Bayesian Degradation Analysis Considering Competing Risks and Residual-Life Prediction for Two-Phase Degradation

A thesis presented to
the faculty of
the Russ College of Engineering and Technology of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Shuluo Ning
August 2012
© 2012 Shuluo Ning. All Rights Reserved.
This thesis titled
Bayesian Degradation Analysis Considering Competing Risks and Residual-Life Prediction for Two-Phase Degradation

by

SHULUO NING

has been approved for
the Department of Industrial and Systems Engineering
and the Russ College of Engineering and Technology by

__________________________
Tao Yuan
Assistant Professor of Industrial and Systems Engineering

__________________________
Dennis Irwin
Dean, Russ College of Engineering and Technology
NING, SHULUO, M.S., August 2012, Industrial and Systems Engineering

Bayesian Degradation Analysis Considering Competing Risks and Residual-Life Prediction for Two-Phase Degradation (67 pp.)

Director of Thesis: Tao Yuan

This thesis develops different Bayesian hierarchical models to study two degradation phenomena. One phenomenon is the accelerated degradation path associated with catastrophic failures, which is called the competing risks problem. The other phenomenon is the two-phase degradation pattern which is caused by the remaining defects or contaminants developed during the manufacturing process. By incorporating prior information accumulated from the field testing into the proposed models, the Bayesian approach will play a significant role in improving the modeling accuracy for the observed degradation data. Gibbs sampling algorithms were developed for the inference of the parameters in the proposed models as well as for the prediction of failure-time distributions and the residual-life distribution. The proposed models and algorithms are validated by experimental data of light emitting diodes (LEDs) and plasma display panels (PDPs), respectively.

Approved:                                  

Tao Yuan

Assistant Professor of Industrial and Systems Engineering
ACKNOWLEDGMENTS

During my graduate studies in Ohio University, several persons and institutions collaborated directly and indirectly with my research. Without their support, it would be impossible for me to fulfill my graduate study.

First and foremost I offer my sincerest gratitude to my advisor, Dr. Tao Yuan, who gave me the opportunity to study with him and guided me throughout my master life with his patience and knowledge. Without his encouragement and effort, this thesis would not have been completed.

Besides my advisor, I would like to thank the rest of my thesis committee members: Dr. Diana Schwercha, Dr. David Koonce and Dr. Andrew Snow for their guidance, insightful comments, and suggestions.

I also would like to thank my girlfriend, Yizhen Ji, who supports me with her care and understanding.

Finally, I would like to thank my family, for their continuous support, inspiration, and love.
# Table of Contents

Abstract ................................................................. 3

Acknowledgments ....................................................... 4

List of Tables .......................................................... 7

List of Figures .......................................................... 8

1 Introduction .......................................................... 9
  1.1 Objective and motivation ....................................... 9
  1.2 Reliability and reliability data analysis ...................... 9
    1.2.1 Failure-time data analysis ............................... 10
    1.2.2 Degradation data analysis ............................... 12
  1.3 Competing risks in reliability .................................. 13
  1.4 Condition monitoring and residual-life distribution .......... 14
  1.5 Maximum likelihood estimation ............................... 14
  1.6 Bayesian statistical inference ............................... 15

2 Literature Review .................................................... 17
  2.1 Degradation analysis .......................................... 17
    2.1.1 Accelerated degradation testing .......................... 17
    2.1.2 Degradation analysis using ML method .................... 18
    2.1.3 Degradation analysis using Bayesian approach .......... 18
  2.2 Competing risks ................................................ 19
  2.3 Residual-Life Distribution ...................................... 20

3 Methodologies ....................................................... 23
  3.1 Competing risks model ......................................... 23
    3.1.1 Mixed-effects model for ADT data ........................ 23
    3.1.2 Bayesian hierarchical model for ADT data ................ 24
    3.1.3 Bayesian hierarchical model for competing risks problem . 25
    3.1.4 Deriving failure-time distributions ...................... 27
  3.2 Bayesian change-point degradation model ...................... 30
    3.2.1 Individual Bayesian Modeling Approach ................... 32
    3.2.2 Hierarchical Bayesian Modeling Approach for Multiple Units . 39
    3.2.3 Prediction of Residual Life Distribution .................. 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Case Study</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>LEDs Example</td>
<td>45</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Accelerated Degradation Model</td>
<td>46</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Competing Risks Model</td>
<td>47</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Failure-time Distribution</td>
<td>48</td>
</tr>
<tr>
<td>4.2</td>
<td>PDP Example</td>
<td>51</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Individual Degradation Modeling</td>
<td>51</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Hierarchical Bayesian Degradation Modeling</td>
<td>53</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Residual Life Distribution of a PDP</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion</td>
<td>60</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                                                 Page
4.1 MLEs for the parameters in degradation model . . . . . . . . . . . . . . . . . . 47
4.2 Bayesian posterior medians for the parameters in degradation model . . . . . 47
4.3 Bayesian inference for the competing risks model . . . . . . . . . . . . . . . . 48
4.4 Maximum likelihood estimates for the parameters in change-point degradation
model (3.12) and the simple linear degradation model(4.1). . . . . . . . . . . . 52
4.5 Bayesian inference for the parameters in the individual change-point degrada-
tion model (3.12). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
4.6 Posterior statistics of the model parameters of the hierarchical Bayesian
change-point degradation model (3.17), $\mu$, $\Sigma$, and $\sigma^2$, when the PDPs #1-#5
are analyzed together. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>pdf of Weibull distribution</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Observed degradation paths of six PDPs: (a) relative luminosity and (b) (log) transformed relative luminosity.</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>LEDs degradation data under 40mA and 35mA stress levels</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of soft failure-time distributions</td>
<td>49</td>
</tr>
<tr>
<td>4.3</td>
<td>Bayesian posterior median of the failure-time distributions $F_{T</td>
<td>S}(t_k)$, $F_{T</td>
</tr>
<tr>
<td>4.4</td>
<td>Predicted residual life distribution, $F_R(t)$, at five different inspection times, $t_k$: posterior median (solid line) and 95% Bayesian confidence bounds (dash lines).</td>
<td>57</td>
</tr>
<tr>
<td>4.5</td>
<td>Posterior medians of the residual life distribution, $F_R(t)$, at five different inspection times, $t_k$.</td>
<td>58</td>
</tr>
<tr>
<td>4.6</td>
<td>Predicted mean degradation paths at five different inspection times, $t_k$: posterior median (solid line) and 95% Bayesian confidence bounds (dash lines).</td>
<td>59</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 Objective and motivation

This thesis proposes to study two problems in applying degradation analysis for the reliability prediction: 1) degradation modeling considering both performance degradation and catastrophic failures and, 2) residual-life prediction for the two-phase degradation phenomenon. As the most popular method in statistics estimation, the maximum likelihood (ML) method is widely used in degradation analysis. However, since the degradation paths are often modeled in complex nonlinear functions, the maximum likelihood estimations (MLEs) need to be derived by approximation methods and are usually uncertain. Instead of the ML method, one of the advantages of the Bayesian statistics is that no approximation methods are needed for parameter estimation. The posterior statistics, such as mean or median, can be regarded as the point estimators for the unknown parameters. Another feature of Bayesian statistics is prior knowledge, which can be derived from previous experience and expert opinion. Bayesian approaches are employed to incorporate prior knowledge and real dataset in the degradation modeling and analysis within this thesis. Sampling methods can perform a significant function in calculating marginal posterior distributions for the unknown parameters.

The following sections will introduce the background on reliability and reliability data analysis.

1.2 Reliability and reliability data analysis

Today’s manufacturers suffer from enormous pressure to design and produce highly reliable products due to the increasing market competition and customers’ expectations [1]. Technically, the definition of reliability is the probability that a system, vehicle, machine, device, and so on will perform its intended function under operating conditions, for a specified period of time [2]. Reliability is usually evaluated by analyzing
reliability data collected in reliability tests or field operations. In general, two types of reliability data are considered: failure-time data and degradation data. Failure-time data describes the length of time in which a component stops working. Degradation data can continuously measure the product performance over time. Failure is defined when the performance degradation exceeds some threshold criterion.

1.2.1 Failure-time data analysis

The statistical analysis of failure-time data performs a significant function in the area of reliability engineering. There are various statistical distributions used in analyzing failure-time data. For example, Exponential, Weibull, and Lognormal distributions are frequently used due to their demonstrated usefulness in a wide range of situations. The Weibull distribution is perhaps the most useful distribution in failure-time data analysis. In the area of reliability, it can be applied in modeling failure-timed data or durability of diverse manufactured items, such as ball bearings, automobile components, and electrical insulation [3]. In this thesis, the Weibull distribution is used to describe catastrophic failure-time data. Its probability density function (pdf) and cumulative distribution function (cdf) have the following forms:

\[
f(t) = \frac{\eta}{\theta} \left(\frac{t}{\theta}\right)^{\eta-1} \exp\left\{ -\left(\frac{t}{\theta}\right)^\eta \right\}, \quad t \geq 0; \quad (1.1)
\]

\[
F(t) = 1 - \exp\left\{ -\left(\frac{t}{\theta}\right)^\eta \right\}, \quad t \geq 0.
\]

where \(\eta > 0\) is the shape parameter, \(\theta > 0\) is the scale parameter, and \(f(t)\) and \(F(t)\) denote, respectively, the pdf and cdf. \(\eta\) is related to the failure mechanism. Figure 1 illustrates the pdf of Weibull distribution when \(\eta < 1\), \(\eta = 1\), and \(\eta > 1\). When \(\eta = 1\), the distribution is identical to the exponential distribution. The scale parameter \(\theta\) influences both the mean and the spread of the distribution. This distribution is fairly flexible and can be used to provide a good description of various types of failure-time data.
However, because most modern products are designed to have a very long life under normal operational conditions, the failure-time data for such products are difficult to collect. For such applications, the accelerated life test (ALT) is widely used in obtaining failure information for highly reliable components and materials. During the ALT, failure information can be collected in short periods of time since the testing units are placed into high levels of one or more accelerating variables, e.g., voltage, current, humidity, temperature, etc. The failure information can then be transformed through an acceleration model to obtain estimates of life or long-term performance at the normal levels of the accelerating variables [2].

Acceleration models are used to describe the relationship between the stress conditions and the lifetime. For example, the Inverse Power Law (IPL) is widely used for estimating the failure-time of products under non-thermal accelerated stress [4]. It is given by

\[ L(V) = KV^n \]  

(1.3)

where \( L \) represents a quantifiable life measure, \( V \) represents the accelerating stress level, \( K \) and \( n \) are the parameters characteristic of the product (\( K > 0 \)).
1.2.2 Degradation data analysis

Today, many advanced devices are designed to work normally without failure for ten years or more. Few units will fail during the regular reliability tests. Therefore, it becomes difficult to use a traditional life test, which only uses failure-times to assess a product’s reliability. In recent years, degradation analysis has become increasingly common and important. Not only can degradation analysis improve the reliability inference over the standard failure-time analysis, but also its additional information is related to failure mechanisms [5].

Typically, degradation data contains two types of data: one is physical degradation as a function of time, e.g., tire wear; the other is performance degradation as a function of time, e.g., power output [2]. One popular method in degradation data analysis is based on the general degradation path model. The actual degradation path of a particular unit can be denoted by $D(t), t \geq 0$. The observed sample degradation path for unit $i$ at the measurement time $t_{ij}$ can be modeled as

$$y_{ij} = D(t_{ij}, \beta_i) + \epsilon_{ij}, \quad i = 1, \ldots, n, \ j = 1, \ldots, n_i, \quad (1.4)$$

where $D(t_{ij}, \beta_i)$ is the actual degradation path of $i$th unit at time $t_{ij}$, $\beta_i = (\beta_{i1}, \ldots, \beta_{ik})$ is a vector of $k$ unknown parameters for unit $i$, $\epsilon_{ij}$ is the measurement error which is usually assumed to be a normal random variable with a zero mean and variance $\sigma^2_{\epsilon}, \ i.e.\ \epsilon_{ij} \sim N(0, \sigma^2_{\epsilon})$. It is also reasonable to assume that the random components of the vector $\beta_i$ are independent of the $\epsilon_{ij}$ deviation [1]. Some of the $\beta_1, \ldots, \beta_k$ parameters will be random from unit to unit. One or more of the $\beta_1, \ldots, \beta_k$ parameters could be constant(s) across all units. $D_f$ is defined as the critical threshold for the degradation path. Thus, the failure time is the time when the actual degradation path exceeds the threshold $D_f$. Finally, the failure-time distribution for a particular product can be predicted by estimated degradation model parameters.
Some display devices, such as LEDs, have such high reliability that we are unable to obtain enough evidence of product degradation under normal operating conditions. Therefore, in order to test highly reliable products, the most common approach is ALT in which testing products are subjected to harsh environments. Accelerated degradation testing (ADT), which is the combination of ALT and degradation, indicates that the degradation analysis is conducted in various stress levels. In order to measure the performance of test units at normal use conditions, the relationship between stress and degradation path need to be extrapolated. Based on the acceleration model (1.3) and degradation path model (1.4), the degradation model parameter $\beta_i$ can be parameterized as $\beta_i = (K, n_i)$. $K$ is a fixed parameter which represents a material property and $n_i$ describes unit-to-unit variability.

