Statistical Modeling and Predictions Based on Field Data and Dynamic Covariates

Zhibing Xu

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Statistics

Yili Hong, Chair
Inyoung Kim
Marion R. Reynolds, Jr.
William H. Woodall

December 1, 2014
Blacksburg, Virginia

Statistical Modeling and Predictions Based on Field Data and Dynamic Covariates

Zhibing Xu

Abstract

Reliability analysis plays an important role in keeping manufacturers in a competitive position. It can be applied in many areas such as warranty predictions, maintenance scheduling, spare parts provisioning, and risk assessment. This dissertation focuses on statistical modeling and predictions based on lifetime data, degradation data, and recurrent event data. The datasets used in this dissertation come from the field, and have complicated structures. The dissertation consists of three main chapters, in addition to Chapter 1 which is the introduction chapter, and Chapter 5 which is the general conclusion chapter. Chapter 2 consists of the traditional time-to-failure data analysis. We propose a statistical method to address the failure data from an appliance used at home with the consideration of retirement times and delayed reporting time. We also develop a prediction method based on the proposed model. Using the information of retirement-time distribution and delayed reporting time, the predictions are more accurate and useful in the decision making. In Chapter 3, we introduce a nonlinear mixed-effects general path model to incorporate dynamic covariates into degradation data analysis. Dynamic covariates include time-varying environmental variables and usage condition. The shapes of the effect functions of covariates may be constrained to be, for example, monotonically increasing (i.e., higher temperature is likely to cause more damage). Incorporating dynamic covariates with shape restrictions is challenging. A modified alternative algorithm and the corresponding prediction method are proposed. In Chapter 4, we introduce a multi-level trend-renewal process (MTRP) model to describe component-level events in multi-level repairable systems. In particular, we consider two-level repairable systems in which events can occur at the subsystem level, or the component (within the subsystem) level. The main goal is to develop a method for estimation of model parameters and a procedure for prediction of the future replacement events at component level with the consideration of the effects from the subsystem replacement events. To explain unit-to-unit variability, time-dependent covariates as well as random effects are introduced into the heterogeneous MTRP model (HMTRP). A Metropolis-within-Gibbs algorithm is used to estimate the unknown parameters in the HMTRP model. The proposed method is illustrated by a simulated dataset.

Key Words: Calibration, Covariate process, Degradation path, Discrete Fourier transfor-

m, Failure reporting delay, Multi-level repairable systems, Nonhomogeneous Poisson process, Nonlinear mixed-effects model, Organic coatings, Poisson-binomial distribution, Prediction interval, Renewal process, Repairable system, Shape-restricted regression splines, Trend-renewal process, Weibull.

To my family.

Acknowledgement

Foremost, I would like to show my deepest gratitude to my research advisor Dr. Yili Hong for his invaluable support, immense knowledge, guidance and help through my research. With the help from Dr. Hong, my programming skills and statistical understanding were greatly improved during my Ph.D. study. His invaluable advice also gave me a lot of inspiration and helped me finish my dissertation. Without his guidance, this research work would not be completed.

Besides my advisor, I would like to thank my committee members Dr. William Woodall, Dr. Marion Reynolds, and Dr. Inyoung Kim for their feedback, assistance, and guidance to my research. My sincere thanks also go to Dr. Jeffrey Birch for always being helpful. I also would like to thank Dr. Xinwei Deng for his valuable advice. I also would like to show my deep gratitude to Dr. David Kingston for the help he gave me in the Department of Chemistry.

I thank my fellow labmates at Virginia Tech: Dr. Yuanyuan Duan, Dr. Khaled Bedair, Caleb King, Yimeng Xie, and Miao Yuan for their valuable advice in my research. Also, I thank my friends: Dr. Shaobin Liu, Dr. Lei Yu, and Dr. Shengzhi Shao for their support when I met difficulty. I am grateful to Dr. Yangyi Xu for the help in finding a job.

I would like to acknowledge Advanced Research Computing (ARC) at Virginia Tech (URL: http://www.arc.vt.edu). Without the computational resources from ARC, this dissertation would not be finished so fast.

I would like to thank my parents for giving birth to me and their selfless love through my life. I also would like to thank my brother and hope he will recover as soon as possible.

Last but not the least, I would like to show my deepest love to my wife, Mingming and daughter, Christina. Without your love and encouragement, I could never go so far.

Zhibing Xu
Blacksburg, VA, USA
October 20, 2014

Contents

List of Figures

List of Tables

Chapter 1 General Introduction

1.1 Background

Reliability of products is of concern to both manufacturers and consumers. A highly reliable product, such as aircrafts, cars, and computers, can operate stably and safely over time under normal conditions. A manufacturer with highly reliable products always has a stronger competitive position in the global market. To have a better understanding about the reliability of products, reliability analysis is necessary.

1.1.1 Prediction in Reliability Analysis

Statistical prediction is often an important task in the reliability analysis. It can help manufacturers predict the expected lifetime of products, the expected number of future events, and other important information. For example, how long a product can continuously work, how many failures will occur in a certain time interval, how many spare parts should be prepared, what is the estimated cost in the warranty period, and how to optimize the maintenance strategy. Prediction analysis is often based on the currently available data and a model that is used to describe the event process. Nelson (1982), and Meeker and Escobar (1998, Chapter 12) introduced general methods in the prediction of future events such as the naïve plug-in method, approximate pivotal method, and the Monte Carlo simulation method. However, different data structures will require different prediction methods. Predictions based on data with complicated structures are often more challenging.

1.1.2 Field Data in Reliability

Reliability data mainly come from two sources: laboratory test data and field data. Laboratory test data are used to solve specific problems based on well designed experiments. Accelerated test methods are used to obtain reliability information in a short time period by testing

units under high levels of stress variables. The stress variables usually include temperature, usage frequency, and load. Compared to the laboratory test data, field data are often more complicated, containing truncation, censoring, reporting delays and a large amount of data. Field data are usually more valuable because they come from reality and include complicated environmental variables and usage conditions (Oh and Bai 2001).

1.1.2.1 Censoring

In field data, it is rare to observe complete data (i.e., the exact event time for each unit) due to the limit of the study time and the difficulty to track the product all of the time. Thus, censoring is common in lifetime data and could consist of right censoring, left censoring, interval censoring, and random censoring. Right censoring occurs when the study stops before the event has happened. For products with multiple failure modes, right censoring can occur due to the occurrence of other failure modes. Escobar and Meeker (1999) described prediction methods based on censored lifetime data. Hong et al. (2009) proposed a prediction method for left truncated and right censored data. Geisser (1993) presented a general procedure using a Bayesian framework.

1.1.2.2 Reporting Delayed

One way of field data collection is through the repair requests from customers in the warranty period or a track study from the manufacturers. Due to various reasons, some failures may be reported immediately, while others may be reported with delays. For those delays outside of the warranty period or study time, we may not be able to collect the failure events. Without the consideration of the reporting delay, the prediction of failure events will be more conservative (i.e., more failures).

