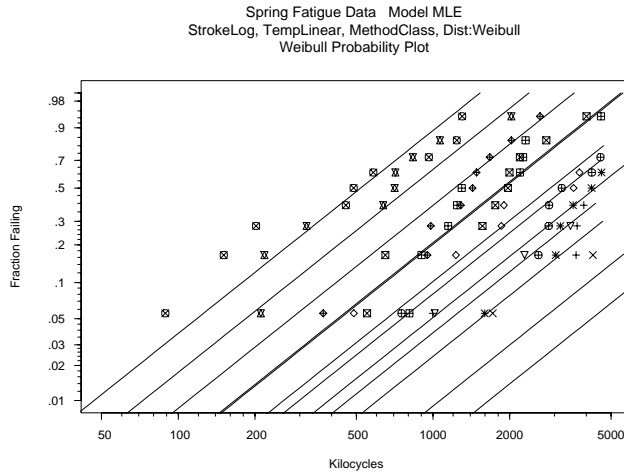


Figure 6: Weibull multiple probability plot of the spring accelerated life test data with ML response surface estimates of  $F(t)$ .

no failures) is  $12 - 5 = 7$ . The approximate  $p$ -value for the test comparing these two different models is  $\Pr(\chi_7^2 > 3.8) = 0.1975$ , indicating that the differences in positions of the fitted lines relative to the floating-scale model estimates can be explained by chance alone.

### 3.3 Testing for interactions

Although the model in (3) appears to fit the data well, it is important to make sure that there is no evidence of interaction in the data. To do such a check, we fit the following model in (4) that adds two-factor interactions to (3).

$$\begin{aligned} \mu &= \beta_0 + \beta_1 \log(\text{Temp}) + \beta_2 \text{Stroke} + \beta_3 \text{Method} \\ &+ \beta_4 \log(\text{Temp}) \times \text{Stroke} + \beta_5 \log(\text{Temp}) \times \text{Method} + \beta_6 \text{Stroke} \times \text{Method} \end{aligned} \quad (4)$$

Details of the model fit are not given here, but the total log likelihood for this model was  $-624.6$  which is very close to the value of  $-623.9$  for model (3). The log likelihood ratio statistics for the comparison is  $Q = 2 \times (-624.6 - (-623.9)) = 2.4$  with a difference in the number of parameters of  $8 - 5 = 3$ . The corresponding  $p$ -value is  $\Pr(\chi_3^2 > 2.4) = 0.716$ . Thus there is no evidence for interaction.

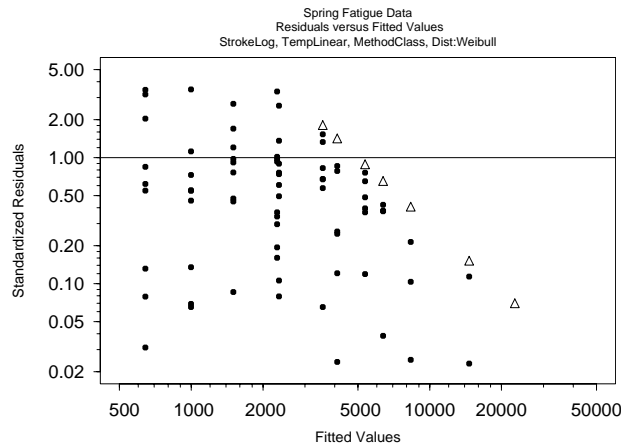


Figure 7: Plot of the spring accelerated life test data residuals versus fitted values.

### 3.4 Residual diagnostics

Residual plots provide another useful tool for detecting and visualizing departures from a fitted model. A traditional residual plot for this purpose plots the residuals versus the fitted values from the model. Such a plot for the spring fatigue life data and model (3) is shown in Figure 7. In this example, there is one vertical line of points for each of the twelve combinations of the explanatory variables (although the fitted values for some of the combinations are so close together that it is difficult to see any separation). The upward-pointing triangles indicate right-censored residuals, correspond to right-censored observations, indicating that the actual residual, had there been no censoring, would have been larger, higher in Figure 7. Recognizing the meaning of the censored residuals, this plot does not indicate any apparent deviation from the assumed model.