### 1.3 Competing risks in reliability

In certain applications, products may fail due to more than one type of failure, or competing risks. The failure-time of products can be defined as the time when the first one of these risks occurs. Examples related to products that fail due to competing risks are shown in Nelson [4]. For instance, failure of the ball or the race may lead to the failure of ball bearing assemblies. Semiconductor devices can fail if either a lead or junction failure occurs. In this thesis, the proposed competing risks model contains two types of failures: soft failures and hard failures. For some devices, there is a continuous loss of performance during usage. For example, LEDs experience a gradual permanent reduction in light output during operation which is called light intensity degradation. The failure of LEDs is usually defined as the light output at a specified level such as 50% of initial performance, which is usually referred to as the definition of soft failure. During the operation, it is also possible that the devices stop working suddenly. These are called hard failures. Unlike soft failures, failure times for hard failures will not correspond exactly with a critical
threshold of degradation. They will be random from unit to unit and over time. In this study, the component may fail due to either the soft failure or hard failure. It is assumed that the failure time of the component is the time when its performance reaches a critical level or a sudden hard failure occurs.

1.4 Condition monitoring and residual-life distribution

One important application of degradation analysis is condition monitoring. In order to understand the performance of a system, condition monitoring performs a significant function in the process of collecting real-time sensor information from a operational system [6]. Not only can it eliminate the unnecessary maintenance for the system, but also it helps in avoiding failures by making failures more predictable. From the condition information, one can identify a degradation signal for capturing the current state of the system and predict the condition of the system in the future [4]. In this thesis, the future condition prediction is referred to as the residual-life estimation for the system. The residual-life distribution provides probabilistic estimates of failure time for use in replacement policies and maintenance management. To reduce the cost of maintenance activities, it is necessary to accurately assess the current state of system degradation and to precisely evaluate the remaining life of specific degrading components by extracting sensory information from condition monitoring [7].

1.5 Maximum likelihood estimation

The ML method is widely used in estimating the values of statistical model parameters given the sample data. Suppose that the likelihood function depends on $k$ parameters $\Theta = (\theta_1, \theta_2, \ldots, \theta_k)$ as follow:

$$L(\Theta) = \prod_{i=1}^{n} f(y_i|\Theta)$$  \hspace{1cm} (1.5)
The objective is to find the estimated values of $\theta_1, \theta_2, \ldots, \theta_k$ that maximize the likelihood function for given data denoted by $y_1, y_2, \ldots, y_n$. Because of the multiplicative form of the likelihood function, finding the derivatives of such functions is tedious. In general, the maximum of the natural logarithm of the likelihood function is used to solve the problem. Then finding first partial derivatives of the log-likelihood function with respect to $\theta_1, \theta_2, \ldots, \theta_k$ and setting these partial derivatives to zero are necessary for deriving the MLEs, i.e.,

$$\frac{\partial \ln L(\Theta)}{\partial \theta_j} = 0, \quad j = 1, \ldots, k$$

### 1.6 Bayesian statistical inference

The Bayesian approach is based on Bayes’ theorem and it is fundamentally different from the classical approach (ML method). The ML method assumes that the parameters are unknown, but fixed quantities. However, Bayesian statistics considers all unknown parameters as random variables and can be described by a joint prior distribution.

Initially, the prior distributions which express the information before any data are involved in the statistical analysis should be defined. They are usually based on historical data and/or expert knowledge. Then the joint posterior distribution $f(\Theta|y)$ of the parameters $\Theta$ can be calculated by observed data $y = (y_1, y_2, \ldots, y_n)$. According to the Bayes theorem, the posterior distribution can be written as follows

$$f(\Theta|y) = \frac{f(y|\Theta)f(\Theta)}{f(y)} \propto f(y|\Theta)f(\Theta) \quad (1.6)$$

The joint posterior distribution is a combination of prior distributions and observed data information, which is expressed by the joint prior distributions $f(\Theta)$ and the likelihood function $f(y|\Theta) = \prod_{i=1}^{n} f(y_i|\Theta)$ respectively.

Bayesian inference is based on the marginal posterior density for each parameter which can be obtained by integrating the joint posterior density function. Multi-level integrations are usually necessary to derive the marginal posterior density, but for
high-dimensional models these integrations are often computationally intractable and do not have closed forms. In these cases, Markov chain Monte Carlo (MCMC) simulation appears to be the easiest way to solve the problems. The idea of MCMC is to create a Markov process with a specified stationary distribution and then run the simulation long enough so that the distribution of current samples approximates the stationary distribution [8].

A particular MCMC algorithm that is useful in many multi-dimensional problems is called Gibbs sampling [8]. For each step of Gibbs sampling, random values for the parameters can be generated from conditional posterior distributions which are easy to derive. This algorithm can be expressed as follows

$$
\theta_1^{(s)} \text{ from } f(\theta_1 | \theta_2^{(s-1)}, \theta_3^{(s-1)}, \ldots, \theta_d^{(s-1)}, y),
$$

$$
\theta_2^{(s)} \text{ from } f(\theta_2 | \theta_1^{(s)}, \theta_3^{(s-1)}, \ldots, \theta_d^{(s-1)}, y),
$$

$$
\theta_3^{(s)} \text{ from } f(\theta_3 | \theta_1^{(s)}, \theta_2^{(s)}, \theta_4^{(s-1)}, \ldots, \theta_d^{(s-1)}, y),
$$

$$
\vdots
$$

$$
\vdots
$$

where $\theta_d$ represents the various parameters in the distribution, $s$ represents the $s$th iteration, and $y$ denotes the data. When the number of iteration is large enough, the sample drawn on one parameter can be regarded as simulated observations from its marginal posterior distribution. Functions of the model parameters can also be conveniently sampled.
2 Literature Review

This chapter reviews the literature related to this thesis. The literature review covers the following topics: degradation analysis, competing risks, and residual-life distribution.

2.1 Degradation analysis

Degradation analysis has attracted increasing interests for reliability evaluation since it can obtain reliability information of products in a short test period. Given the data collected during the test, the purpose is to estimate the model parameters and to predict lifetime distributions under normal conditions. For analyzing general/accelerated degradation model, both the ML method and Bayesian approach have been applied in previous research.

2.1.1 Accelerated degradation testing

ADT has been studied by many researchers since it can provide the researchers with more opportunities to draw quick inference on the lifetime distribution of highly reliable test items under normal use conditions. Meeker et al. [1] provided a practical guide to degradation analysis along with some standard acceleration models. Pan and Crispin [9] developed a hierarchical modeling approach to analyzing LEDs accelerated degradation process with nonlinear function and random coefficients. Park and Bae [10] provided three methods to extrapolate the lifetime distribution in a normal environment directly from accelerated degradation data without any assumptions for the degradation model. Modeling accelerated variables into stochastic degradation paths were also discussed in recent years. Padgett and Tomlinson [11] and Park and Padgett [12] both used a Gaussian process to describe a general degradation model with several accelerated variables in which observed failure times and degradation measures can be combined. A continuous cumulative damage approach was proposed to make inference on statistical parameters of
the product and system lifetime distributions. An extended ADT model, including Brownian motion with the linear drift model, was presented by Liao and Elsayed [13]. They developed this model to predict field reliability by considering the stress variations.

2.1.2 Degradation analysis using ML method

Most existing degradation studies employed the ML method to estimate the model parameters. Lu and Meeker [5] proposed a nonlinear mixed-effects degradation model to fit observed degradation data. Because MLEs of random-effect parameters are computationally intensive and do not have closed forms, they used a two-stage approximation method to estimate the parameters and confidence intervals. Lu et al. [14] introduced a random-coefficients regression model for analyzing hot-carrier-induced degradation in semiconductor devices and estimated the model parameters using the ML method. They also made statistical inference on failure-time distribution and conducted simulation studies to examine the quality of ML estimation. Bae and Kvam [15] developed a nonlinear random-coefficients model to describe the degradation path of vacuum fluorescent displays. They considered and evaluated four different approximation methods for the MLEs and derived failure-time distribution using Monte Carlo simulation. Bae et al. [16] presented a bi-exponential model with random-coefficients to analyze the nonlinear degradation paths caused by nano-contamination in plasma display panels. The MLEs of the parameters in this model were estimated by Lindstrom and Bates algorithm [17].

2.1.3 Degradation analysis using Bayesian approach

However, the Bayesian approach has not been widely applied in degradation analysis. Wakefield et al. [18] used a Bayesian method to analyze repeated degradation measures in both linear and non-linear population models. Pettit and Young [19] modeled degradation process as a Wiener process and assumed that failure-times follow an an Inverse Gaussian
distribution. Then they proposed a fully Bayesian approach for integrating failure-time data with degradation data to derive failure-time distributions. Fatigue crack growth in components, influenced by cyclic loading, was modeled as a nonlinear random coefficients degradation model by Robinson and Crowder [20]. They applied Bayesian inference in estimating model parameters and predicting the failure-time distribution of currently tested units and future units. Onar et al. [21] developed Bayesian linear mixed-effects models to describe the degradation paths for rut depth.

### 2.2 Competing risks

There are numerous studies on competing risks in failure-time analysis. Most of these competing risks models are under fixed experimental conditions or accelerated life tests. Bai and Bai [22] described how electronic devices and components can fail when both intrinsic failures and extrinsic failures exist. Intrinsic failures are caused by wear-out, while extrinsic failures are due to defects generated during the production. They applied a mixture of two Weibull distributions to model the two failure modes in accelerated life test. Cox [23] considered some exponential models that can be used for interpreting failure data when there are two or more types of failures. Chan and Meeker [24] proposed a mixture of lognormal and Weibull models to describe infant-mortality and wear-out failures. Park [25] employed an Expectation-Maximization (EM) algorithm, which consists of an expectation step (E-step) and a maximization step (M-step), to estimate the parameters of a competing risk model. Exponential and lognormal examples were presented to illustrate the parameter estimation in his article.

The majority of existing reliability methods for competing risks concentrate on products subject to hard failures. However, along with the improvement of the quality and reliability, few or no hard failures will occur at normal conditions, even at accelerated conditions. In general, many products, such as semiconductors and microelectronics, may
also be under continuous deterioration over time (degradation). In existing literature, competing risk problems involving both performance degradation and catastrophic failures (hard failures) were seldom investigated. Bagdonavicius and Nikulin [26] introduced theoretical models for accelerated degradation processes with catastrophic failures. Zuo et al. [27] developed a mixture model which is suitable for modeling both catastrophic failures and degradations failures. A maintenance model which jointly models the degradation processes and sudden failures was shown in Zhu et al. [28]. Zhao and Elsayed [29] considered a model of competing risk under accelerated conditions with two types of failure: degradation and hard failure. The degradation process was modeled by a Brownian motion process, while the time of hard failures was modeled as Weibull random variables. Based on real experimental data, they used the ML method to estimate the parameters, and then predicted the lifetime distribution.

Unlike Zhao and Elsayed’s [29] work, this thesis presents a general degradation path model to analyze degradation data and applies Bayesian hierarchical approach in analyzing the competing risks model.