1.1.2.3 Dynamic Covariates

With the advancement of technology, sensors are becoming more widely used. Sensors can help record the environmental variables and the usage conditions (i.e., dynamic covariates)

as well as the time to failure. These dynamic covariates can be the load amount, use rate, temperature, humidity, etc. Dynamic covariates are expected to give more information for reliability analysis. For example, the lifetime of solar panels is affected by the humidity of the air and the amount of ultraviolet radiation. The degradation of LED lights is affected by the frequency of on-off switches. The wearing of rubber tires is influenced by road conditions. Dynamic covariates also can help explain individual variability. For example, organic coating under high temperature will degrade faster than in normal temperature. Thus, incorporation of dynamic covariates in the analysis of reliability is becoming more important. Singpurwalla (1995), Nelson (1990, Chapter 10), Bagdonavičius and Nikulin (2001a), Bagdonavičius and Nikulin (2001b, Chapter 2), Lawless and Crowder (2004), Hong and Meeker (2011), and Hong et al. (2014) introduced the methods to incorporate dynamic covariates.

Because sensors can record environmental variables and usage conditions in seconds, the next generation of reliability data will be very large in size (Hong and Meeker 2013). The number of covariates will probably be much larger than the number of observations and so extracting useful dynamic covariates information will be a great challenge in future.

1.1.2.4 Multi-level Repairable System

The structure of a system can be complicated (e.g., containing several levels). For example, the structure of a truck (system level) includes engine subsystem, brake subsystem, gas supply subsystem, etc. Subsystems are also composed of many components. Failures may occur in some critical subsystems. For a repairable system, the system can be restored back to an operational state after a repair event. To fix the problem in a repairable system, we may replace the whole subsystem or just replace one component inside the subsystem. The replacement of a subsystem may change the lifetime of the components in that subsystem. The complicated data structure and the effect from the subsystem replacement events increase the challenge in the modeling and prediction of component replacement events.

1.1.2.5 Degradation Data in Reliability

Many materials such as metal, organic coating and polymers degrade over time due to the physical changes (e.g., wearing or fatigue) and/or chemical changes (e.g., aging or oxidation). Once the degradation reaches a critical value, the material is defined as a failure. For those materials with degradation measurements, degradation analysis provides a powerful way to estimate the reliability. Instead of recording only the failure times, the degradation data will contain a series of degradation measurements over time. Compared to the traditional time-to-failure data analysis, degradation data analysis can give a better and more timely estimation of highly reliable products, even when no failure is observed. There are two primarily analytic methods in current degradation data analysis: general degradation path model and stochastic model. Lu and Meeker (1993), Wang and Coit (2007) and Hong et al. (2014) described the method of general degradation path model. Whitmore (1995), Padgett and Tomlinson (2004), and Park and Padgett (2006) applied stochastic processes to model degradation data.

1.1.2.6 Recurrent Event in Reliability

The observation of recurrent events is common in repairable systems. Current research mainly focuses on the estimation and prediction of future events and the expected number of events in a certain future time interval. Nelson (2003) and Cook and Lawless (2007) gave a good introduction to the statistical methods in the analysis of recurrent events. Nonhomogeneous Poisson process and renewal process are two popular models in recurrent event modeling. Berman (1981) introduced an inhomogeneous gamma process. Kijima (1989) proposed two types of virtual age models. Wang and Pham (1996) proposed a quasi renewal process. Lawless and Thiagarajah (1996) introduced a proportional intensity model and Lindqvist et al. (2003) proposed a trend-renewal process.

1.2 Motivation

The goal of this dissertation is to model complex reliability data and develop prediction procedures to address prediction problems arising from different situations. As we mentioned before, the structure of field data is complicated and the capacity of current software packages is far from what is needed to analyze modern reliability data. In this dissertation, three modeling and prediction methods are developed which can be generally applied to other reliability data sharing some key characteristics.

1.2.1 Failure Time and Retirement Time

To make the products more competitive, manufacturers will upgrade their products and develop more functions on the new generation of products. These actions will stimulate customers to replace their older version for the latest version, even when the older version is still in working condition. The shorter retirement time of products can affect the prediction of failure events and mislead managers into making incorrect decisions. However, there is little research considering retirement so far. In Chapter 2, a dataset from a home appliance is considered. In this dataset, 120,921 units were staggered installed at home since January 1996. A potential dangerous failure mode was reported after 55 months and the failure data were collected before 118 months from the first installation. To consider the retirement effect in the prediction, a separate survey was made to collect the retirement information. It found the average retirement time was around 98 months. Besides the retirement effect, reporting delay was also considered. Based on this dataset, we develop a statistical prediction procedure that considers the impact of product retirements and reporting delays. The statistical prediction method can be applied to other lifetime data with similar characteristics (i.e., include retirement or reporting delay or both).

1.2.2 Nonlinear Mixed-effects Model and Dynamic Covariates

Traditional reliability analysis is based on failure-to-time data. However, many products are designed with high reliability and low risk of failures. It is hard to collect enough failure-to-time data to give an accurate estimation for highly reliable products in a limited time, even under accelerated test conditions. Compared to traditional lifetime data analysis, degradation data analysis can give a better estimation and prediction for the lifetime of products under limited study time and small sample units. Life-affecting environmental variables (i.e., dynamic covariates) are recorded as well as degradation measures over time. It is important to incorporate the dynamic covariates into the degradation path model. In Chapter 3, we propose a nonlinear mixed-effects model incorporating dynamic covariates to model the degradation data. Shape restricted regression splines are used in the proposed model and a modified alternating algorithm is developed in this chapter. The performance of the algorithm is evaluated by simulation. Based on the relationship between failure and degradation measurements, we also propose a method to obtain the lifetime distribution. An outdoor weathering dataset is used for illustration of the proposed method and a comparison between linear mixed-effects model fitting (Hong et al. 2014) and nonlinear mixed-effects model fitting is made.

1.2.3 Recurrent Events in Multi-level Repairable Systems

A multi-level repairable system may experience recurrent events in the subsystem level or the component level. The replacement of the subsystem may affect the risk of components having failures. So far, most of the research, however, has treated the subsystem and component independently in the modeling of component events (e.g., Martz et al. 1988, and Martz and Waller 1990). The ignorance of the effect from the subsystem event may lead to bias in estimation and prediction of the component events. In Chapter 4, we introduce a multi-level trend-renewal process (MTRP) to model the component events with the consideration of the subsystem event effect. To explain the unit-to-unit variability, time-dependent covariates and random effects are considered in the modeling and prediction. The prediction of the

component events is achieved based on the Monte Carlo simulation. The performance of the estimation and prediction procedures are validated by simulation studies.

1.3 Overview

The rest of this dissertation is organized as follows. In Chapter 2, a statistical method is proposed to model and predict the lifetime of products with the consideration of retirement times and reporting delays. Chapter 2 is mainly based on Xu et al. (2014b), which is to appear in *IEEE Transactions on Reliability*. Chapter 3 introduces a modified alternative algorithm to estimate the unknown parameters in a nonlinear mixed-effects model incorporating dynamic covariates. Chapter 3 is based on Xu et al. (2014a), which is under review by *Applied Stochastic Models in Business and Industry* after a minor revision. Chapter 4 proposes a multi-level trend-renewal process to model the component events in multi-level repairable systems. Chapter 4 will be turned into a paper and will be submitted to *Technometrics*.

Bibliography

V. Bagdonavičius and M. S. Nikulin. Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, 7:85–103, 2001a.

V. Bagdonavičius and M. S. Nikulin. *Accelerated Life Models: Modeling and Statistical Analysis*. Chapman & Hall/CRC, Boca Raton, FL, 2001b.