### 3.5 Comparison of old and new springs

One of the purposes of the experiment was to compare the new and the old processing methods with respect to fatigue life. Figure 8 is a conditional model plot, giving estimates of the quantiles of the fatigue life distribution for the new and old methods, conditional on values of the other factors fixed at  $\text{Temp} = 600$  and  $\text{Stroke} = 20$ , the nominal values of these variables. The densities are actually smallest extreme value densities, corresponding to Weibull distributions being plotted on a log response axis.

Figure 8 suggests that the new method will produce springs with a longer fatigue life. Quantifying this finding, the coefficient corresponding to  $\text{MethodOld}$  in Table 3 is  $-1.272$ . Under the

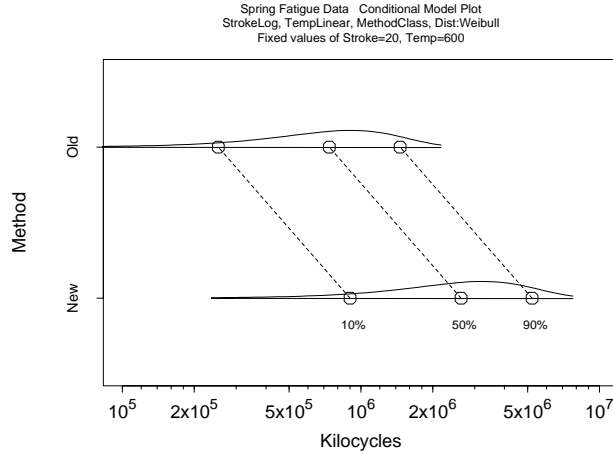


Figure 8: Conditional model plot showing the relationship between spring life and stroke displacement.

constant Weibull shape parameter model, this quantity has the interpretation of the difference between quantiles of the distributions of the old and new methods, on the log scale (e.g.,  $\mu_{\text{Old}} - \mu_{\text{New}}$ ). Alternatively, we can say that the ML estimate of any given quantile (such as B10) of the fatigue life distribution for the new method is  $\exp(\mu_{\text{Old}} - \mu_{\text{New}}) = \exp[-(\mu_{\text{New}} - \mu_{\text{Old}})] = \exp(1.272) = 3.57$  time larger than the corresponding ML estimate for the fatigue life distribution for the old method. Again taking numbers from Table 3, a 95% confidence interval for this improvement factor is

$$[\exp(0.9832672), \exp(1.561510)] = [2.67, 4.77].$$

Thus there is strong statistical evidence that the new processing method results in springs with a longer fatigue life distribution.

### 3.6 Estimate of B10 at nominal conditions

Figure 9 is similar to Figure 6, except that it shows the ML estimate for  $F(t)$  for the new processing method with Temp = 600 and Stroke = 20, the nominal values of these variables. This plot provides a visualization of the distribution of primary interest and the amount of extrapolation needed to make the desired inference.

The ML estimate of the 0.10 quantile (B10) of the fatigue life for the new processing method at Temp = 600 and Stroke = 20 is 900,221 kilocycles. A normal approximation 95% confidence interval for B10 is [237,471, 3,412,615] kilocycles. The details for how to compute this estimate and the associated confidence interval are given in Section 19.2.4 of Meeker and Escobar (1998).

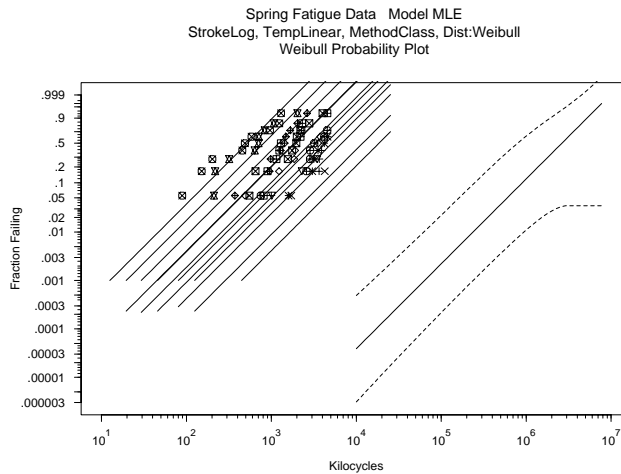


Figure 9: Weibull multiple probability plot of the spring accelerated life test data with ML response surface estimates with extrapolation to stroke displacement of 20 mils and normal approximation 95% pointwise confidence intervals.