2.3 Residual-Life Distribution

Residual-life distribution can be used to predict the failure information of a system for use in maintenance or replacement plans. Gebraeel [6, 30, 31] made several contributions on computing residual-life distributions based on degradation signals. Gebraeel et al. [6] developed a Bayesian updating approach to estimate the stochastic parameters in linear and log-linear degradation signal models. The random error terms were modeled by normal random variables and a Brownian motion process, respectively, in two models. Then they used these models to predict residual-life distributions for a single degraded system. Chakraborty et al. [30] extended their research [6] by assuming that the stochastic parameters within the linear degradation signal model follow a gamma
distribution instead of normal distribution. They also applied a simulation-based algorithm in estimating residual-life distributions and concluded that using the gamma prior distribution can perform better prediction than that presented in [6]. In contrast to the constant environmental or operational conditions used in [6] and [30], Gebraeel and Pan [31] presented a new stochastic degradation model which concentrates on computing and continuously updating residual-life distributions of partially degraded components under time-varying environment. Different from the previous research, Gebraeel et al. [32] developed a two-stage prognostic methodology on the residual-life predictions in which prior degradation signals are absent. They first used failure time information along with the characteristics of the Bernstein distribution to derive prior distributions and then computed residual-life distributions by utilizing real-time degradation signals with estimated stochastic parameters.

All of these studies considered linear or log-linear degradation paths. It is desirable to extend these studies to consider non-linear degradation paths. This thesis presents a residual-life prediction for products exhibiting two-phase degradation. A two-phase degradation model can also be called: a degradation model with a change-point. Bae and Kvam [33] analyzed the Plasma Display Panel (PDP) degradation data with a change-point model and used ML method to estimate the unknown time change-point within the degradation path. Ng [34] proposed an independent-increments stochastic process with a random change-point to model the nonlinear degradation paths of PDP. He developed an expectation-maximization (EM) algorithm for the change-point regression model with an analytically intractable likelihood function to obtain the ML estimators efficiently. Other applications exhibiting two-phase degradation patterns include direct methanol fuel cells (DMFCs) [35] and organic light emitting diodes (OLEDs) [10]. The degradation paths of both membrane electrode assemblies (MEAs) in DMFCs and OLEDs were fitted by a bi-exponential model with random-coefficients in previous studies [10, 35].
In this thesis, a Bayesian hierarchical change-point degradation model will be developed for analyzing PDP degradation data.
3 Methodologies

There are two problems considered in this chapter: One is using a hierarchical Bayesian approach to analyze the competing risks problem involving accelerated degradation path and catastrophic failures; the other is proposing a change-point degradation model to predict the residual-life distribution.

3.1 Competing risks model

The first problem is motivated by an experimental study concerning the degradation of LEDs [36]. According to Mitsuo [36], the true degradation path $D(t)$ of a LEDs light intensity at $t$ can be expressed as

$$D(t) = D_0 \exp(-\gamma \cdot t), \quad t \geq 0 \quad (3.1)$$

where $D_0$ is the initial intensity, and $\gamma$ is the degradation rate. Here, the degradation-stress relationship with IPL for degradation rate $\gamma$ is modeled as

$$\gamma = k_1 \cdot I^m \quad (3.2)$$

where $I$ is the stress level (mA).

For this accelerated degradation path, a mixed-effects model and a Bayesian hierarchical model are presented to describing the ADT data. In the following two sections, these two models will be introduced respectively.

3.1.1 Mixed-effects model for ADT data

During the experiment, each LED item may experience different sources of variation. Therefore, one can easily analyze the behavior of a particular degradation over time as well as the unit-to-unit variability in degradation path by using a mixed-effects model. Based on the actual degradation path model (3.1) and acceleration factor (3.2), the
observed degradation path model for LEDs can be expressed as

\[ y_{ij} = \ln \left( \frac{D_i(t)}{D_i(0)} \right) = D(t_{ij}, \beta_i) + \epsilon_{ij} = -\gamma_i \cdot t_{ij} + \epsilon_{ij} = -k_1 I^{m_i} \cdot t_{ij} + \epsilon_{ij} \]  

(3.3)

where \( \beta_i = (k_1, m_i) \), \( m_i \sim N(\mu_m, \sigma^2_m) \), \( \epsilon_{ij} \sim N(0, \sigma^2_e) \), and \( I \) is the known accelerated stress level. It is assumed that rate-accelerated parameter \( k_1 \) is an unknown fixed effect for all units and the random effect \( m_i \) expresses the variation between the units. \( \epsilon_{ij} \) is also assumed to be independent of the parameter vector \( \beta_i \).

The likelihood function for the above mixed-effects accelerated degradation model (3.3) is given by

\[ L(k_1, \sigma^2_e, \mu_m, \sigma^2_m) = \prod_{i=1}^{n} \left[ \int \left( \prod_{j=1}^{n_i} f(d | k_1, \sigma^2_e, m_i) \right) \times f(m_i) dm_i \right] \]  

(3.4)

where \( f(d | k_1, \sigma^2_e, m_i) = \frac{1}{\sqrt{2\pi \sigma_e}} \exp \left[ -\frac{(y_{ij} - k_1 I^{m_i} \cdot t_{ij})^2}{2\sigma^2_e} \right] \), and \( f(m_i) \) is the pdf of \( m_i \).

Maximizing (3.4) with respect to \( (k_1, \sigma^2_e, \mu_m, \sigma^2_m) \) will not have closed-form expressions since the actual degradation path \( D(t) \) is a nonlinear function. In order to provide approximate MLEs for these parameters, a modification of the method of Lindstrom and Bates [17] was implemented by Pinheiro and Bates [37]. This maximum likelihood function can be computed by a R/S-Plus function named \textit{nlme} [38] which was written for giving better approximation for accelerated degradation data analysis.

For analyzing the same dataset which describes unit-to-unit variability, a Bayesian hierarchical model can play the same role. The following section will introduce a hierarchical Bayesian model to analyze the ADT data.

### 3.1.2 Bayesian hierarchical model for ADT data

By using the Bayesian approach, for accelerated degradation model (3.3), prior distributions for \( k_1, \sigma^2_e, \mu_m, \sigma^2_m \) should be assigned to complete the model specification. Independent normal prior distributions are assumed for \( k_1 \) and \( \mu_m \); and independent Inverse-Gamma (IG) prior distributions are assigned to \( \sigma^2_m \) and \( \sigma^2_e \). IG distribution is the
conjugate prior for the normal variance which is called precision parameter of the normal distribution. The prior distributions of these parameters can be expressed as

\[ k_1 \sim N(\mu_1, \sigma^2_1), \quad \mu_m \sim N(\mu_0, \sigma^2_0) \]
\[ \sigma^2_m \sim IG(\alpha_m, \beta_m), \quad \sigma^2_e \sim IG(\alpha_e, \beta_e) \]

where \( \mu_1, \sigma^2_1, \mu_0, \sigma^2_0, \alpha_m, \beta_m, \alpha_e, \) and \( \beta_e \) are pre-specified hyper parameters. Based on the Bayes theorem, the joint posterior distribution for all the model parameters conditioning on data can be derived as follows

\[
\begin{align*}
    f(k_1, \mu_m, \sigma^2_m, \sigma^2_e) \propto & \prod_{i=1}^{n} \prod_{j=1}^{n_i} \frac{1}{\sigma_e} \exp \left[ \frac{(y_{ij} + k_1I_{ij}^{m_i} \cdot t_{ij})^2}{2\sigma^2_e} \right] \cdot \frac{1}{\sigma_m} \exp \left[ \frac{-(m_i - \mu_m)^2}{2\sigma^2_m} \right] \\
    & \times f(k_1) f(\mu_m) f(\sigma^2_m) f(\sigma^2_e)
\end{align*}
\]

(3.5)

where \( n \) is the number of tested units, \( n_i \) is the number of measurements on the \( i \)th tested unit, \( y_{ij} \) is the \( j \)th measurement observed at time \( t_{ij} \) for the \( i \)th tested item, and \( f(k_1), f(\mu_m), f(\sigma^2_m), \) and \( f(\sigma^2_e) \) are the pdf of prior distributions for \( k_1, \sigma^2_m, \mu_m, \) and \( \sigma^2_m, \) respectively.

This section presents a Bayesian hierarchical model to describe testing items suffering from degradation over time. Another Bayesian hierarchical model considered two types of failures for testing items will be introduced in next section.

### 3.1.3 Bayesian hierarchical model for competing risks problem

This section proposes a Bayesian hierarchical model to describe the LEDs suffering from two competing failure modes involving accelerated degradation path (soft failures) and catastrophic failures (hard failures). As mentioned in Section 3.1.2, (3.3) can be used to model the degradation process. On the other hand, for the hard failure, the hard failure-time is assumed to follow a Weibull distribution. Its probability density function can be defined as (1.1). Moreover, IPL is used to describe life-stress relationship for scale parameter \( \theta \) as follow

\[ \theta = k_2 \cdot T^l \]
where \( k_2 \) and \( l \) are model parameters to be determined. The prior distributions for the parameters of hard failure model are given by

\[
k_2 \sim N(\mu_2, \sigma^2_2), \quad l \sim N(\mu_l, \sigma^2_l), \quad \eta \sim \text{Gamma}(a, b)
\]

Thus, based on Bayes theorem, the joint posterior distribution for all parameters of this competing risks problem can be expressed as

\[
f(k_1, k_2, \beta, l, \mu_m, \sigma^2_m, \sigma^2_e|\mathbf{d}) \propto \prod_{i=1}^{n} \prod_{j=1}^{n_i} \frac{1}{\sigma^2_e} \exp \left[ \frac{(y_{ij} + k_1 t_{ij} m_i + t_{ij})^2}{2\sigma^2_e} \right] \cdot \frac{1}{\sigma^2_m} \exp \left[ -\frac{(m_i - \mu_m)^2}{2\sigma^2_m} \right] \\
\times \left[ \frac{\eta}{k_2 l} \right]^{\eta-1} \delta_i \exp \left[ -\left( \frac{t_i}{k_2 l} \right)^\eta \right] \times f(k_1) f(k_2) f(\eta) f(l) \\
f(\mu_m) f(\sigma^2_m) f(\sigma^2_e)
\]

(3.6)

where \( \delta_i = 1 \), if hard failure occurs; \( \delta_i = 0 \), if hard failure does not occur.

For the Bayesian hierarchical model (3.5) and (3.6), inference on each parameter is based on its marginal posterior distribution. As mentioned in section 1.6, Gibbs sampling is applied in this multidimensional problem since multiple levels of integration for marginal posterior distributions are often analytically intractable. It is usually straightforward to derive the conditional posterior distributions required for Gibbs sampling. For Bayesian hierarchical model (3.6), the conditional posterior distribution of each parameter can be easily obtained:

1. The conditional posterior distribution of \( k_1 \) is

\[
k_1|k_2, \eta, l, \mu_m, \sigma^2_m, \sigma^2_e, \mathbf{d} \sim N \left( \frac{\mu_1 \sigma^2_e - \sigma^2_1 \sum_{j=1}^{n_i} (I_{m_i} t_{ij})^2 y_{ij}}{\sigma^2_e + \sigma^2_1 \sum_{j=1}^{n_i} (I_{m_i} t_{ij})^2}, \frac{\sigma^2_2 \sigma^2_1}{\sigma^2_e + \sigma^2_1 \sum_{j=1}^{n_i} (I_{m_i} t_{ij})^2} \right)
\]

2. The conditional posterior distribution of \( \mu_m \) is

\[
\mu_m|k_1, k_2, \eta, l, \sigma^2_m, \sigma^2_e, \mathbf{d} \sim N \left( \frac{\sigma^2_m \mu_0 + \sigma^2_0 \sum_{i=1}^{n} m_i}{\sigma^2_m + n \sigma^2_0}, \frac{\sigma^2_2 \sigma^2_0}{\sigma^2_m + n \sigma^2_0} \right)
\]
(3) The conditional posterior distribution of $\sigma_m^2$ is

$$
\sigma_m^2 | k_1, k_2, \eta, l, \mu_m, \sigma_e^2, d \sim IG \left( \frac{n}{2} + \alpha_m, \frac{1}{2} \sum_{i=1}^{n} (m_i - \mu_m)^2 + \beta_m \right)
$$

(4) The conditional posterior distribution of $\sigma_e^2$ is

$$
\sigma_e^2 | k_1, k_2, \eta, l, \mu_m, \sigma_m^2, d \sim IG \left( \frac{\sum_{i=1}^{n} N_i}{2} + \alpha_e, \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} + k_1 m_i t_{ij})^2 + \beta_e \right)
$$