M. Berman. Inhomogeneous and modulated gamma processes. *Biometrika*, 68:143–152, 1981.

R. J. Cook and J. F. Lawless. *The Statistical Analysis of Recurrent Events*. Springer-Verlag, New York, 2007.

L. A. Escobar and W. Q. Meeker. Statistical prediction based on censored life data. *Technometrics*, 41:113–124, 1999.

S. Geisser. *Predictive Inference: An Introduction*. Chapman and Hall, New York, 1993.

Y. Hong and W. Q. Meeker. A model for field-failure prediction using dynamic environmental data. In N. Balakrishnan, M. Nikulin, and V. Rykov, editors, *Mathematical and Statistical Methods in Reliability. Applications to Medicine, Finance and Quality Control*, chapter 16. Birkhauser: Boston, 2011.

Y. Hong and W. Q. Meeker. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics*, 55(2):135–149, 2013.

Y. Hong, W. Q. Meeker, and J. D. McCalley. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics*, 3:857–879, 2009.

Y. Hong, Y. Duan, W. Q. Meeker, D. L. Stanley, and X. Gu. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics*, 2014. in press, DOI:10.1080/00401706.2014.915891.

M. Kijima. Some results for repairable systems with general repair. *Journal of Applied Probability*, 26:89–102, 1989.

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

J. Lawless and K. Thiagarajah. A point-process model incorporating renewals and time trends, with application to repairable systems. *Technometrics*, 38:131–138, 1996.

B. Lindqvist, G. Elvebakk, and K. Heggland. The trend-renewal process for statistical analysis of repairable systems. *Technometrics*, 45:31–44, 2003.

C. J. Lu and W. Q. Meeker. Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 34:161–174, 1993.

H. F. Martz and R. A. Waller. Bayesian reliability analysis of complex series/parallel systems of binomial subsystems and components. *Technometrics*, 32:407–416, 1990.

H. F. Martz, R. A. Waller, and E. T. Fickas. Bayesian reliability analysis of series systems of binomial subsystems and components. *Technometrics*, 30:143–154, 1988.

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

W. Nelson. *Applied Life Data Analysis*. John Wiley & Sons, New York, 1982.

W. Nelson. *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses, (Republished in a paperback in Wiley Series in Probability and Statistics, 2004)*. John Wiley & Sons, New York, 1990.

W. B. Nelson. *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*. ASA-SIAM, Philadelphia, 2003.

Y. S. Oh and D. S. Bai. Field data analyses with additional after-warranty failure data. *Reliability Engineering & System Safety*, 72:1–8, 2001.

W. J. Padgett and M. A. Tomlinson. Inference from accelerated degradation and failure data based on Gaussian process models. *Lifetime Data Analysis*, 10:191–206, 2004.

C. Park and W. J. Padgett. Stochastic degradation models with several accelerating variables. *IEEE Transactions on Reliability*, 55:379–390, 2006.

N. D. Singpurwalla. Survival in dynamic environments. *Statistical Science*, 10:86–103, 1995.

H. Wang and H. Pham. A quasi renewal process and its applications in imperfect maintenance. *International Journal of Systems Science*, 27:1055–1062, 1996.

P. Wang and D. Coit. Reliability assessment based on degradation modeling with random or uncertain failure threshold. In *Proceedings of the 2007 Reliability & Maintainability Symposium (RAMS)*, Orlando, FL, 2007.

G. A. Whitmore. Estimation degradation by a Wiener diffusion process subject to measurement error. *Lifetime Data Analysis*, 1:307–319, 1995.

Z. Xu, Y. Hong, and R. Jin. Nonlinear general path models for degradation data with dynamic covariates. *Applied Stochastic Models in Business and Industry*, 2014a. Under minor revision.

Z. Xu, Y. Hong, and W. Q. Meeker. Assessing risk of a serious failure mode based on limited field data. *IEEE Transactions on Reliability*, PP:1–12, 2014b.

Chapter 2 Assessing Risk of a Serious Failure Mode Based on Limited Field Data

Zhibing Xu[1], Yili Hong[1], and William Q. Meeker[2]

[1]Department of Statistics, Virginia Tech, Blacksburg, VA 24061

[2]Department of Statistics, Iowa State University, Ames, IA 50011

A paper accepted by *IEEE Transactions on Reliability*

Abstract

Many consumer products are designed and manufactured so that the probability of failure during the technological life of the product is small. Most product units in the field retire before they fail. Even though the number of failures of such products is small, there is still a need to model and predict field failures for purposes of risk assessment in applications that involve safety. Challenges in the modeling and prediction of failures arise because the retirement times are often unknown, few failures have been reported, and there are delays in field failure reporting. Motivated by an application to assess the risk of failure for a particular product, we develop a statistical prediction procedure that considers the impact of product retirements and reporting delays. Based on the developed method, we provide the point predictions for the cumulative number of reported failures over a future time period, and corresponding prediction intervals to quantify uncertainty. We also conduct sensitivity analysis to assess the effects of different assumptions on failure-time and retirement distributions.

Key Words: Calibration, discrete Fourier transform, failure reporting delay, Poisson-binomial distribution, prediction interval, Weibull.

2.1 Introduction

2.1.1 Motivation

Product reliability is important to both the manufacturers and consumers. Many products have high reliability with only a small fraction failing (e.g., 1% or less). There are, however, some failure modes that can lead to risk of loss of property or life. Examples include material anomalies in rotating components in aircraft engines that lead to premature cracking and fracture, failure of electrical insulation in home appliances giving rise to the risk of fire or electrical shock, failure of electrical connections in defibrillators, and the explosion of a laptop battery.

Although the particular technical details and nature of the available data and other information will differ from application to application, there is a common scenario that we have seen in numerous different applications. At some point in time (which may range from months to years) after product introduction, a few failures have been reported. Often the particular failure mode is one that had not been anticipated. Sometimes the problem was caused by just a single batch of raw material, or an unreported and untested change in a component or material made by a vendor. Generally, management (or in some cases government agencies) will want engineers to determine whether there is a serious problem, and will often ask for a formal risk assessment. This action then leads to some or all of the following questions.

1. Were the reported failures anomalies (e.g., cause by extreme product abuse or a few defective units that got shipped) or is the problem more widespread? Usually, it is the latter, but wishful thinking will cause some to believe the former.

2. Is there a small proportion of defective units failing rapidly or will all units (that remain in service) eventually fail prematurely?

3. What is the risk (e.g., potential cost, both tangible and intangible) of future failures from this product?

4. Should there be a product recall?

5. How can we fix the problem so that future production will not have the failure mode of concern?

This chapter focuses on statistical methods for answering question 3.

In some applications, the risk of failure is lessened because of product retirement, before product failure occurs. Retirements are often a result of product performance degradation or technical obsolescence. For example, cell phones and laptop computers are typically retired after two or three years of use. Ironically, for some products, the risk of a serious failure is sometimes lessened by the occurrence of an innocuous failure mode. For example, an implanted defibrillator that has a broken electrical connection would be removed from service if its rechargeable battery fails before the unit is called upon to be used. In such applications, possible retirement/innocuous failure events should be part of the risk assessment. Another complicating feature of some field data is the delayed reporting of failures.

Motivated by several different but similar applications, we develop a statistical procedure to predict the field failures of products, considering the impact of product retirement and reporting delays. Based on the developed method, we provide point predictions for the cumulative number of reported failures at a future point in time, and the corresponding prediction interval (PI) to quantify uncertainty.