Although the ML estimate exceeds the target value of 500,000 kilocycles by a sizable margin, the lower confidence bound leaves some doubt as to whether the new spring will meet the reliability target if operated at 20 mils.

## 4 Effect of stroke displacement on spring life

### 4.1 Conditional model plot

Figure 10 is a conditional model plot showing the fatigue life of springs manufactured with the new method at  $\text{Temp} = 600$ , as a function of Stroke. This plot also provides a useful visualization of extrapolation to  $\text{Stroke} = 20$  mils. Putting life on the horizontal axis is common practice in the fatigue literature when plotting cycles as a function of time.

The confidence intervals shown in Figure 9 shows that there is a large amount of uncertainty in the estimate of  $F(t)$  at the nominal use conditions. This is primarily due to the large amount of extrapolation. It is important to recognize, however, that the width of the confidence interval reflects only statistical uncertainty due to a limited sample size, and does not account for possible model error. In applications (such as the present application) model error could be substantially large. We investigate potential model error in the next section.

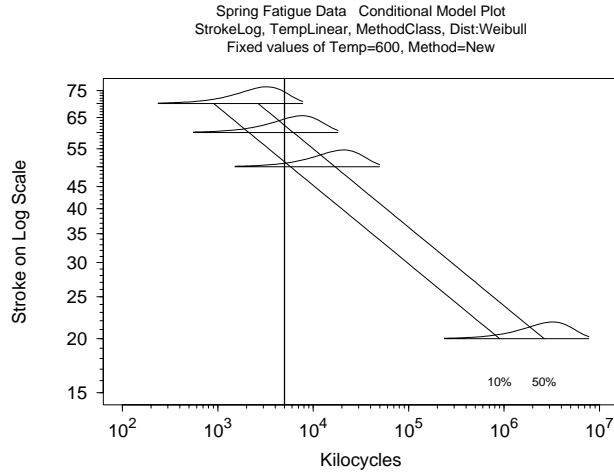


Figure 10: Conditional model plot showing the relationship between spring life and stroke displacement.

## 4.2 Sensitivity to the assumed form of the stroke-life relationship

The authors of the articles in Saltelli, Chan, and Scott (2000) describe a number of different approaches to sensitivity analysis. One general, but useful, approach to sensitivity analysis is to expand the formulation of the model and to investigate the effect of perturbing the parameter(s), over a range of the expansion, to see the effect on answers to questions of interest. We take this approach here by extending the model in (3). We start by replacing  $\log(\text{Stroke})$  with the more general Box-Cox transformation on Stroke. In particular, we fit the model

$$\mu = \beta_0 + \beta_1 X + \beta_2 \text{Temp} + \beta_3 \text{Method} \quad (5)$$

where

$$X = \begin{cases} \frac{(\text{Stroke})^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(\text{Stroke}) & \lambda = 0. \end{cases} \quad (6)$$

Note that because  $X$  is a continuous function of  $\lambda$ , (6) provides a continuum of transformations for possible evaluation and assessment. The Box-Cox transformation parameter  $\lambda$  can be varied over some range of values (e.g.,  $-1$  to  $2$ ) to see the effect of different assumed stroke-life relationships on the fitted model and inferences of interest. The results from the analyses can be displayed in a number of different ways.

For the spring fatigue life example, Figure 11 is a plot of the 0.10 Weibull quantile estimates versus  $\lambda$  between  $-1$  and  $2$ . Approximate confidence intervals are also given. Note that  $\lambda = 0$

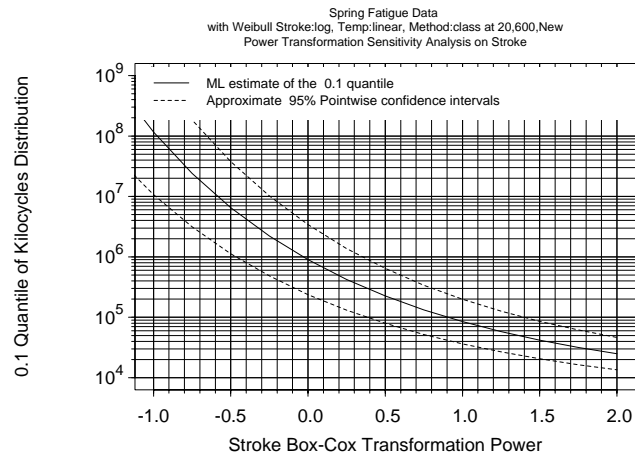


Figure 11: Plot of the 0.10 quantile of spring life versus the Stroke displacement Box-Cox transformation parameter with 95% confidence limits.