(5) The conditional posterior probability density function of $k_2$ is

$$
f(k_2 | k_1, \eta, l, \mu_m, \sigma_m^2, \sigma_e^2, d) \propto \prod_{i=1}^{n} \left[ \frac{\eta(t_i)^{\eta-1} \delta_i}{(k_2 l_i)^\eta} \right] \cdot \exp \left[ - \sum_{i=1}^{n} \left( \frac{t_i}{k_2 l_i} - \frac{(k_2 - \mu_2)^2}{2 \sigma_2^2} \right) \right]
$$

(6) The conditional posterior probability density function of $\eta$ is

$$
f(\eta | k_1, k_2, l, \mu_m, \sigma_m^2, \sigma_e^2, d) \propto \prod_{i=1}^{n} \left[ \frac{\eta(t_i)^{\eta-1} \delta_i}{(k_2 l_i)^\eta} \right] \cdot \exp \left[ - \sum_{i=1}^{n} \left( \frac{t_i}{k_2 l_i} \right)^\eta - \frac{\eta}{b} \right] \cdot \eta^{\eta-1}
$$

(7) The conditional posterior probability density function of $l$ is

$$
f(l | k_1, k_2, \eta, \mu_m, \sigma_m^2, \sigma_e^2, d) \propto \prod_{i=1}^{n} \left[ \frac{\eta(t_i)^{\eta-1} \delta_i}{(k_2 l_i)^\eta} \right] \cdot \exp \left[ - \sum_{i=1}^{n} \left( \frac{t_i}{k_2 l_i} \right)^\eta - \frac{(l - \mu_l)^2}{2 \sigma_l^2} \right]
$$

In this thesis, WinBUGS [39], a specialized software package for implementing the Gibbs sampling, is used to complete MCMC simulation and posterior inference.

3.1.4 Deriving failure-time distributions

Deriving failure-time distributions is a critical issue related to competing risks model analysis. Firstly, soft failure-time $T_s$ of a randomly selected unit is defined as the first hitting time that actual degradation path reaches the pre-specified degradation level $D_f$.

The distribution function of $T_s$ is as follow

$$
F_{T_s}(t|\Theta) = Pr(T_s \leq t|\Theta) = Pr(D(t, k_1, m) + \epsilon \leq D_f|\Theta) = \Phi \left( \frac{D_f - D(t, k_1, m)}{\sigma_\epsilon} \right)
$$

(3.7)
where $\Theta = (k_1, \mu_m, \sigma^2_m, \sigma^2_e)$, and $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution. Secondly, based on the same procedure, the distribution function of hard failure-time $T_h$ can be derived as follows:

$$F_{T|H}(t|\Psi) = Pr(T_h \leq t|\Psi) = 1 - \exp \left( \frac{t}{k_2} \right)$$

where $\Psi = (\eta, k_2, l)$. Finally, in the competing risk model, distribution function of failure-time $T_c$ can be expressed as

$$F_{T|C}(t) = 1 - (1 - F_{T|S}(t))(1 - F_{T|H}(t)) \quad (3.8)$$

To estimate $F_{T|S}(t|\Theta)$ given by (3.7), Bae et al. [16] introduced a simulation-based algorithm to compute the MLEs of $F_{T|S}(t|\Theta)$. The parameters $\Theta$ are estimated with MLEs $\hat{\Theta} = (\hat{k}_1, \hat{\mu}_m, \hat{\sigma}^2_m, \hat{\sigma}^2_e)$ obtained through approximation method by Pinheiro and Bates [37]. Then ML method is used to estimate $F_{T|S}(t|\Theta)$ by replacing $\Theta$ with their estimates $\hat{\Theta}$. However, since there is no closed-form expression for $F_{T|S}(t|\Theta)$, $\hat{F}_{T|S}(t|\hat{\Theta})$ can be evaluated by using Monte Carlo simulation as in Lu and Meeker [5]. For this method, the estimated model parameters $\hat{\mu}_m$ and $\hat{\sigma}^2_m$ are used to generate the $N$ simulated realizations $\hat{m}$. From $N$ values of $\hat{m}$, the $N$ soft failure-times $\hat{t}$ can be computed by substitute $\hat{m}$ into $D_f = D(t, \hat{k}_1, m) + \epsilon$ where the values of $\epsilon$ are generated from $N(0, \sigma^2_e)$. For any desired values of $t$, $F_{T|S}(t|\Theta)$ is estimated from the simulated empirical distribution

$$\hat{F}_{T|S}(t) = \frac{\text{number of } (\hat{t} \leq t)}{N}$$

For hard failure-time distribution, the desired values of $F_{T|H}(t)$ can be easily simulated from WeiBull distribution with parameters $\Psi$. Then, based on (3.8), we can derive failure-time distribution $F_{T|C}(t)$ of the competing risks model. However, The aforementioned procedure for deriving the failure time distribution depends on the estimation of $\Theta$ and does not incorporate the uncertainty in the estimate of $\Theta$. 

This thesis proposes another method to estimate the failure-time distribution of the competing risk model by modifying the Bayesian approach given in Robinson and Crowder [20]. In the Bayesian framework, model parameters $\Theta$ are considered as random variables and functions of the model parameters, such as $F_T(t)$, are also random variables. The Bayesian posterior distribution functions of soft failure-time $T_s$ and hard failure-time $T_h$ are given by

$$f(F_{T|S}(t)|d) \propto \int_\Theta Pr(D(t,k,m) + \epsilon \leq D_j|\Theta) f(\Theta|d) d\Theta$$

$$f(F_{T|H}(t)|d) \propto \int_\Psi Pr(T_h \leq t|\Psi) f(\Psi|d) d\Psi$$

where $f(\Theta|d)$ and $f(\Psi|d)$ are the joint posterior distribution of $\Theta$ and $\Psi$, respectively. In general, these integrations are analytically intractable, but it can be implemented numerically in the Gibbs sampler. The procedure lists as follows

For a pre-specified set of time values $t_k (k = 1, \ldots, M)$, where $M$ is the total number of time points), the soft failure-time distribution $F_{T|S}(t_k)$ is firstly estimated. When the $i$th iteration of the Gibbs sampling procedure is executed, the Gibbs sampler simulates a value for each of the model parameters at the $i$th iteration. Denote their values as $\Theta^{(i)} = (k_1^{(i)}, \mu_m^{(i)}, \sigma_m^{2(i)}, \sigma_e^{2(i)})$, and a value for $F_{T|S}(t_k)$ is also generated, denote by $F_{T|S}^{(i)}(t_k)$. To simulate the value $F_{T|S}^{(i)}(t_k)$, $n$ values of $m$ according to $m \sim N(\mu_m^{(i)}, \sigma_m^{2(i)})$ for large number of $n$ are randomly generated. They can be denoted by $m_1^{(i)}, \ldots, m_n^{(i)}$. Consider the $j$th value $m_n^{(i)}$, we can calculate the expected response value at time $t_k$, denoted by $D(t_k, k_1^{(i)}, m_j^{(i)})$. Then, $F_{T|S,j}^{(i)}(t_k) = \Phi \left[ \frac{D_j-D(t_k,k_1^{(i)},m_j^{(i)})}{\sigma_e^{(i)}} \right]$. A simulated value for $F_{T|S}^{(i)}(t_k)$ is obtained by averaging $F_{T|S,1}^{(i)}(t_k), \ldots, F_{T|S,n}^{(i)}(t_k)$.

This procedure can be easily modified to simulate the hard failure-time distribution $F_{T|H}(t_k)$ and and failure-time distribution $F_{T|C}(t_k)$ for the competing risk model. At $i$th iteration of the Gibbs sampling, $\eta^{(i)}$, $k_2^{(i)}$, $l^{(i)}$ are simulated. At time $t_k$, we can simulate a value of $F_{T|H}^{(i)}(t_k)$ from Weibull distribution. Finally, based on (3.8), one value of $F_{T|C}(t_k)$
for $i$th iteration at time $t_k$ can be derived. After the Gibbs sampling terminates, we obtained three random samples for $F_{T|S}(t_k)$, $F_{T|H}(t_k)$, and $F_{T|C}(t_k)$ from their posterior distributions respectively. Posterior inference for $F_{T|S}(t_k)$, $F_{T|H}(t_k)$, and $F_{T|C}(t_k)$ can then be derived based on the sample statistics.

### 3.2 Bayesian change-point degradation model

This study is largely motivated by the nonlinear degradation paths of PDP observed by an industrial collaborator. A degradation test for six PDPs was executed to assess the reliability of PDP at a constant stress level. Six individual PDP degradation paths, which were also analyzed in Bae et al. [16], consist of measurements of relative luminosity inspected regularly. As shown in Figure 3.1, after a rapid decrease in the light intensity at the initial stage of degradation testing, the paths decrease more slowly. Bae et al. [16] elucidated this phenomenon in terms of incomplete burn-in (or “aging”) in PDP manufacturing process. In general, degradation paths caused by an incomplete burn-in procedure appear in the two-phase pattern. A change-point in the two-phase degradation path represents the transition time between the burn-in phase and the subsequent inherent degradation path. Bae et al. [16] showed that simple linear or log-linear degradation models are inadequate to fit the two-phase degradation of PDPs. By using prior information accumulated from a number of field tests, we propose to employ Bayesian change-point regression models to analyze the two-phase PDP degradation data.

In the following sections, we first introduce a Bayesian change-point regression modeling approach for the degradation path of an individual unit, and then a hierarchical Bayesian change-point modeling approach using the degradation data of multiple units is also presented.
Figure 3.1: Observed degradation paths of six PDPs: (a) relative luminosity and (b) (log) transformed relative luminosity.
3.2.1 Individual Bayesian Modeling Approach

There have been different forms of change-point regression. A simple regression model involving a change-point can be formulated as

\[ y_j = \begin{cases} 
\alpha_1 + \beta_1 t_j + \epsilon_j, & j = 1, \ldots, \tau, \\
\alpha_2 + \beta_2 t_j + \epsilon_j, & j = \tau + 1, \ldots, n,
\end{cases} \]  

(3.9)

where \( y_j \) and \( t_j \) are the response and the independent variable for \( j = 1, 2, \ldots, n \), respectively, \( \tau \) is an unknown index corresponding to the change-point (i.e., the change occurs between \( t_\tau \) and \( t_{\tau+1} \)), and \( n \) is the number of observations for a single degradation path. It is assumed that \( \beta_1 \neq \beta_2 \) to accommodate exactly one change-point in the relationship between \( y_j \) and \( t_j \). The random errors \( \epsilon_j \) are usually assumed to be independent and identically distributed (\( iid \) \( N(0, \sigma^2) \) random variables [5]. One may specify a continuity constraint as \( t_\tau \leq \gamma < t_{\tau+1} \) such that the regression function is continuous, where \( \gamma \) is the intersection point of the two linear lines, i.e.,

\[ \gamma = (\alpha_1 - \alpha_2) / (\beta_2 - \beta_1). \]

ML estimates of the model parameters may be computed using the software package \textit{Segcurve} developed by Beem [40]. Carlin et al. [41] developed a Bayesian approach to the estimation of the model parameters using Gibbs sampling without considering the continuity constraint. This study proposes to apply the change-point regression to model the logarithm of relative luminosity of PDP over time. In practice, it is impossible to observe a genuine change-point in a degradation path since the light luminosity of PDPs cannot be measured continuously. Instead, if it is assumed that the degradation path is continuous, the change-point can be predicted by the estimation of \( \gamma \). Hence, the continuity constraint should be imposed.
An alternative form of change-point regression, which explicitly imposes the continuity constraint, is given by [42]

\[
y_j = \begin{cases} 
\alpha + \beta_1(t_j - \gamma) + \epsilon_j, & t_j \leq \gamma, \\
\alpha + \beta_2(t_j - \gamma) + \epsilon_j, & t_j > \gamma,
\end{cases}
\]  

(3.10)

for \( j = 1, 2, \ldots, n \), where \( \gamma \) is the change-point, and \( \alpha \) is the expected response at the change-point. Muggeo [43] considered another change-point regression model given by

\[
y_j = \begin{cases} 
\alpha t_j + \epsilon_j, & t_j \leq \gamma, \\
\alpha t_j + \beta(t_j - \gamma) + \epsilon_j, & t_j > \gamma,
\end{cases}
\]  