2.1.2 Related work

There is a large amount of literature describing statistical prediction, and some of this previous work has focused on the prediction of the number of failures in a future time period and the construction of a corresponding PI. Nelson (1982), and Meeker and Escobar (1998) introduced general methods to obtain PIs for reliability applications. Engehardt and Bain (1978) provided an exact PI for the number of failures in a repairable system based on maximum likelihood (ML) estimation. Mee and Kushary (1994) gave simulation-based methods for computing PIs for selected order statistics from future samples from a Weibull distribution. Nelson (2000),

and Nordman and Meeker (2002) proposed PI procedures based on a Weibull distribution with a known shape parameter. Geisser (1993), and Tian et al. (2011) described Bayesian approaches to obtain a PI. De Menezes et al. (2006) used subsampling to obtain the PI for the number of failures in a future time interval. For censored failure-time data, Escobar and Meeker (1999), and Hong et al. (2009) described methods to obtain PIs for a future number of failures. Lawless and Fredette (2005) proposed an effective, easy-to-use procedure to construct frequentist PIs. Yang (2010) proposed a prediction method for warranty cost based on the accelerated life test plans. Park and Kulasekera (2004) described a parametric method to deal with the competing risks problem. Few published works, however, have considered prediction in the presence of the unknown retirement times and reporting delays. In one exception, Zhao et al. (2010) compared the difference of predictions between the models that account for retirement and that do not account for retirement, based on a specific retirement rate assumption. In this chapter, we propose a general statistical procedure to predict the future number of field failures in the presence of retirement and reporting delays, and we develop a PI procedure to quantify the uncertainties in prediction.

2.1.3 Overview

Our approach to the field-failure prediction uses the following steps.

- *Failure-time modeling*- We first construct a failure-time model based on assumptions for the retirement-time distribution and reporting delays. Then, we estimate the parameters of the failure-time distribution using ML.

- *Derivation of the probability of future failures* - Based on the failure-time and retirement-time distributions, as well as the ML estimates from Step 1, the probability of a reported failure in a future time interval can be estimated, providing the basis of prediction of the cumulative number of reported failures in a specified future time period.

- *Prediction* - Based on the probability of a reported failure and the number of units that are at risk, one can obtain a point prediction for the cumulative number of reported

failures by a specified future point in time and a corresponding PI.

- *Prediction Interval* - We use a method based on the concept of a predictive distribution and bootstrap calibration to construct PIs for the future number of failures.

- *Sensitivity analysis* - The predictions are based on uncertain assumptions about the failure-time and the retirement-time distributions. Thus, it is prudent to assess the effect of deviations from these assumptions.

The rest of the chapter is organized as follows. In Section 2.2, we introduce the failure-time distribution, the retirement-time distribution, and the reporting delay distribution. In Section 2.3, we develop an ML procedure to estimate the unknown failure-time distribution parameters. Section 2.4 shows in detail how to use the failure-time distribution and estimated parameters to predict the number of reported future failures, and how to compute a corresponding PI. In Section 2.5, sensitivity analysis is used to compare the prediction results with different parameters and distributions. Section 2.6 contains some concluding remarks, and describes possible areas for future research.

2.2 Data and Failure-Time Model

2.2.1 The Data

In this chapter, we use a dataset from a product that is used at home that we call product B. To protect proprietary and sensitive information, we have disguised the data by changing the time scale, and using a randomly chosen subset of the original dataset. Although our methods were motivated by this specific application, the developed method is general, and can be applied to other situations with unknown retirement times and reporting delays.

The company manufactured 14 batches of product B over time, and there were 120,921 units in total. The units were put into service at different times between January 1996 and June 1997 (staggered entry). We define the first installation time (i.e., January 1996) to be time 0. Then the installation times for the 14 different batches were 0, 2, 3, 4, 5, 6, 7, 10,

12, 13, 14, 15, 16, and 17 months, respectively. After 55 months (a little less than five years) later, a potentially dangerous failure mode was reported. Subsequently, 32 additional failures were reported by the data freeze date (DFD), which was at 118 months (about 9.8 years) after the first units were introduced into service. Figure 2.1 illustrates the staggered entry pattern. The figure shows, for each batch, the number of units that had been installed, the time where failures were reported, and the number of units that had not being reported as failing by the DFD.

Table 2.1 shows the 32 reported failure times (months in service before failure). The failure times are denoted by t_i, $i = 1, \cdots, r$ where r is the total number of reported failures ($r = 32$ here). The failure times were recorded to the nearest month. For example, failure time 91 indicates that a unit failed between 90.5 and 91.5 months after its installation. Table 2.1 also lists the age that the failed unit would have been at the DFD if it had not failed; these times are denoted by $\mathcal{A}_i, i = 1, \cdots, r$.

Table 2.2 shows the number of units installed, the number of failures reported, the number of units not being reported by the DFD, and the ages of units at the DFD (denoted by $\mathcal{A}_j$) for the 14 batches of product B. The number of units installed, denoted by $n_j, j = 1, \cdots, J$, where J is the total number of batches ($J = 14$ here), ranges between 5,795 and 12,233. The number of units that were not reported is denoted by w_j. Although the number of reported failures from each batch and the overall fraction failing is small (0.026%), it is worthy to investigate such data, and predict the number of failures in the future due to the potential serious consequences of the failure.

2.2.2 Model for Time to Failure

Let T denote the product failure time. We use the log-location-scale family of distributions to model the distribution of T. Among those members in the log-location-scale family, the Weibull and lognormal distributions are the two most commonly-used distributions for describing failure times. In particular, the cumulative distribution function (cdf) and proba-

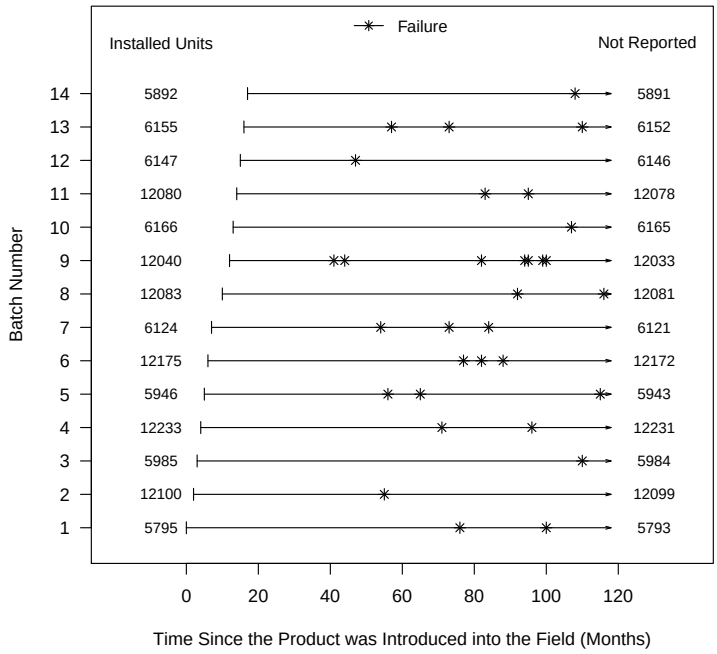


Figure 2.1: Event plot for the dataset with $*$ indicating the failure times for the reported failures.