corresponds to the log transformation that is commonly used in fatigue life versus stress models and  $\lambda = 1$  corresponds to no transformation (or, more precisely, a linear transformation that affects the regression parameterization but not the underlying structure of the model). Figure 11 shows that fatigue life decreases by more than an order of magnitude as  $\lambda$  moves from 0 to 1. In particular, the ML estimate of the 0.10 quantile decreases from 900,221 kilocycles to 84,925 kilocycles when  $\lambda$  is changed from 0 to 1.

Figure 12 is a profile likelihood plot for the Box-Cox  $\lambda$  parameter, providing a visualization of what the data say about the value of this parameter. In this case the peak is at a value close to 0; this is in agreement with the commonly used fatigue life/stress model. Values of  $\lambda$  close to 1 are less plausible, but cannot be ruled out, based on the data alone.

The engineers, based on previous experience with the same failure mode and similar materials, felt that the actual value of  $\lambda$  was probably near to 0 (corresponding to the log transformation) and almost certainly less than 1. Then a conservative decision could be made by designing with an assumed value of  $\lambda = 1$ . In this case, the somewhat optimistic evaluation in Section 4 would become somewhat pessimistic, relative to the 500,000 kilocycle target.

### 4.3 Sensitivity to the assumed form of the temperature-life relationship

The same kind of sensitivity study can be done with the temperature variable. We omit the details here, as they are similar to the presentation in Section 4.2, but we present the graphical results.

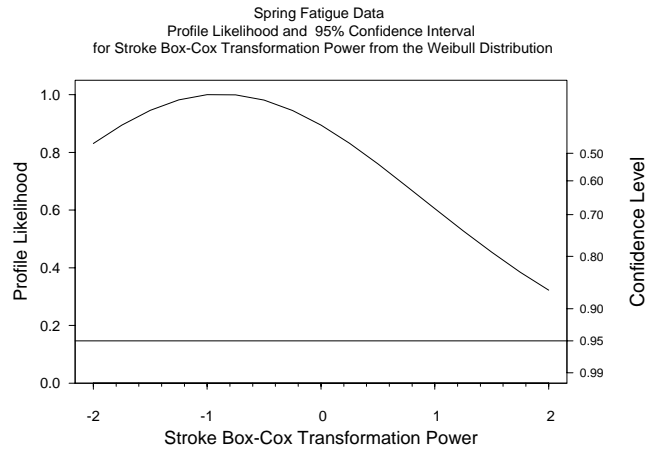


Figure 12: Profile likelihood plot for the Stroke Box-Cox transformation parameter in the spring life model.

Figure 13 is a conditional model plot showing the fatigue life of springs manufactured with the new method at Stroke = 20, as a function of Temp. This plot suggests that fatigue decreases somewhat with the level of this processing temperature. The nominal value of 600°F is, however, in the high-life region. The engineers were also concerned about the veracity of the assumed relationship between fatigue life and Temp, especially because they had no previous experience modeling this variable. Figure 14 is similar to Figure 11, except that it is the temperature relationship that is being perturbed. The figure shows that the estimates of the fatigue life distribution are not highly sensitive to the assumed relationship. Interestingly, the profile likelihood for the Temp Box-Cox transformation parameter (not shown here) is perfectly flat. Thus the data do not allow any information about the nature of the transformation.

#### 4.4 Sensitivity to the assumed distribution

The Weibull distribution seemed to fit the data well and it is easy to demonstrate that, in the tails of the distribution, the Weibull distribution will give predictions of life that are less optimistic relative to the lognormal distribution. To confirm this we can compare directly estimates from the two different distributions. Figure 15 is similar to Figure 11, except that instead of giving confidence intervals, plots of the 0.10 Weibull and lognormal quantiles versus  $\lambda$  are shown. The plot shows that the lognormal distribution estimates are optimistic relative to the Weibull distribution by about a factor 2 over this range of  $\lambda$ .