(3.11)

for \( j = 1, 2, \ldots, n \), where \( \alpha \) is the slope of the line before the change-point \( \gamma \), and \( \alpha + \beta \) is the slope of the line after the change-point. The model (3.11) also imposes the continuity constraint, and assumes that the expected response at \( t = 0 \) is zero. Because the relative luminosity at time zero is 100% and the logarithm of the initial relative luminosity is zero, the model (3.11) is suitable for modeling the PDP degradation. As shown in Figure 3.1-(b), the slope of the line before the change point is expected to be less than the slope of the line after the change point, and that both slopes are negative. In order to conveniently incorporate this prior knowledge related to the degradation pattern into the prior distribution of the model parameters, this study re-parameterizes the model (3.11) and proposes a change-point degradation model given by

\[
y_j = g(t_j; \theta) + \epsilon_j = \begin{cases} 
\alpha t_j - \beta t_j + \epsilon_j, & t_j \leq \gamma, \\
\alpha t_j - \beta \gamma + \epsilon_j, & t_j > \gamma,
\end{cases}
\]  

(3.12)

for \( j = 1, 2, \ldots, n \), where \( \theta = (\alpha, \beta, \gamma)' \), \( g(t_j; \theta) \) denotes the expected degradation path (i.e., the mean regression function), and \( \epsilon_j \sim N(0, \sigma^2) \), for \( j = 1, 2, \ldots, n \). The model (3.12) explicitly imposes the continuity constraint and \( g(0; \theta) = 0 \). In this new change-point regression model, \((\alpha - \beta)\) and \( \alpha \) are, respectively, the slope of the linear function before
and after the change-point. Specifying $\alpha < 0$ and $\beta > 0$ will result in the desired degradation pattern if this model is applied to model the PDP degradation. Note that the change-point regression model (3.12) is different from the model in Bae et al. [33], which used the model (3.9).

In this study, the change-point regression model (3.12) is applied to model the observed two-phase degradation of individual PDPs, as well as to predict the residual life distribution of a monitored unit, mainly under Bayesian formula. For the purpose of comparison, an ML algorithm to estimate the model parameters is also developed. Assuming that $\gamma \in [t_\tau, t_{\tau+1})$, we can formulate a constrained nonlinear optimization problem which maximizes the log-likelihood function given by

$$L_\tau(\theta) = -\frac{n}{2} \log(\sigma^2) - \sum_{j=1}^{\tau} \frac{(y_j - \alpha t_j + \beta t_j)^2}{2\sigma^2} - \sum_{j=\tau+1}^{n} \frac{(y_j - \alpha t_j + \beta \gamma)^2}{2\sigma^2},$$

subject to $t_\tau \leq \gamma < t_{\tau+1}$. After considering all the intervals $[t_\tau, t_{\tau+1})$, for $\tau = 1, 2, \ldots, n - 1$, the result with maximum log-likelihood value is selected. In this study, the `fmincon` function in Matlab® is used to solve the constrained nonlinear optimization problem given by Eq. (3.13).

Next, a Bayesian algorithm is developed for the inference of the change-point regression model. Let us denote $\vartheta = (\alpha, \beta)'$. The model (3.12) can be represented as a standard regression form

$$y = X(\gamma) \vartheta + \epsilon,$$

(3.14)
where \( y, X(\gamma), \) and \( \epsilon \) are defined as

\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_n
\end{bmatrix}, \quad
\begin{bmatrix}
t_1 & -t_1 \\
\vdots & \vdots \\
t_\tau & -t_\tau \\
-t_\tau & t_\tau + 1 & -\gamma \\
\vdots & \vdots \\
t_n & -\gamma
\end{bmatrix}, \quad \text{and}
\begin{bmatrix}
\epsilon_1 \\
\vdots \\
\epsilon_n
\end{bmatrix},
\]

respectively, where \( t_\tau \leq \gamma \) and \( t_{\tau+1} > \gamma \).

In a Bayesian approach, the change-point regression model requires prior distributions for \( \theta \) and \( \sigma^2 \). In existing studies related to Bayesian change-point regression analysis, independent priors have usually been employed for the change-point \( \gamma \) (or to the index \( \tau \)) and other regression parameters [41, 42]. When analyzing the change-point regression model (3.9), for instance, Carlin et al. [41] assumed a multivariate normal prior to \((\alpha_1, \beta_1, \alpha_2, \beta_2)^\prime\) and a discrete uniform prior to the index \( \tau \), which is independent to that multivariate normal prior. However, it is necessary to consider the dependence between the change-point and other regression parameters if the purpose of the change-point regression analysis is to derive the failure time distribution or the residual life distribution. Note that in the change-point regression model (3.12), \((\alpha - \beta)\) and \(\alpha\) are, respectively, the slope of the linear function before and after the change-point. In application to luminosity degradation in displays, we expect that \(\alpha - \beta < \alpha < 0\), i.e., \(\alpha < 0\) and \(\beta > 0\), as shown in the PDP degradation data in Figure 3.1-(b). In addition, this study considers the case where there exists a change-point in a time interval \((0, \phi)\). Therefore, a truncated trivariate normal prior is assumed to \(\theta\), i.e.,

\[
\theta = (\alpha, \beta, \gamma) \sim N_3(\mu, \Sigma) I_{\{\alpha < 0, \beta > 0, 0 < \gamma < \phi\}},
\]
where $N_3(\mu, \Sigma)$ denotes the trivariate normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$, and $I_{\{\cdot\}}$ is the indicator function. $\sigma^2$ is assumed to have an $IG$ prior with parameters $a_\sigma$ and $b_\sigma$, i.e., $\sigma^2 \sim IG(a_\sigma, b_\sigma)$, equivalently, $\sigma^{-2} \sim G(a_\sigma, b_\sigma)$.

Given the observed data $d$, the joint posterior distribution of the model parameters can be derived according to the Bayes’ formula

$$f(\theta, \sigma^2 | d) \propto f(d | \theta, \sigma^2) f(\theta) f(\sigma^2)$$

$$\propto \sigma^{-n} \exp \left[ -\frac{(y - X(\gamma)\vartheta)' (y - X(\gamma)\vartheta)}{2\sigma^2} \right] \exp \left[ -\frac{(\theta - \mu)' \Sigma^{-1} (\theta - \mu)}{2} \right]$$

$$\times \left( \sigma^2 \right)^{-a_\sigma-1} \exp \left( -\frac{b_\sigma}{\sigma^2} \right) I_{\{ \alpha < 0, \beta > 0 \}},$$

(3.15)

where $f(\theta, \sigma^2 | d)$ is the joint posterior distribution, $f(d | \theta, \sigma^2)$ is the likelihood, and $f(\theta)$ and $f(\sigma^2)$ are the prior distributions for $\theta$ and $\sigma^2$, respectively.

Because of the high-dimensional integration involved, it may be analytically intractable to derive the marginal posterior distributions for parameters. Gibbs sampling algorithm can be also applied in the inference of the model (3.12). Each iteration of the algorithm consists of the following three steps:

(a) Sample $\vartheta$ from the conditional posterior distribution $f(\vartheta | \sigma^2, \gamma, d)$. It can be shown that

$$f(\vartheta | \sigma^2, \gamma, d) \propto \exp \left[ -\frac{(y - X(\gamma)\vartheta)' (y - X(\gamma)\vartheta)}{2\sigma^2} \right]$$

$$\times \exp \left[ -\frac{(\vartheta - \mu_{\vartheta \gamma})' \Sigma_{\vartheta \gamma}^{-1} (\vartheta - \mu_{\vartheta \gamma})}{2} \right] I_{\{ \alpha < 0, \beta > 0 \}},$$

$$\propto \exp \left[ -\frac{(\vartheta - \tilde{\vartheta})' \left( \sigma^{-2}X(\gamma)'X(\gamma) + \Sigma_{\vartheta \gamma}^{-1} \right) (\vartheta - \tilde{\vartheta})}{2} \right] I_{\{ \alpha < 0, \beta > 0 \}},$$

where $\tilde{\vartheta} = \left( \sigma^{-2}X(\gamma)'X(\gamma) + \Sigma_{\vartheta \gamma}^{-1} \right)^{-1} \left( \sigma^{-2}X(\gamma)'y + \Sigma_{\vartheta \gamma}^{-1} \mu_{\vartheta \gamma} \right)$. Because $\theta = (\alpha, \beta, \gamma)'$ is assumed to have a truncated trivariate normal prior, the prior distribution of $\vartheta = (\alpha, \beta)'$ conditional on $\gamma$ is then truncated bivariate normal denoted by...
$N_2(\mu_{\theta y}, \Sigma_{\theta y})I_{(\alpha < 0, \beta > 0)}$, where the mean vector $\mu_{\theta y}$ and the covariance matrix $\Sigma_{\theta y}$ are given by

$$\mu_{\theta y} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} (\gamma - \mu_3),$$

and

$$\Sigma_{\theta y} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} - \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{31} \\ \Sigma_{32} \end{bmatrix},$$

respectively. Herein $\mu_i$ denotes the $i$th element of the vector $\mu$, and $\Sigma_{ij}$ is the $ij$th element of the matrix $\Sigma$. The conditional posterior distribution of $\theta$ is, therefore, truncated bivariate normal with mean vector $\bar{\theta}$ and covariance matrix

$$\left(\sigma^{-2}X(\gamma)'X(\gamma) + \Sigma_{\theta y}^{-1}\right)^{-1},$$

restricted to the region $\{\theta = (\alpha, \beta)' : \alpha < 0, \beta > 0\}$.

(b) Sample $\gamma$ from its conditional posterior distribution $f(\gamma|\theta, \sigma^2, d)$. The prior distribution of $\gamma$ conditional on $\theta$ is truncated normal with mean $\mu_{\gamma|\theta}$ and variance $\sigma_{\gamma|\theta}^2$, which are given by

$$\mu_{\gamma|\theta} = \mu_3 + \begin{bmatrix} \Sigma_{31} \\ \Sigma_{32} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} (\bar{\theta} - [\mu_1, \mu_2]'),$$

and

$$\sigma_{\gamma|\theta}^2 = \Sigma_{33} - \begin{bmatrix} \Sigma_{13} \\ \Sigma_{23} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{31} \\ \Sigma_{32} \end{bmatrix},$$

respectively, restricted to $\{\gamma : 0 < \gamma < \phi\}$. The conditional posterior distribution $f(\gamma|\theta, \sigma^2, d)$ is then given by

$$f(\gamma|\theta, \sigma^2, d) \propto \exp \left[ -\frac{(\gamma - X(\gamma)\theta)'(\gamma - X(\gamma)\theta)}{2\sigma^2} \right] \exp \left[ -\frac{(\gamma - \mu_{\gamma|\theta})^2}{2\sigma_{\gamma|\theta}^2} \right] I_{(0 < \gamma < \phi)}.$$
into \( m \) sub-intervals \((t_1, t_2), [t_2, t_3), \ldots, [t_{m-1}, t_m), [t_m, \phi)\). It can be shown that in the
\( k \)th sub-interval denoted by \( I_k \), \( f(\gamma|\vartheta, \sigma^2, d) \) has the form of
\[
f(\gamma|\vartheta, \sigma^2, d) = \frac{1}{C} \exp \left( -\frac{a_k \gamma^2 + b_k \gamma + d_k}{2} \right) \mathbb{1}_{(\gamma \in I_k)},
\]
where
\[
a_k = (n - k) \beta \sigma^{-2} + \sigma_{\gamma \vartheta}^{-2},
\]
\[
b_k = 2 \beta \sigma^{-2} \sum_{i=k+1}^{n} (y_i - \alpha t_i) - \sigma_{\gamma \vartheta}^{-2} \mu_{\gamma \vartheta},
\]
\[
d_k = \sigma^{-2} \left[ \sum_{i=1}^{k} (y_i - \alpha t_i + \beta t_i)^2 + \sum_{i=k+1}^{n} (y_i - \alpha t_i)^2 \right] + \sigma_{\gamma \vartheta}^{-2} \mu_{\gamma \vartheta}^2,
\]
\[
C = \sum_{k=1}^{m} c_k, \text{ with } c_k = \int_{l_k}^{u_k} \exp \left( -\frac{a_k \gamma^2 + b_k \gamma + d_k}{2} \right) d\gamma.
\]
Herein \( l_k \) and \( u_k \) denote, respectively, the lower limit and the upper limit of the \( k \)th sub-interval. \( c_k \) can be easily evaluated numerically. To sample a value for \( \gamma \), we first choose the sub-interval according to the discrete probabilities \( p_k = c_k / C \), for \( k = 1, 2, \ldots, m \). If the \( k \)th sub-interval is chosen, the new value for \( \gamma \) can then be sampled according to the truncated normal distribution with mean \( -b_k / 2a_k \) and variance \( 1 / a_k \), restricted to the \( k \)th sub-interval \( I_k \).