Table 2.1: Failure Time t_i and Corresponding Age of the unit's Batch at the DFD $\mathcal{A}_i$.

t_i	$\mathcal{A}_i$	t_i	$\mathcal{A}_i$	t_i	$\mathcal{A}_i$	t_i	$\mathcal{A}_i$
91	101	88	106	106	108	60	113
41	102	82	106	47	111	51	113
94	102	32	106	77	111	92	114
57	102	29	106	66	111	67	114
32	103	83	106	76	112	107	115
69	104	87	106	82	112	53	116
81	104	70	106	71	112	100	118
94	105	82	108	110	113	76	118

Table 2.2: Installed quantities, number of failures reported, number not reported, and the age of the batch at the DFD.

Batch j	Installed n_j	Reported r_j	Not Reported $w_j = n_j - r_j$	Batch Age at DFD $\mathcal{A}_j$
1	5795	2	5793	118
2	12100	1	12099	116
3	5985	1	5984	115
4	12233	2	12231	114
5	5946	3	5943	113
6	12175	3	12172	112
7	6124	3	6121	111
8	12083	2	12081	108
9	12040	7	12033	106
10	6166	1	6165	105
11	12080	2	12078	104
12	6147	1	6146	103
13	6155	3	6152	102
14	5892	1	5891	101

bility density function (pdf) of the Weibull distribution can be expressed as

$$F_T(t;\mu_T,\sigma_T) = \Phi_{\text{sev}}\left[\frac{\log(t)-\mu_T}{\sigma_T}\right] \quad \text{and} \quad f_T(t;\mu_T,\sigma_T) = \frac{1}{\sigma_T t}\phi_{\text{sev}}\left[\frac{\log(t)-\mu_T}{\sigma_T}\right] \quad t>0,$$

where $\Phi_{\text{sev}}(w) = 1-\exp[-\exp(w)]$, and $\phi_{\text{sev}}(w) = \exp[w-\exp(w)]$ are the standard smallest extreme value cdf and pdf, respectively. Here, μ_T is the location parameter, and σ_T is the scale parameter of the distribution of $\log(T)$.

We use the Weibull distribution to describe the distribution of T. In our sensitivity analysis, the lognormal distribution is considered as an alternative to describe the failure-time distribution. By replacing Φ_{sev}, and ϕ_{sev} with Φ_{nor}, and ϕ_{nor}, the cdf, and pdf of the lognormal distribution are

$$F_T(t;\mu_T,\sigma_T) = \Phi_{\text{nor}}\left[\frac{\log(t)-\mu_T}{\sigma_T}\right] \quad \text{, and } f_T(t;\mu_T,\sigma_T) = \frac{1}{\sigma_T t}\phi_{\text{nor}}\left[\frac{\log(t)-\mu_T}{\sigma_T}\right] \quad t>0,$$

respectively.

The cdf, and pdf of the Weibull distribution can also be re-expressed as

$$F_T(t;\eta_T,\beta_T) = 1-\exp\left[-\left(\frac{t}{\eta_T}\right)^{\beta_T}\right] \quad \text{and} \quad f_T(t;\eta_T,\beta_T) = \frac{\beta_T}{\eta_T}\left(\frac{t}{\eta_T}\right)^{\beta_T-1}\exp\left[-\left(\frac{t}{\eta_T}\right)^{\beta_T}\right],$$

where $t>0$, $\eta_T = \exp(\mu_T)$ is the Weibull scale parameter (also the approximate 0.63 quantile),

and $\beta_T = 1/\sigma_T$ is the Weibull shape parameter. The value of β_T indicates the shape of the hazard function, which is given by

$$h(t;\beta_T,\eta_T) = \frac{f(t;\beta_T,\eta_T)}{1-F(t;\beta_T,\eta_T)} = \frac{\beta_T}{\eta_T}\left(\frac{t}{\eta_T}\right)^{\beta_T-1}.$$

In particular, $\beta_T > 1$ indicates an increasing hazard function; $\beta_T = 1$ indicates a constant hazard function; and $\beta_T < 1$ indicates a decreasing hazard function. More information about the log-location-scale family of distributions can be found in Chapter 4 of Meeker and Escobar (1998).

2.2.3 Product Retirement Distribution

Product retirement occurs when a unit is removed from service before it fails. Let R be the time of retirement. To avoid prediction bias, it is important to incorporate the retirement information into the failure-time model. However, there is no tracking of the retirement time at the individual product level. Thus, the retirement information at the population level has to be used.

A marketing survey had been conducted, providing information about the length of time that people tended to use certain appliances before replacing them with new models. This survey provided information about the mean and standard deviation of appliances like product B. By assuming that the retirement times s-independently and identically follow a Weibull distribution, we obtain that the mean retirement time is $\mathbf{E}(R) = 98$ months (approximately 8.2 years), and the shape parameter β_R is between 1.5 and 2. The cdf of the retirement time can be expressed as

$$F_R(r) = \Pr(R \le r) = 1 - \exp\left[-\left(\frac{r}{\eta_R}\right)^{\beta_R}\right], \quad r > 0. \tag{2.1}$$

The value of shape parameter $\beta_R > 1$ indicates that the retirement hazard function is an increasing function of product age. The mean of the Weibull distribution is $\mathbf{E}(R) = \eta_R\Gamma(1 + 1/\beta_R)$, implying that the Weibull characteristic life parameter η_R is between 108.6 and 110.6 months. Due to the nature of the serious failure mechanism (because it has no symptoms before it occurs), it can reasonably be assumed to be s-independent of the time of retirement.

Table 2.3: Probability mass function for the distribution of reporting delays.

δ (months)	0	1	2	3 - 5	6 - 9	10 - 15	≥ 16
$\Pr(\Delta = \delta)$	0.62	0.31	0.04	0.004	0.003	0.001	0

2.2.4 Failure Reporting Delay

In this application, there were known delays in the reporting of failures. Because these delays were potentially important to the estimation of the failure-time distribution, there was need to consider them in modeling and prediction. We denote the length of the delay by Δ, which is assumed to be s-independent of the failure time. The reporting time is equal to $T + \Delta$, where T is the product's failure time.

Based on available records, no reporting delays had been longer than 15 months. Thus, the probability of a reporting delay greater than or equal to 16 months is equal to zero, and all delay times are between 0 month to 15 months. Based on historical information, the distribution of delays is approximated by a discrete distribution given in Table 2.3. A particular delay time is denoted by δ, and the corresponding probability is denoted by $\Pr(\Delta = \delta)$. Table 2.3 indicates that around 62% of failures would be reported to the company without any delay. Note that $\sum_\delta \Pr(\Delta = \delta) = 1$. Because the failure process and reporting process are not related, it is reasonable to assume that the reporting delay Δ is s-independent of failure time T.

2.3 Maximum Likelihood Estimation

2.3.1 Construction of the Likelihood Function

Let t_i denote the realized failure time of unit i, which is the amount of time between when the unit was installed and when it failed. If a unit failed, and the failure was reported before the DFD, the failure that occurred at time t_i was recorded. Because the failure times were recorded to the nearest month, the actual failure time for observation i is in the interval

$(t_{i1}, t_{i2}]$, where $t_{i1} = \max(0, t_i - .5)$, and $t_{i2} = \min(t_i + 0.5, \mathcal{A}_i)$. The probability of a failure before retirement with failure time between t_{i1} and t_{i2} is

$$\Pr[(T \leq R) \cap (t_{i1} < T \leq t_{i2})] = \int_{t_{i1}}^{t_{i2}} f_T(t)[1 - F_R(t)]dt \tag{2.2}$$

where the factor $1-F_R(t)$ represents the probability that the unit retires after time t. One can consider the failure time T and retirement time R as in a competing-risks model (e.g., Chapter 2 of Crowder, 2001). Figure 2.2 illustrates the computing of the likelihood contribution in (2.2) in which the shaded area shows the likelihood contribution.