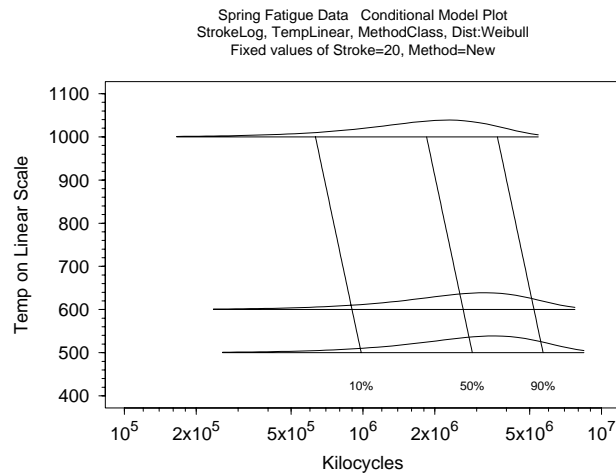


Figure 13: Conditional model plot showing the relationship between spring life and processing temperature.

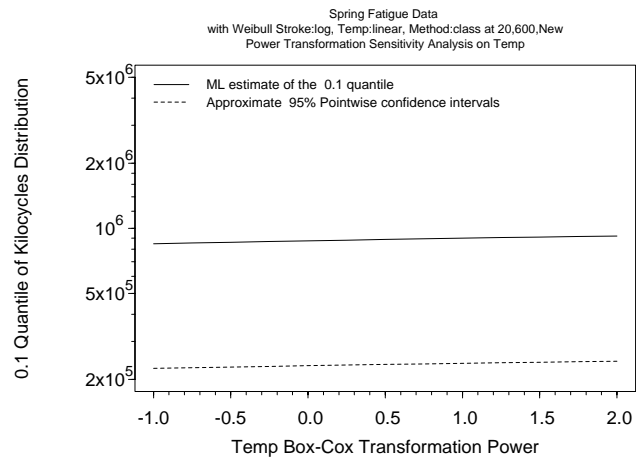


Figure 14: Plot of the 0.10 quantile of spring life versus the temperature Box-Cox transformation parameter with 95% confidence limits.

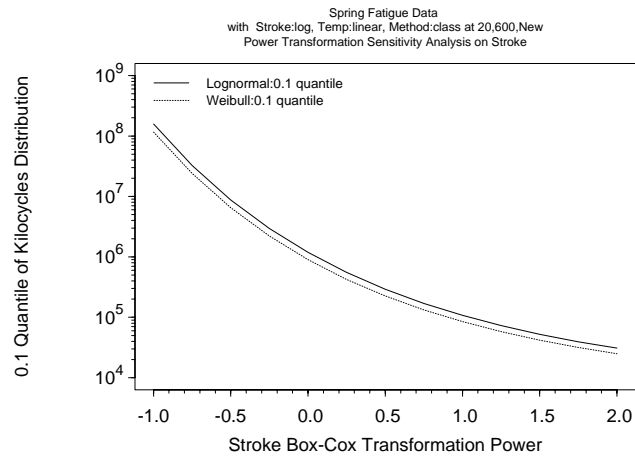


Figure 15: Plot of the 0.10 quantile of spring life versus the stroke displacement Box-Cox transformation parameter for the Weibull and lognormal distributions.

## 5 Concluding Remarks

It is often suggested that the extrapolation involved in accelerated testing requires the use of a model based on the physics or chemistry of the failure mechanism. While such a theoretical basis for extrapolation is clearly important and desirable, there are situations where important decisions need to be made on the basis of a model that is not firmly grounded in such theory. Due to limited time and resources that preclude the timely development of such theory, alternative approaches are needed. Most commonly, empirical models are used, based on previous experience and engineering judgment. Because of the uncertainty involved in the use of such models, sensitivity analyses becomes especially important.

In this paper we have illustrated how one can use simple sensitivity analyses to explore the effect of extrapolation and the sensitivity to perturbations to the assumed model. Even when a model based on the physics or chemistry of the failure mode is available, such sensitivity analyses is still important.