(c) Sample \( \sigma^2 \) from the conditional posterior distribution \( f(\sigma^2|\theta, d) \), which is
\[
f(\sigma^2|\theta, d) \propto (\sigma^2)^{-\frac{a}{2}} \exp \left[ -\frac{(y - X(\gamma)\vartheta)' (y - X(\gamma)\vartheta)}{2\sigma^2} \right] \times (\sigma^2)^{-a-1} \exp \left( -\frac{b_{\sigma}}{\sigma^2} \right)
\]
\[
\propto (\sigma^2)^{-\frac{a}{2}} \exp \left[ -\frac{b_{\sigma} + 0.5 (y - X(\gamma)\vartheta)' (y - X(\gamma)\vartheta)}{\sigma^2} \right]
\]
\[
\sim IG \left( a_{\sigma} + n/2, b_{\sigma} + 0.5 (y - X(\gamma)\vartheta)' (y - X(\gamma)\vartheta) \right).
\]
Each iteration of Gibbs sampling draws values of model parameters from their conditional posterior distributions. Draws for functions of the model parameters, such as the expected degradation path \( g(t_j; \theta) \), can also be obtained in Gibbs sampling. For example, in each
iteration of Gibbs sampling, a value for \( g(t_j; \theta) \) can be computed using the \( \theta \) simulated in the same iteration. The change-point regression model and the Gibbs sampling algorithm suggested in this section are used to analyze each observed degradation path individually and also used to predict the residual life distribution of a monitored unit, which will be discussed in Chapter 4.

3.2.2 Hierarchical Bayesian Modeling Approach for Multiple Units

This section describes a hierarchical Bayesian change-point model for the degradation paths of multiple units. In the hierarchical Bayesian model, the \( j \)th response measured on the \( i \)th individual unit is modeled by

\[
y_{i,j} = \begin{cases} 
\alpha_i \tau_{i,j} - \beta_i \tau_{i,j} + \epsilon_{i,j}, & j = 1, \ldots, \tau_i, \\
\alpha_i \tau_{i,j} - \beta_i \gamma_i + \epsilon_{i,j}, & j = \tau_i + 1, \ldots, n_i,
\end{cases}
\]

(3.16)

for \( i = 1, 2, \ldots, N \), where \( y_{i,j} \) is the \( j \)th response on the \( i \)th unit, \( \tau_{i,j} \) is the \( j \)th measurement time on the \( i \)th unit, \( \tau_i \) is the number of measurements on the \( i \)th unit, and \( N \) is number of units tested. The measurement errors \( \epsilon_{i,j} \) are again assumed to be iid normal with mean zero and variance \( \sigma^2 \) [5].

A three-stage hierarchical Bayesian degradation model for multiple units is proposed in this study. The first stage of the model is the change-point regression given by Eq. (3.16). The first-stage model can be equivalently written as,

\[
y_i = X_i(\gamma_i) \theta_i + \epsilon_i, \tag{3.17}
\]
where $y_i$, $X_i(\gamma_i)$, $\vartheta_i$, and $\epsilon_i$ are defined as

$$y_i = \begin{bmatrix} y_{i,1} \\ \vdots \\ y_{i,n_i} \end{bmatrix}, \quad \vartheta_i = \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}, \quad X_i(\gamma_i) = \begin{bmatrix} t_{i,1} & -t_{i,1} \\ \vdots & \vdots \\ t_{i,r_i} & -t_{i,r_i} \\ t_{i,r_i+1} & -\gamma_i \\ \vdots & \vdots \\ t_{i,n_i} & -\gamma_i \end{bmatrix}, \quad \text{and} \quad \epsilon_i = \begin{bmatrix} \epsilon_{i,1} \\ \vdots \\ \epsilon_{i,n_i} \end{bmatrix}$$

respectively, for $i = 1, 2, \ldots, N$.

In the second stage, all the coefficient vectors $\theta_i = (\alpha_i, \beta_i, \gamma_i)'$ are assumed to come from a common truncated trivariate normal distribution, i.e.,

$$\theta_i \sim N_3(\mu, \Sigma)I_{\{\alpha_i < 0, \beta_i > 0, 0 < \gamma_i < \phi\}}, \quad i = 1, 2, \ldots, N. \quad (3.18)$$

The third stage of the model specifies prior distributions to $\sigma^2$, $\mu$ and $\Sigma$. Prior distributions play a critical role in the Bayesian analysis. Conjugate priors, if available, may simplify the posterior computation. However, if prior information is not available, noninformative priors are desired. Note that the conjugate prior for $\sigma^2$ is inverse-Gamma. The conditional conjugate prior for the mean vector and covariance matrix of non-truncated multivariate normal distribution is the independent normal-Wishart prior, i.e., a multivariate normal prior for the mean vector and an inverse-Wishart (IW) prior for the covariance matrix [41]. The inverse-Wishart distribution of a $\nu \times \nu$ symmetric positive-definite random matrix $Z$ has the following probability density function

$$f(Z|\rho, S) = \left[2^{\nu/2} \pi^{(\nu-1)/4} \prod_{i=1}^{\nu} \Gamma\left(\frac{\rho + 1 - i}{2}\right)\right]^{-1} |S|^{\rho/2} |Z|^{-(\nu+\rho+1)/2} \exp\left[-\frac{1}{2} \text{tr}(SZ^{-1})\right],$$

where $\rho$ is the degree of freedom, and $S$ is a $\nu \times \nu$ symmetric positive-definite scale matrix. The inverse-Wishart prior for $\Sigma$ is computationally convenient, however, it is usually difficult to choose the parameters $\rho$ and $S$ if a noninformative prior is needed for $\Sigma$.\[\]
Letting $\rho$ be equal to $\nu + 1$ and $S$ be a $\nu \times \nu$ identity matrix $I_\nu$, is equivalent to putting the noninformative uniform priors defined over the interval $(-1, 1)$ to the correlation terms (i.e., the off-diagonal elements of the matrix $\Sigma$) [44]. The $IW(\nu + 1, I_\nu)$ prior may be reasonable for correlation estimation but is quite constraining on the variance terms (i.e., the diagonal terms of $\Sigma$) [45]. In this study, we adopt the scaled inverse-Wishart prior discussed in [44, 45]. The covariance matrix $\Sigma$ is decomposed into variance and correlation components as

$$\Sigma = \Delta Q \Delta,$$

where the diagonal matrix $\Delta = \text{diag}(\delta_1, \ldots, \delta_\nu)$ with $\delta_i > 0$, and $Q$ is a $\nu \times \nu$ symmetric positive-definite matrix. Then we assume the $IW(\rho, S)$ prior distribution to the matrix $Q$, and independent Gamma priors, $G(\alpha, \beta)$, to $\delta_i$, $i = 1, 2, \ldots, \nu$. For example, assuming the $IW(\nu + 1, I_\nu)$ prior for $Q$ and the $G(1, 0.001)$ prior for $\delta_i$'s can result in a proper and noninformative prior for the covariance matrix $\Sigma$.

Applying the Bayes’ theorem, we can derive the joint posterior distribution for all the model parameters conditioning on the data $d$ as follows

$$f(\theta_1, \ldots, \theta_N, \mu, \delta_1, \ldots, \delta_\nu, \sigma^2 | d) \propto$$

$$\prod_{i=1}^{N} \frac{1}{\sigma^\nu} \exp \left[ -\frac{(y_i - \mu)^T \gamma_i (y_i - \mu)}{2\sigma^2} \right] I_{(\alpha_i < 0, \beta_i > 0, 0 < \gamma_i < \phi)}$$

$$\times [P(\mu, \Sigma)]^{-N/2} |\Delta|^{-N/2} |Q|^{-N/2} |\Delta|^{-N/2} \exp \left[ -\frac{\sum_{i=1}^{N} (\theta_i - \mu)^T \Delta^{-1} Q^{-1} \Delta^{-1} (\theta_i - \mu)}{2} \right]$$

$$\times \exp \left[ -\frac{(\mu - \xi)^T C^{-1} \mu - \xi}{2} \right] \times |Q|^{-(\nu + \rho + 1)/2} \exp \left[ -\frac{\text{tr}(\Sigma Q^{-1})}{2} \right] \times \prod_{i=1}^{\nu} \delta_i^{\alpha_i - 1} \exp \left( -b_i \delta_i \right)$$

$$\times (\sigma^2)^{-\left(\alpha_0 + 1\right)} \exp \left( -\frac{b_0 \sigma^2}{2} \right),$$

where $\xi$ and $C$ are, respectively, the mean vector and the covariance matrix of the bivariate normal prior distribution for $\mu$, and $[P(\mu, \Sigma)]$ is a normalizing constant, attributable to the truncation in the second-stage model (3.18).

An MCMC algorithm implementing Gibbs sampling is developed to fit the three-stage hierarchical Bayesian degradation model. Each iteration of the algorithm consists of the following steps:
(a) Sample $\mathbf{\theta}_i = (\alpha_i, \beta_i)'$ from the conditional posterior distribution $f(\mathbf{\theta}_i|\sigma^2, \gamma_i, \mu, \Sigma, \mathbf{d})$ for $i = 1, 2, \ldots, N$. Step (a) of the sampling algorithm described in Section 3.2.1 can be applied here.

(b) Sample $\gamma_i$ from the conditional posterior distribution $f(\gamma_i|\sigma^2, \mathbf{\theta}_i, \mu, \Sigma, \mathbf{d})$ for $i = 1, 2, \ldots, N$. Step (b) of the sampling algorithm described in Section 3.2.1 can be applied here.

(c) Sample $\sigma^2$ from the conditional posterior distribution $f(\sigma^2|\theta_1, \ldots, \theta_N, \mathbf{d})$, which is

$$IG\left(a_{\sigma} + \sum_{i=1}^{N} n_i, b_{\sigma} + \sum_{i=1}^{N} \left(\frac{(y_i - \mathbf{X}_i(\gamma_i)\theta_i)'(y_i - \mathbf{X}_i(\gamma_i)\theta_i)}{2}\right)\right).$$

(d) Sample $\mu$ from its conditional posterior distribution

$$f(\mu|\theta_1, \ldots, \theta_N, \Sigma, \mathbf{d}) \propto [P(\mu, \Sigma)]^{-N} \exp\left[\frac{-\sum_{i=1}^{N} (\theta_i - \mu)'\Sigma^{-1}(\theta_i - \mu) + (\mu - \xi)'C^{-1}(-\xi)}{2}\right].$$

Because of the truncation used in the second-stage model (3.18), the $N(\xi, \mathbf{C})$ distribution is no longer the conditional conjugate prior for $\mu$. Therefore, the conditional posterior distribution $f(\mu|\theta_1, \ldots, \theta_N, \Sigma, \mathbf{d})$ is not a standard distribution. However, Griffiths [46] developed a convenient sampling algorithm for the parameters of a truncated multivariate normal distribution. For each $\theta_i$ vector that follows the truncated $N_3(\mu, \Sigma)$ distribution, a latent vector $\theta_i^*$ can be computed using the inverse-cdf method such that the latent vector $\theta_i^*$ follows the non-truncated $N_3(\mu, \Sigma)$ distribution. Then, it is routine to sample $\mu$ and $\Sigma$ conditional on the $\theta_i^*$ vectors. Conditional on the $\theta_i^*$ vectors, it can be shown that the conditional posterior distribution of $\mu$ is trivariate normal with mean vector

$$(N\Sigma^{-1} + \mathbf{C}^{-1})^{-1} \left(\Sigma^{-1} \sum_{i=1}^{N} \theta_i^* + \mathbf{C}^{-1} \xi\right)$$

and covariance matrix $(N\Sigma^{-1} + \mathbf{C}^{-1})^{-1}$. 
(e) Sample $Q$ from its conditional posterior distribution given by

$$f(Q|\theta^*_1, \ldots, \theta^*_N, \mu, \delta_1, \ldots, \delta_i, d) \propto |Q|^{-N/2} \exp \left[ -\sum_{i=1}^{N} (\theta^*_i - \mu') \Delta^{-1} Q^{-1} \Delta^{-1} (\theta^*_i - \mu) \right] |Q|^{-\nu + \rho + 1} \exp \left[ -\text{tr}(SQ^{-1}) \right],$$

which is

$$IW \left( \rho + N, S + \Delta^{-1} \left( \sum_{i=1}^{N} (\theta^*_i - \mu)(\theta^*_i - \mu)' \right) \Delta^{-1} \right).$$

(f) Sample $\delta_i$ from the conditional posterior distribution given by

$$f(\delta_i|\theta^*_1, \ldots, \theta^*_N, \mu, \delta_j; j \neq i, d) \propto \delta_i^{-N+a_i} \exp \left[ -\left( b_\delta \delta_i + \frac{V_{ij}}{2\delta_i} + \frac{1}{\delta_i} \sum_{j \neq i} (Q^{-1})_{ij} V_{ij} / \delta_i \right) \right],$$

for $i = 1, 2, \ldots, N$, where $(Q^{-1})_{ij}$ is the $i,j$th element of the matrix $Q^{-1}$, and $V_{ij}$ is the $i,j$th element of the matrix $\left( \sum_{i=1}^{N} (\theta^*_i - \mu)(\theta^*_i - \mu)' \right)$. Slice sampling can be applied here to sample $\delta_i$ from its conditional posterior distribution [47].