To account for reporting delay, (2.2) needs to be adjusted. Here, the reporting delay is incorporated into the model by conditioning on the observed value of Δ. In particular, the probability of actually failing in the interval $(t_{i1}, t_{i2}]$ and having the failure reported before the DFD is

$$\begin{aligned}
\pi_i &= \Pr[(T \leq R) \cap (t_{i1} < T \leq t_{i2}) \cap \text{Reported}] \\
&= \sum_{\delta} \Pr(\Delta_i = \delta) \Pr[(T \leq R) \cap (t_{i1} < T \leq t_{i2}) \cap \text{Reported} | \Delta_i = \delta] \\
&= \sum_{\delta} \Pr(\Delta_i = \delta) \int_{t_{i1}}^{t_{i2}} \mathbf{1}(t + \delta; t_{i1}, \mathcal{A}_i) f_T(t)[1 - F_R(t)]dt
\end{aligned} \tag{2.3}$$

where

$$\mathbf{1}(t + \delta; t_{i1}, \mathcal{A}_i) = \begin{cases} 1 & \text{when } t_{i1} < t + \delta \leq \mathcal{A}_i \\ 0 & \text{otherwise} \end{cases}.$$

The indicator function $\mathbf{1}(t + \delta; t_{i1}, \mathcal{A}_i)$ accounts for the censoring that arises because we only know about failures that are reported before the DFD. For purposes of numerical computation, (2.3) can be re-expressed as

$$\pi_i = \int_{t_{i1}}^{t_{i2}} \sum_{\delta} \Pr(\Delta_i = \delta) \mathbf{1}(t + \delta; t_{i1}, \mathcal{A}_i) f_T(t)[1 - F_R(t)]dt.$$

For those units that were not reported as failures before DFD, the probability that a unit in

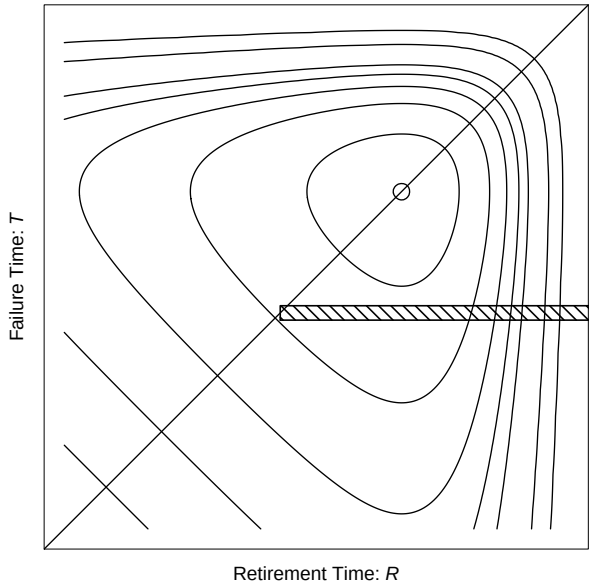


Figure 2.2: Illustration of the likelihood contribution in (2.2) relative to the joint distribution of T and R. The probability under the shaded region is the likelihood contribution.

installation batch j has not been reported as a failure before DFD is

$$\begin{aligned}\xi_j &= 1 - \Pr[(T \le R) \cap (0 < T \le \mathcal{A}_j) \cap \text{Reported}] \\ &= 1 - \sum_{\delta} \Pr(\Delta_j = \delta) \Pr[(T \le R) \cap (0 < T \le \mathcal{A}_j) \cap \text{Reported} | \delta] \\ &= 1 - \sum_{\delta} \Pr(\Delta_j = \delta) \int_0^{\mathcal{A}_j} \mathbf{1}(t+\delta; 0, \mathcal{A}_j) f_T(t)[1 - F_R(t)]dt \end{aligned} \tag{2.4}$$

where

$$\mathbf{1}(t+\delta; 0, \mathcal{A}_j) = \begin{cases} 1 & \text{when } 0 \le t+\delta \le \mathcal{A}_j \\ 0 & \text{otherwise} \end{cases}.$$

Equation (2.4) can be re-expressed as

$$\xi_j = 1 - \int_0^{\mathcal{A}_j} \sum_{\delta} \Pr(\Delta_j = \delta) \mathbf{1}(t+\delta; 0, \mathcal{A}_j) f_T(t)[1 - F_R(t)]dt.$$

The log-likelihood function based on the data in Tables 2.1 and 2.2 is

$$\mathcal{L}(\boldsymbol{\theta} | \text{DATA}) = \sum_{i=1}^{r} \log(\pi_i) + \sum_{j=1}^{J} w_j \log(\xi_j), \tag{2.5}$$

where $\boldsymbol{\theta} = (\eta_T, \beta_T)'$. Here, the first summation is over the reported failures, the second summation is over the installation batches in Table 2.2, and w_j is the number of units from batch j that have not been reported as failures.

2.3.2 Parameter Estimates

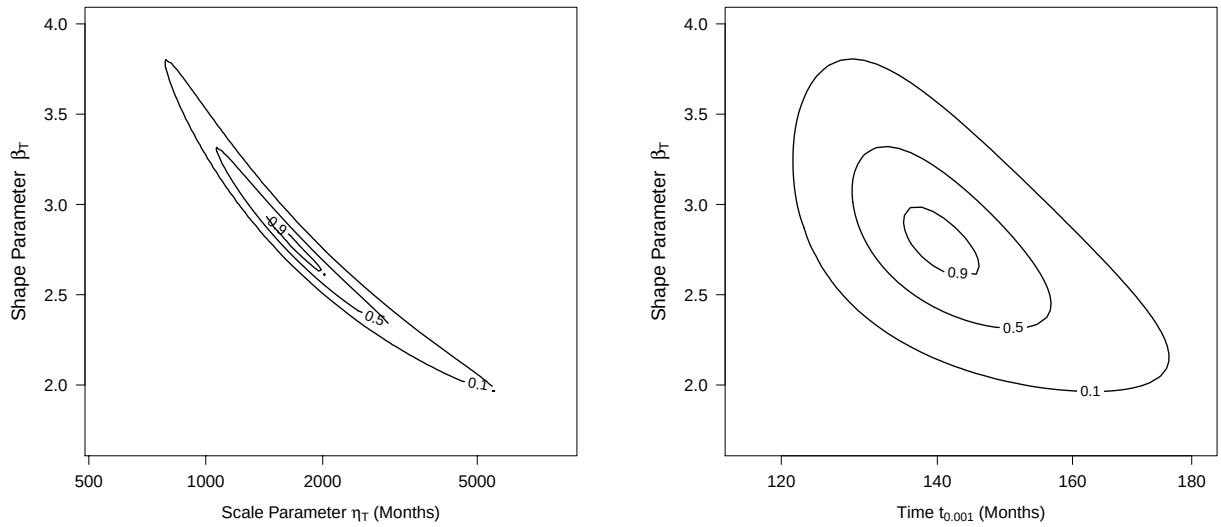
The ML estimator of $\boldsymbol{\theta}$ is denoted by $\widehat{\boldsymbol{\theta}} = (\widehat{\eta}_T, \widehat{\beta}_T)'$. To make the numerical optimization more stable, we optimized the loglikelihood function using an alternative parametrization $t_{0.001}$ and β_T, instead of the original parametrization η_T and β_T. Here, $t_{0.001}$ is the 0.001 quantile of the product failure-time distribution. The effect of the reparametrization is shown in the contour plots of the log relative likelihood as in Figure 2.3. Under the original parametrization, the shape of the log relative likelihood is elongated, indicating a strong correlation between $\widehat{\eta}_T$ and $\widehat{\beta}_T$. Such strong correlation will make the numerical optimization less stable. Under the alternative parametrization, the log-likelihood is better behaved. In addition, it is more meaningful to parameterize it using $t_{0.001}$ instead of using η_T. This is because η_T is the time by which 63.2% of the population will fail, which is hard to reach with the consideration of retirement time. Due to the invariance property of ML estimators, the ML estimates obtained under the alternative parametrization can be transformed to the ML estimates for the original parameters for subsequent computations. Here, $\widehat{\eta}_T = \exp(\widehat{\mu}_T) = \exp[\log(\widehat{t}_{0.001}) - z_{0.001}\widehat{\sigma}_T]$, where $z_{0.001} = \Phi^{-1}_{\text{sev}}(0.001)$ is the 0.001 quantile of the standard smallest extreme value distribution. The standard error of $\widehat{\boldsymbol{\theta}}$ can be obtained by using the delta method. More details about the delta method can be found in page 626 of Meeker and Escobar (1998). The CI can be computed by the transformation as in Hong et al. (2008).