## A SPLIDA Commands for the Analyses

This appendix gives explicit direction on how to use the SPLIDA software to do the analyses described in this paper.

1. Use the data frame `NewSpring` (an example built into SPLIDA) to make the life data object

`NewSpring.ld`, using `Stroke`, `Temp`, and `Method` as explanatory variables. `Status` is the censoring variable. Use `Splida -> Make/edit/summary/view data object -> Summary/view data object` or the object browser to view `NewSpring.ld`.

2. First use `Splida -> Multiple regression (ALT) data analysis -> Censored data pairs plot` to get scatter plots for all pairs of variables. Note that censored observations are denoted in the plots by an open triangle ( $\Delta$ ) symbol.
3. Use `Splida -> Multiple regression (ALT) data analysis -> Censored data scatter plot` to make a scatter plot of the lifetimes versus `Temp` and lifetime versus `Stroke`. Use a log axes for life.
4. Use `Splida -> Multiple regression (ALT) data analysis -> Probability plot and ML fit for individual conditions` to obtain a multiple probability plot analysis to allow a comparison of the failure-time distributions for the different combinations of levels of the explanatory variables. Choose the Weibull distribution for analyses.
5. Use `Splida -> Multiple regression (ALT) data analysis -> Prob plot and ML fit for indiv cond: common shapes (slopes)` to fit a model that has the same distribution shape at each combination of the explanatory variables, but allows the scale of the distribution to float.
6. Use `Splida -> Multiple regression (ALT) data analysis -> Probability plot and fit of regression (acceleration) model` to begin fitting different regression models to the data. Choose `Stroke`, `Temp`, and `Method` (in that order) as the explanatory variables. Specify the nominal use conditions and `New method as 20;600;New` under “Specify new data for evaluation” on the `Basic` page of the dialog box. On the `Model` page, specify log relationships for the `Stroke` and `Temp` variables. Visit the `Tabular output` page and request tables of failure probabilities and quantiles at `20;600;New`. Click `Apply` and examine the results.
7. Use `Splida -> Regression residual analysis -> Residuals versus fitted values` to obtain plots of the residuals versus the fitted values to look for evidence of lack of fit for models that you fit.
8. Use `Splida -> Multiple regression (ALT) data analysis -> Conditional model plot` to get a plot of the estimates of the Spring life distribution as a function of `Stroke` when temperature is  $600^{\circ}\text{F}$  with the `New` processing method. On the `Plot options` page, request evaluation for `Stroke` over the range of 15 to 90 mils.

9. Do a similar evaluation, letting temperature vary from 400°F to 1000°F when Stroke=20, for the new processing method.
10. Use Splida -> Multiple regression (ALT) data analysis -> Sensitivity analysis plot to check the sensitivity of estimates of B01 to the assumed relationships for Stroke when Stroke is 20 mils and Temp is 600°F, and the processing method is “New.” On the Other inputs page, Specify the evaluation powers by entering -2,2,.5 in the appropriate cell. Then click on “Apply.”

**Questions:**

1. Compare the plots and the total log likelihood values from steps 4 and 5. What does this tell you? (The difference in the number of parameters estimated in the two steps is  $22 - 12 = 10$ . A log likelihood difference of  $\chi^2_{.80,10}/2 = 6.721$  or more between the log likelihood values might be considered to be big enough to be statistically important.)
2. What information did the conditional model plot in step 8 above provide?
3. Overall, which explanatory variable (Temp or Stroke) seems to have a more important effect over the range of values used in the experiment? Explain why.
4. If the old processing method had to be used, what would be a safe level of stroke displacement such that one could be 95% confidence that B10 will exceed 500 megacycles (500,000 kilocycles) if the processing temperature is 600°F?
5. For the temperature relationship sensitivity analysis, the likelihood profile plot for Box-Cox parameter is perfectly flat? Why?
6. In Section 4.2 it was noted that the Box-Cox transformation is a continuous function of the parameter  $\lambda$ . To see this, show that

$$\lim_{\lambda \rightarrow 0} \frac{x^\lambda - 1}{\lambda} = \log(x)$$

7. Explain why the Box-Cox transformation with  $\lambda = 1$  is, in terms of inferences on the failure time distribution, the same as no transformation.
8. Repeat the analysis in this paper using instead the lognormal distribution. How do the results compare within the range of the data? How do they compare outside of the range of the data?