This section proposed a three-stage hierarchical Bayesian change-point degradation model to analyze the degradation paths of multiple units, and a Gibbs sampling algorithm was developed for the inference of the model parameters.

### 3.2.3 Prediction of Residual Life Distribution

This section describes the prediction of residual life distribution of a monitored unit using the Bayesian method. At a given inspection or measurement time, $t_c$, the objective is to determine the distribution of the time until the degradation signal falls below the failure threshold $y^*$. Let $T_R$ denote the residual useful life of the monitored unit at time $t_c$. When the expected degradation path is a decreasing function of time under given model parameters, the cumulative distribution function of $T_R$ is defined as

$$F_R(t|\theta, \sigma^2) = \text{Pr} \left( T_R \leq t | \theta, \sigma^2 \right) = \text{Pr} \left( g(t + t_c; \theta) + \epsilon \leq y^* \right) = \Phi \left( \frac{y^* - g(t + t_c; \theta)}{\sigma} \right).$$

In the Bayesian framework, the parameters are treated as random variables. Let $d_c$ denote degradation measurements observed on the monitored unit by time $t_c$. Uncertainty in the
model parameters is expressed by the joint posterior distribution \( f(\theta, \sigma^2|d_c) \), which combines information from two sources: (1) the joint prior distribution summarized from a unit’s population; and (2) real-time degradation data of the monitored unit. The derivation of \( f(\theta, \sigma^2|d_c) \) is given by Eq.(3.15). Then the predictive posterior distribution of \( F_R(t) \) is defined as

\[
f(F_R(t)|d_c) = \int \Pr (g(t + t_c; \theta) + \epsilon \leq y^*) f(\theta, \sigma^2|d_c) d\theta d\sigma^2
= \int \Phi \left( \frac{y^* - g(t + t_c; \theta)}{\sigma} \right) f(\theta, \sigma^2|d_c) d\theta d\sigma^2. \tag{3.20}
\]

The integral required in Eq.(3.19) has no closed-form solution. Estimation of \( f(F_R(t)|d_c) \), however, can be conveniently embedded in the Gibbs sampling algorithm presented in Section 3.2.1. At a given inspection time \( t_c \), we pre-specify a set of \( M \) residual life values \( t_m, m = 1, 2, \ldots, M \), for which \( f(F_T(t_m)|d_c) \) will be estimated. In each iteration of the Gibbs sampling algorithm discussed in Section 3.2.1, Steps (a)-(c) draw values for \( \theta \) and \( \sigma^2 \) from their conditional posterior. Then Step (d) is added to compute the values \( F_T(t_m|\theta, \sigma^2) \), for \( m = 1, 2, \ldots, M \), according to Eq.(3.19) and using \( \theta \) and \( \sigma^2 \) values obtained at the same iteration. Upon the completion of the Gibbs sampling algorithm, a random sample for \( F_R(t_m) \) is simulated from \( f(F_R(t_m)|d_c) \). Sample statistics can then be used to construct point and interval estimates for \( F_R(t_m) \).
4 Case Study

4.1 LEDs Example

In this section, an experimental dataset reported by Zhao and Elsayed [29] is used to validate the competing risks model. In their experiment, they assumed that LEDs failed can be either due to degradation failure (soft failure) or sudden failure (hard failure) which is the result of solder and heat sink [48]. This is a typical competing risks problem which can be appropriately described by the proposed model in this thesis. Moreover, accelerated life testing was also considered in their experiment. The light density of testing LEDs were measured under two different accelerated stress levels: 40mA and 35mA. The experimental data from 40mA and 35mA conditions are used to estimate the parameters in the proposed models. Based on the validated model, the failure-time distribution is predicted at 28mA which is the normal operation condition. In order to fit the proposed models, the inverse LEDs light density [29] is transformed into the original decreasing degradation data and then taken logarithm. Some experimental data shown in Figure 4.1 are used to present the degradation paths under 40mA and 35mA operating conditions in the competing risks problem discussed above.
4.1.1 Accelerated Degradation Model

First, the Bayesian hierarchical degradation model (3.5) introduced in section 3.1.2 is fitted by the degradation dataset under 40mA and 35mA levels. Noninformative prior distributions are assigned to model parameters:

\[ k_1 \sim N(0, 10^6), \mu_m \sim N(0, 10^6), \sigma_m^2 \sim IG(0.001, 0.001), \sigma_e^2 \sim IG(0.001, 0.001) \]

Bayesian posterior inference can be implemented by WinBUGS, a special software that can complete Gibbs sampling procedures. Next, for the purpose of comparison, the mixed-effects model (3.4) is also applied to analyze the same degradation dataset. The MLEs of model parameters are derived by \textit{nlme} function in R. Table 4.1 and Table 4.2 compare the estimated results obtained by Bayesian hierarchical method and ML method. The results indicate that the two methods provide very similar results.

Figure 4.1: LEDs degradation data under 40mA and 35mA stress levels
Table 4.1: MLEs for the parameters in degradation model

<table>
<thead>
<tr>
<th>Subject</th>
<th>Mixed-effects Model</th>
<th>95% interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{k}_1$</td>
<td>$5.158 \times 10^{-11}$</td>
<td>$(-2.556 \times 10^{-10}, 3.588 \times 10^{-10})$</td>
</tr>
<tr>
<td>$\mu_m$</td>
<td>5.063</td>
<td>(3.420, 6.706)</td>
</tr>
<tr>
<td>$\sigma^2_m$</td>
<td>0.0082</td>
<td>(0.0048, 0.0142)</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>0.0296</td>
<td>(0.0235, 0.0373)</td>
</tr>
</tbody>
</table>

Table 4.2: Bayesian posterior medians for the parameters in degradation model

<table>
<thead>
<tr>
<th>Subject</th>
<th>Bayesian Hierarchical Model</th>
<th>95% interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{k}_1$</td>
<td>$6.248 \times 10^{-11}$</td>
<td>$(4.840 \times 10^{-13}, 4.587 \times 10^{-10})$</td>
</tr>
<tr>
<td>$\mu_m$</td>
<td>5.006</td>
<td>(4.454, 6.350)</td>
</tr>
<tr>
<td>$\sigma^2_m$</td>
<td>0.0091</td>
<td>(0.0059, 0.0147)</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>0.0296</td>
<td>(0.0247, 0.0361)</td>
</tr>
</tbody>
</table>

### 4.1.2 Competing Risks Model

Now, the Bayesian hierarchical competing risks model (3.6) is analyzed using both the degradation data and hard failure-times. For hard failures, the following noninformative prior distributions are assigned to the parameters of Weibull distribution:

$$k_2 \sim N(0, 10^6), \ l \sim N(0, 10^6), \ \eta \sim IG(0.001, 0.001)$$

Table 4.3 summarizes the posterior medians and 95% confidence intervals for all the model parameters.
Table 4.3: Bayesian inference for the competing risks model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Posterior median</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_m$</td>
<td>5.006</td>
<td>(4.454, 6.350)</td>
</tr>
<tr>
<td>$\sigma^2_m$</td>
<td>0.0091</td>
<td>(0.0059, 0.0147)</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>0.0296</td>
<td>(0.0247, 0.0361)</td>
</tr>
<tr>
<td>$k_1$</td>
<td>$6.248 \times 10^{-11}$</td>
<td>$(4.840 \times 10^{-13}, 4.587 \times 10^{-10})$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>$6.469 \times 10^5$</td>
<td>$(1316, 2.525 \times 10^{11})$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>4.012</td>
<td>(1.904, 7.398)</td>
</tr>
<tr>
<td>$l$</td>
<td>2.038</td>
<td>(0.345, 5.486)</td>
</tr>
</tbody>
</table>

4.1.3 Failure-time Distribution

After obtaining MLEs and Bayesian posterior statistics for the proposed model parameters, the two simulation based methods introduced in Section 3.1.4 are used to predict failure-times distributions. For the purpose of comparison, soft failure-time distributions derived by different methods are plotted together in Figure 4.2. Under each stress level of 40mA, 35mA, and 28mA, Figure 4.2 indicates that the predicted soft failure-time distribution curves are very similar and partially overlapping. For competing risks problem, Figure 4.3 plots the posterior median of the failure-time distributions $F_{T|S}(t_k), F_{T|H}(t_k),$ and $F_{T|C}(t_k)$. When sampling the failure-time distribution in procedures discussed in Section 3.1.4, the number of simulated degradation parameters, $m$, is selected to be 2000.
Figure 4.2: Comparison of soft failure-time distributions
Figure 4.3: Bayesian posterior median of the failure-time distributions $F_{TS}(t_k)$, $F_{TH}(t_k)$, and $F_{TC}(t_k)$
4.2 PDP Example

In this section, the PDP degradation data shown in Figure 3.1 are used to illustrate the application of the models and methods proposed in this study. In the change-point degradation models, the degradation measurement $y$ is the logarithm of relative luminosity, and $t$ is measurement time in thousand hours.

4.2.1 Individual Degradation Modeling

First, we check whether there is a change-point in the degradation path. We individually fit the simple two-phase degradation model (3.12) to the (log) transformed relative luminosity of the six PDP degradation paths. The parameters of the two-phase model are estimated by the ML algorithm proposed in Section 3.2.1. MLEs of the two-phase model parameters are summarized in Table 4.4. For the purpose of comparison, the simple linear degradation model is applied to analyze the (log) transformed relative luminosity of the six PDP degradation paths. The simple linear degradation model is

\[
y_j = h(t_j; \zeta) + \epsilon_j = \zeta t_j + \epsilon_j, \quad j = 1, 2, \ldots, n,
\]

and $\epsilon_j \sim N(0, \sigma_h^2)$, where $h(t, \zeta) = \zeta t$ denotes the expected degradation path of the (log) relative luminosity. MLEs of the linear regression model parameters, $\hat{\zeta}$ and $\hat{\sigma}_h^2$, are also listed in Table 4.4. The likelihood ratio test is employed to compare the simple linear degradation model (the null model) and the change-point degradation model (the alternative model) and yields $p$-value $\approx 0$ for all the six PDP degradation paths, strongly supporting the two-phase degradation model.

Next, we fit the individual change-point degradation model to the six PDPs using the Bayesian approach described in Section 3.2.1. To reflect an absence of prior knowledge, noninformative prior distributions are used. Especially, we assign $\mu = (0, 0, 0)'$, $\Sigma = \text{diag}(10^6, 10^6, 10^6)$, $a_\sigma = 1$ and $b_\sigma = 0.001$. In addition, we assume $\phi = t_n$, which
Table 4.4: Maximum likelihood estimates for the parameters in change-point degradation model (3.12) and the simple linear degradation model (4.1).