Table 2.4 shows the estimates of the Weibull shape and scale parameters based on the Weibull retirement distribution assumption with $\mathbf{E}(R) = 98$, and $\beta_R = 1.5$. Under this assumption, the Weibull shape parameter estimate is $\widehat{\beta}_T = 2.788$, which is larger than 1, indicating that the failure-time distribution hazard function is increasing, which is in agreement with the known physical degradation cause of failure.

It is important to consider retirement when one needs to determine the fraction reported. As an illustration, Figure 2.4 shows the failure-time distributions with and without adjustment of retirement. Here, the distribution with adjustment of retirement is plotted based on $\int_0^t \widehat{f}_T(x)[1 - F_R(x)]dx$, and the distribution without adjustment of retirement time is plotted

Table 2.4: ML estimates and standard errors for parameters under $\mathbf{E}(R) = 98, \beta_R = 1.5$.

Parameter	ML est.	Std. err.	95% lower	95% upper
β_T	2.788	0.428	2.064	3.766
η_T	1670.902	730.451	709.316	3936.066



(a) Original Parametrization (b) Alternative Parametrization

Figure 2.3: Comparison of shape of the log relative likelihood function of the two parametrizations with $\mathbf{E}(R) = 98$, and $\beta_R = 1.5$.

based on $\int_0^t \widehat{f}_T(x)dx$, where $\widehat{f}_T(x)$ is the estimated pdf based on ML estimates in Table 2.4 given $\mathbf{E}(R) = 98$ and $\beta_R = 1.5$. The figure illustrates the large effect that the retirement distribution plays in determining the fraction reported as a function of time. That is, the failure probability of the distribution without considering retirement is much larger than the one with retirement as time goes.

2.4 Prediction of Future Number of Reports

Prediction is important in reliability analysis, even for highly reliable products, because failure events have the potential to cause serious damage, or have other serious consequences such as an automobile accident, the explosion of power transformers, and the failure of power systems

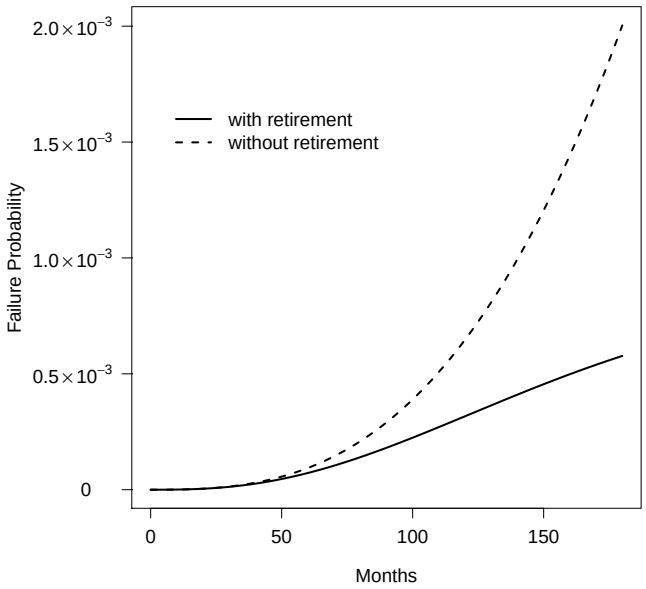


Figure 2.4: Failure-time distributions with and without adjustment of retirement when $\mathbf{E}(R) = 98$, and $\beta_R = 1.5$.

in satellites. The prediction for highly reliable products is quite challenging due to the limited number of the failure events. In this section, we propose a method of prediction for the number of reported failures in the future. We consider both point prediction and the construction of corresponding prediction intervals.

2.4.1 Probability of Being Reported Before Retirement

Based on the model for the failure-time distribution, the ML estimates $\widehat{\eta}_T$ and $\widehat{\beta}_T$, and the assumed values of η_R and β_R, one can predict the number of reported failures that will occur before a specified future point in time. Let $\mathcal{A}_j$ denote the age of the units in batch j at the DFD, $j = 1, \cdots, J$.

For a product unit in batch j that has not been reported as a failure before the DFD, it may fail and be reported in the future time interval $(\mathcal{A}_{jl1}, \mathcal{A}_{jl2}]$. Here $\mathcal{A}_{jl1} = \max(\mathcal{A}_j, \mathcal{A}_j + l - .5)$, $\mathcal{A}_{jl2} = \mathcal{A}_j + l + .5$, and l is the number of months after the DFD, $l = 0, 1, 2, \cdots$. The

corresponding probability of being reported is

$$\begin{aligned} h_j(l) &= \Pr(T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}],\, T \le R \,|\, \text{not being reported by } \mathcal{A}_j) \\ &= \frac{\Pr(T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}],\, T \le R, \text{not being reported by } \mathcal{A}_j)}{\Pr(\text{not being reported by } \mathcal{A}_j)} \\ &= \frac{\Pr(T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}],\, T \le R)}{1-\Pr(\text{being reported by } \mathcal{A}_j)}. \end{aligned} \tag{2.6}$$

Here, T denotes the failure time of a unit that has not been reported as a failed unit by the DFD, and Δ denotes the random reporting delayed time. Let

$$\gamma_j(l) = \Pr(T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}], T \le R),$$

and let $\xi_j = 1 - \Pr(\text{being reported by } \mathcal{A}_j)$ in (2.6). In particular,

$$\begin{aligned} \gamma_j(l) &= \Pr[T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}], T \le R] \\ &= \sum_{\delta} \int_0^{\mathcal{A}_{jl2}} \Pr[T+\Delta \in (\mathcal{A}_{jl1}, \mathcal{A}_{jl2}], T \le R | T=t, \Delta=\delta] f_T(t) \Pr(\Delta=\delta) dt \\ &= \sum_{\delta} \int_0^{\mathcal{A}_{jl2}} \mathbf{1}(t+\delta; \mathcal{A}_{jl1}, \mathcal{A}_{jl2})[1-F_R(t)] f_T(t) \Pr(\Delta=\delta) dt \\ &= \sum_{\delta} \Pr(\Delta=\delta) \int_0^{\mathcal{A}_{jl2}} \mathbf{1}(t+\delta; \mathcal{A}_{jl1}, \mathcal{A}_{jl2})[1-F_R(t)] f_T(t) dt \end{aligned}$$

where

$$\mathbf{1}(t+\delta; \mathcal{A}_{jl1}, \mathcal{A}_{jl2}) = \begin{cases} 1 & \text{when } \mathcal{A}_{jl1} < t+\delta \le \mathcal{A}_{jl2} \\ 0 & \text{otherwise} \end{cases}.$$