## References

- Condra, L. W. (2001), *Reliability Improvement with Design of Experiments*, Second Edition. Marcel Dekker: New York.
- Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*. John Wiley & Sons: New York.
- Nelson, W. (1990), *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*. John Wiley & Sons: New York:
- Pascual, F. G. and Meeker, W. Q. (1999), Estimating Fatigue Curves with the Random Fatigue-Limit Model (with discussion). *Technometrics* 41, 277-302.
- Saltelli, A., Chan, K, and Scott, E. M., (2000), Editors, *Sensitivity Analyses*, New York: John Wiley & Sons.
- Wu, C. F. J. and Hamada, M. (2000) *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley & Sons: New York:

Table 4: The spring accelerated life test failure time data.

---

Kilocycles	Stroke	Temp	Method	Status	Weight
5000	50	500	New	Suspended	9
3464	60	500	New	Failed	1
1016	60	500	New	Failed	1
2287	60	500	New	Failed	1
5000	60	500	New	Suspended	6
2853	70	500	New	Failed	1
3199	70	500	New	Failed	1
752	70	500	New	Failed	1
2843	70	500	New	Failed	1
4196	70	500	New	Failed	1
2592	70	500	New	Failed	1
4542	70	500	New	Failed	1
5000	70	500	New	Suspended	2
997	50	500	Old	Failed	1
3904	50	500	Old	Failed	1
3674	50	500	Old	Failed	1
3644	50	500	Old	Failed	1
5000	50	500	Old	Suspended	5
2193	60	500	Old	Failed	1
2785	60	500	Old	Failed	1
4006	60	500	Old	Failed	1
1967	60	500	Old	Failed	1
1756	60	500	Old	Failed	1
650	60	500	Old	Failed	1
1995	60	500	Old	Failed	1
1563	60	500	Old	Failed	1
551	60	500	Old	Failed	1
211	70	500	Old	Failed	1
319	70	500	Old	Failed	1
712	70	500	Old	Failed	1
707	70	500	Old	Failed	1

---

*continued on next page*

---

Kilocycles	Stroke	Temp	Method	Status	Weight
2029	70	500	Old	Failed	1
638	70	500	Old	Failed	1
1065	70	500	Old	Failed	1
834	70	500	Old	Failed	1
218	70	500	Old	Failed	1
4241	50	1000	New	Failed	1
1715	50	1000	New	Failed	1
5000	50	1000	New	Suspended	7
3158	60	1000	New	Failed	1
3545	60	1000	New	Failed	1
4188	60	1000	New	Failed	1
4583	60	1000	New	Failed	1
1595	60	1000	New	Failed	1
3030	60	1000	New	Failed	1
5000	60	1000	New	Suspended	3
2196	70	1000	New	Failed	1
808	70	1000	New	Failed	1
2257	70	1000	New	Failed	1
1147	70	1000	New	Failed	1
1296	70	1000	New	Failed	1
1243	70	1000	New	Failed	1
2309	70	1000	New	Failed	1
4563	70	1000	New	Failed	1
901	70	1000	New	Failed	1
489	50	1000	Old	Failed	1
3756	50	1000	Old	Failed	1
1230	50	1000	Old	Failed	1
3562	50	1000	Old	Failed	1
1898	50	1000	Old	Failed	1
1855	50	1000	Old	Failed	1
5000	50	1000	Old	Suspended	3
1670	60	1000	Old	Failed	1
1481	60	1000	Old	Failed	1
371	60	1000	Old	Failed	1
2630	60	1000	Old	Failed	1

---

*continued on next page*

---

---

Kilocycles	Stroke	Temp	Method	Status	Weight
1285	60	1000	Old	Failed	1
2031	60	1000	Old	Failed	1
951	60	1000	Old	Failed	1
1429	60	1000	Old	Failed	1
980	60	1000	Old	Failed	1
963	70	1000	Old	Failed	1
1240	70	1000	Old	Failed	1
1301	70	1000	Old	Failed	1
455	70	1000	Old	Failed	1
151	70	1000	Old	Failed	1
488	70	1000	Old	Failed	1
202	70	1000	Old	Failed	1
89	70	1000	Old	Failed	1
583	70	1000	Old	Failed	1

---