<table>
<thead>
<tr>
<th>Subject</th>
<th>Change-point model (3.12)</th>
<th>Linear model (4.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td>1</td>
<td>-0.0584</td>
<td>0.3189</td>
</tr>
<tr>
<td>2</td>
<td>-0.0602</td>
<td>0.4632</td>
</tr>
<tr>
<td>3</td>
<td>-0.0608</td>
<td>0.4088</td>
</tr>
<tr>
<td>4</td>
<td>-0.0620</td>
<td>0.4033</td>
</tr>
<tr>
<td>5</td>
<td>-0.0523</td>
<td>0.1656</td>
</tr>
<tr>
<td>6</td>
<td>-0.0649</td>
<td>0.2689</td>
</tr>
</tbody>
</table>

reflects our prior assumption that there is a change-point in the interval $(0, t_n)$. Posterior simulation involving the Gibbs sampling is implemented using Matlab®. The Gibbs sampling runs for 100,000 iterations and the first 50,000 iterations are discarded. Convergence is monitored and verified by running multiple chains starting from dispersed initial values. Table 4.5 summarizes posterior means and posterior medians of the parameters in the two-phase degradation model, along with the 95% Bayesian confidence intervals (in parentheses). Table 4.4 and Table 4.5 indicate that the Bayesian method and the ML method provide very similar results. This is because noninformative priors are used in the Bayesian analysis and the Bayesian inference is largely based on the data information contained in the likelihood function.
Table 4.5: Bayesian inference for the parameters in the individual change-point degradation model (3.12).

<table>
<thead>
<tr>
<th>Subject</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$ (thousand hours)</th>
<th>$\sigma^2$ ($\times 10^{-4}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>median</td>
<td>mean</td>
<td>median</td>
</tr>
<tr>
<td>(95% interval)</td>
<td></td>
<td></td>
<td>(95% interval)</td>
<td>(95% interval)</td>
</tr>
<tr>
<td>1</td>
<td>-0.05829</td>
<td>-0.05831</td>
<td>0.3189 0.3160</td>
<td>0.5735 0.5694</td>
</tr>
<tr>
<td></td>
<td>(-0.06726, -0.04919)</td>
<td></td>
<td>(0.2556, 0.4076)</td>
<td>(0.4099, 0.7490)</td>
</tr>
<tr>
<td>2</td>
<td>-0.06186</td>
<td>-0.06185</td>
<td>0.4973 0.4855</td>
<td>0.4952 0.5021</td>
</tr>
<tr>
<td></td>
<td>(-0.07264, -0.05170)</td>
<td></td>
<td>(0.4079, 0.6259)</td>
<td>(0.3688, 0.6197)</td>
</tr>
<tr>
<td>3</td>
<td>-0.06035</td>
<td>-0.06037</td>
<td>0.4029 0.4019</td>
<td>0.5940 0.5887</td>
</tr>
<tr>
<td></td>
<td>(-0.06992, -0.05069)</td>
<td></td>
<td>(0.3417, 0.4728)</td>
<td>(0.4795, 0.7208)</td>
</tr>
<tr>
<td>4</td>
<td>-0.06240</td>
<td>-0.06241</td>
<td>0.4108 0.4068</td>
<td>0.5336 0.5351</td>
</tr>
<tr>
<td></td>
<td>(-0.07180, -0.05298)</td>
<td></td>
<td>(0.3407, 0.5061)</td>
<td>(0.4058, 0.6636)</td>
</tr>
<tr>
<td>5</td>
<td>-0.05196</td>
<td>-0.05198</td>
<td>0.1667 0.1646</td>
<td>0.7472 0.7461</td>
</tr>
<tr>
<td></td>
<td>(-0.06105, -0.04283)</td>
<td></td>
<td>(0.1395, 0.2083)</td>
<td>(0.4945, 0.9648)</td>
</tr>
<tr>
<td>6</td>
<td>-0.06378</td>
<td>-0.06401</td>
<td>0.2629 0.2594</td>
<td>0.4945 0.4734</td>
</tr>
<tr>
<td></td>
<td>(-0.07234, -0.05404)</td>
<td></td>
<td>(0.1960, 0.3608)</td>
<td>(0.3116, 0.7337)</td>
</tr>
</tbody>
</table>

4.2.2 Hierarchical Bayesian Degradation Modeling

Next, we analyze the observed degradation data of multiple PDPs together using the hierarchical Bayesian change-point degradation model presented in Section 3.2.2 to derive the failure time distribution of a randomly selected unit from the population. We arbitrarily choose analyze the degradation data of PDPs #1 – #5, PDP #6 is reserved to illustrate the Bayesian updating method for residual life prediction of a single unit in Section 4.2.3. Noninformative priors are assumed in the third stage of the hierarchical
model. Specially, we assign the following noninformative priors:

$$\mu \sim N(0, 10^6 I_3), \quad Q \sim IW(4, I_3), \quad \delta_i \sim G(1, 0.001), \quad i = 1, 2, 3, \text{ and } \sigma^2 \sim IG(1, 0.001),$$

and additionally, we assume $\phi = 5$ thousand hours. Table 4.6 summarizes the posterior means and medians of the model parameters $\mu, \Sigma, \text{ and } \sigma^2$.

Table 4.6: Posterior statistics of the model parameters of the hierarchical Bayesian change-point degradation model (3.17), $\mu, \Sigma, \text{ and } \sigma^2$, when the PDPs #1-#5 are analyzed together.

<table>
<thead>
<tr>
<th>Posterior statistics</th>
<th>Parameters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$</td>
<td>$\Sigma$</td>
<td>$\sigma^2$</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-0.0596</td>
<td>$2.865 \times 10^{-5}$</td>
<td>$-8.589 \times 10^{-5}$</td>
<td>$3.411 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td>0.3554</td>
<td>$-8.589 \times 10^{-5}$</td>
<td>$5.147 \times 10^{-2}$</td>
<td>$-1.588 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td>0.5615</td>
<td>$3.411 \times 10^{-5}$</td>
<td>$-1.588 \times 10^{-3}$</td>
<td>$7.622 \times 10^{-3}$</td>
</tr>
<tr>
<td>median</td>
<td>-0.0596</td>
<td>$5.900 \times 10^{-6}$</td>
<td>$-2.376 \times 10^{-5}$</td>
<td>$1.639 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>0.3552</td>
<td>$-2.376 \times 10^{-5}$</td>
<td>$2.131 \times 10^{-2}$</td>
<td>$-3.070 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>0.5598</td>
<td>$1.639 \times 10^{-6}$</td>
<td>$-3.070 \times 10^{-4}$</td>
<td>$1.414 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

### 4.2.3 Residual Life Distribution of a PDP

This section uses PDP #6 to illustrate the Bayesian updating method for residual life prediction of a single unit. It is necessary to employ informative priors for the parameters $\theta$ and $\gamma$ and to consider the dependence between them because the change-point is involved. The priors for the parameters are summarized from the unit’s population. When the inspection times have not exceeded the change-point of the monitored unit, the estimation of the change-point will be largely based on the prior distributions of the model parameters. In this example, we use the posterior medians for the second-stage parameters obtained from analyzing PDPs #1 – #5 as our prior distributions for $\theta$ and $\gamma$. That is, for
the monitored unit (PDP #6), we assume that the prior distribution of the parameter vector \( \theta \) is the truncated trivariate normal with mean \( \mu \) and covariance matrix \( \Sigma \), where \( \mu, \Sigma \) are the median values given in Table 4.6. For the parameter \( \sigma^2 \), Gebraeel et al. [6] used a fixed value summarized from the unit’s population. In this study, we also fix the value \( \sigma^2 \) to the posterior median shown in Table 4.6 (i.e., a point mass prior for \( \sigma^2 \)). Figure 4.4 shows posterior medians of the residual life distribution along with their 95% confidence bounds at five different inspection times. At each inspection time \( t_k \), we combine the first \( k \) observations of PDP #6 and the prior distributions of the model parameters to predict the residual lifetime of PDP #6. As shown in Figure 4.4, when more inspection data available, the uncertainty on the residual life distribution, which is measured by the width of confidence bounds, is reduced. Figure 4.5 compares posterior medians of the residual life distribution at the five inspection times. Generally, when \( t_k \) increases, that is, when the unit has been operated for longer period, the residual life should decrease. However, in Figure 4.5, we observe that the residual life predicted at \( t_5 = 219 \) hours is shorter than the residual lives predicted at \( t_{10} = 502 \) hours and \( t_{15} = 868 \) hours. This should be caused by the involvement of the change-point in the degradation model. Table 4.4 and Table 4.5 indicate that the change-point occurs at about 500 hours for PDP #6. When the residual life is predicted at \( t_5 = 219 \) hours, the prediction of the change point as well as the degradation in the second phase is largely based on prior information summarized from the unit’s population. Therefore, the prediction of the residual life at \( t_5 = 219 \) hours may be highly biased. Starting from \( t_{10} = 502 \) hours, the estimation of the model parameters including change-point and the prediction of the residual life distribution are improved with observations after the change-point. Figure 4.6 shows the expected degradation paths (posterior medians and 95% confidence bounds) predicted at different inspection times. In this figure, the filled circles are observed data points at a given inspection time, while the open circles are “unobserved” data points at that inspection time. Again, large uncertainty
on the expected degradation path is observed when $t_5 = 219$ hours because of the lack of observations of the second phase of the degradation. With more data points available in the second phase of the degradation, the uncertainty becomes less and the prediction of the expected degradation path is more accurate.
Figure 4.4: Predicted residual life distribution, $F_R(t)$, at five different inspection times, $t_k$: posterior median (solid line) and 95% Bayesian confidence bounds (dash lines).
Figure 4.5: Posterior medians of the residual life distribution, $F_R(t)$, at five different inspection times, $t_k$. 
Figure 4.6: Predicted mean degradation paths at five different inspection times, $t_k$: posterior median (solid line) and 95% Bayesian confidence bounds (dash lines).
5 Conclusion

In this thesis, competing risks problem involving accelerated degradation and catastrophic failures was investigated. The competing risks model was validated by real experimental data which were collected in different operating conditions. For the propose of comparison, both mixed-effects model and Bayesian hierarchical model were employed to fit the accelerated degradation data. ML method and Bayesian approach are used to estimate the unknown model parameters. The comparison between MLEs and Bayesian posterior statistics show that both methods provide very similar results. Moreover, by incorporating both prior knowledge and the data, Bayesian approach is more suitable for the case of lack of information provided in the dataset. By considering hard failure-times which can be regarded as Weibull random variables, another Bayesian hierarchical model was proposed to fit the competing risks problem. Gibbs sampling algorithms were developed for the inference of the parameters in the competing risks models as well as for the prediction of failure-time distributions. The predicted failure-time distributions indicate that 1) Based on the estimated results of ML method and Bayesian approach, the predictions of soft failure-time distributions are very closed; 2) Soft failure are the determinate factor in competing risks; 3) Hard failure-times are usually longer than the soft failure-times since hard failures may have occurred after the degradation path crossed the threshold criterion.

For PDP example, this thesis proposed change-point regression models to characterize the two-phase degradation pattern of products, which is generally caused by remaining defects or contaminants during their manufacturing process. An individual Bayesian model was developed to analyze each unit separately and to predict the residual life of a single monitored unit. A hierarchical Bayesian model was developed to analyze multiple units and used to summarize prior information for the model parameters from the population. And then the prior information was used in the residual life prediction of a
single unit. Through the predicted residual-life distributions at five different inspection times, the prediction of residual-life at $t_5 = 219$ hours is uncertain. At $t_5$, the degradation path is still in the first phase of the degradation and does not cross the change-point. Therefore, no observation in the second phase of degradation was involved. As more observations in the second phase of degradation are considered, the uncertainty becomes less and the prediction of the residual-life is more accurate. In practice, after obtaining the residual-life predictions for operational devices, the maintenance plan can be rearranged and optimized. Thus, the hazards of failures and costs for repairing can be also reduced.
REFERENCES


[34] T. S. Ng, “An application of the EM algorithm to degradation modeling,” *IEEE Transactions on Reliability*, vol. 57, no. 6, pp. 2-13, 2008.