Here, the indicator function $\mathbf{1}(t+\delta; \mathcal{A}_{jl1}, \mathcal{A}_{jl2})$ constrains the reporting time between $\mathcal{A}_{jl1}$ and $\mathcal{A}_{jl2}$. The factor $1 - F_R(t)$ represents the probability that the unit has not retired before it

fails. Numerical integration is needed to compute γ_j. The quantity ξ_j is computed as

$$\begin{aligned}\xi_j &= 1 - \Pr[\text{being reported by } \mathcal{A}_j] \\ &= 1 - \Pr[(T \le R) \cap (0 < T \le \mathcal{A}_j) \cap \text{Reported}] \\ &= 1 - \sum_{\delta} \Pr(\Delta_j = \delta) \Pr[(T \le R) \cap (0 < T \le \mathcal{A}_j) \cap \text{Reported} | \delta] \\ &= 1 - \sum_{\delta} \Pr(\Delta_j = \delta) \int_0^{\mathcal{A}_j} \mathbf{1}(t+\delta; 0, \mathcal{A}_j) f_T(t)[1 - F_R(t)]dt, \end{aligned} \tag{2.7}$$

which is similar to (2.4). The only difference is that we use (2.4) to estimate parameters, but use (2.7) evaluated at the ML estimates to provide an estimate of the probability that a unit in batch j is not reported as a failure before $\mathcal{A}_j$. Based on the probability function $h_j(l) = \gamma_j(l)/\xi_j$, the probability that a unit in batch j is reported as a failure within s months after the DFD is

$$\rho_j(s) = \sum_{l=0}^{s} h_j(l), \ s = 1, 2, \cdots . \tag{2.8}$$

Because the ages of units at the DFD from different batches are not the same, the $\rho_j(s)$ are different for different batches.

2.4.2 Point Prediction

For unit k in batch j that has not been reported as a failure by $\mathcal{A}_j$, we use $I_{jk}(s)$ as an indicator for being reported in future time interval $(0, s + 0.5]$, where $s = 1, 2, \cdots$. The distribution of $I_{jk}(s)$ is Bernoulli$[\rho_j(s)]$. Thus, the cumulative number of reported failures at time s is

$$N(s) = \sum_{j=1}^{J} \sum_{k=1}^{w_j} I_{jk}(s), \tag{2.9}$$

which is the sum of s-independent and non-identical indicators. The point prediction (estimate of the expected number failing) for the number of reports up to time s is

$$\widehat{N} = \widehat{N}(s) = \sum_{j=1}^{J} w_j \widehat{\rho}_j(s).$$

Here, w_j is the number of units not reported as having failed by the DFD in batch j, and $\widehat{\rho}_j(s)$ is obtained by evaluating (2.8) at the ML estimates η_T and β_T, and the assumed values of η_R and β_R.

2.4.3 Prediction Interval

This section introduces a method of computing PIs for the cumulative number of reported failures at a future time point. The cumulative number of reported failures $N(s)$, as given in (2.9), is the sum of s-independent and non-identically distributed Bernoulli random variables which follows a Poisson-binomial distribution. Hong (2013) gave an exact expression for the cdf of the Poisson-binomial distribution based on a discrete Fourier transform. In particular, the cdf of $N(s)$, denoted by $F_N(n), n = 0, 1, \cdots, n^*$, is

$$F_N(n) = \frac{1}{n^*+1} \sum_{l=0}^{n^*} \left\{ \frac{\exp(-\boldsymbol{i}\omega l n) - \exp(-\boldsymbol{i}\omega l)}{1 - \exp(-\boldsymbol{i}\omega l)} \prod_{j=1}^{J} [1 - \rho_j(s) + \rho_j(s)\exp(\boldsymbol{i}\omega l)]^{w_j} \right\}$$

where $n^* = \sum_j w_j$, $\boldsymbol{i} = \sqrt{-1}$, and $\omega = 2\pi/(n^*+1)$.

Using the predictive distribution given in Lawless and Fredette (2005), a $100(1-\alpha)\%$ PI for $N = N(s)$, denoted by $\left[\underset{\sim}{N}, \widetilde{N}\right]$, is obtained by solving

$$F_N(\underset{\sim}{N}; \widehat{\boldsymbol{\theta}}) = v_{\alpha/2} \quad \text{, and} \quad F_N(\widetilde{N}; \widehat{\boldsymbol{\theta}}) = v_{1-\alpha/2}. \tag{2.10}$$

Here, v_α is the α quantile of the distribution of the random quantity $F_N(N; \widehat{\boldsymbol{\theta}})$, where both N and $\widehat{\boldsymbol{\theta}}$ are treated as random variables. We use a bootstrap simulation procedure to approximate the quantile v_α. In particular, v_α is approximated from B bootstrap samples by the α sample quantile of $F_N(N^{*b}; \widehat{\boldsymbol{\theta}}^{*b}), b = 1, \cdots, B$. Here, N^* is simulated from $F_N(n; \widehat{\boldsymbol{\theta}})$ given the ML estimate $\widehat{\boldsymbol{\theta}}$, and $\widehat{\boldsymbol{\theta}}^*$ is the ML estimates obtained from bootstrap samples. We use the random weighted bootstrap proposed by Newton and Raftery (1994) instead of the ordinary bootstrap because of the heavy censoring, and the complicated data structure. The specific procedure for such a bootstrap is described as follows.

1. Simulate random weights W_i and W_{jk} from a positive distribution with the property of $\mathrm{Var}(W) = [\mathbf{E}(W)]^2$, where W_i is the random weight for reported failure unit i, and W_{jk} is the random weight of the not-reported unit k in batch j by the DFD. We sample W_i and W_{jk} from the exponential distribution with a mean of one.

2. Based on the random weights W_i and W_{jk}, we calculate the random weighted likelihood

$$\mathcal{L}^*(\boldsymbol{\theta}|\text{ DATA}) = \sum_{i=1}^{r} W_i \log(\pi_k) + \sum_{j=1}^{J}\sum_{k=1}^{w_j} W_{jk}\log(\xi_j). \tag{2.11}$$

3. Obtain the estimated $\widehat{\boldsymbol{\theta}}^*$ by maximizing (2.11).

4. Repeat the above steps B times to obtain the bootstrap samples $\widehat{\boldsymbol{\theta}}^{*b}, b = 1, 2, \ldots, B$.

Following Lawless and Fredette (2005), we construct a calibrated PI for N by using the following steps.

1. Simulate $N^{*b} \sim F_N(n, \widehat{\boldsymbol{\theta}}), b = 1, 2, \ldots, B$, where $\widehat{\boldsymbol{\theta}}$ is the vector of ML estimates from the original data.

2. Compute $v^{*b} = F_N(N^{*b}, \widehat{\boldsymbol{\theta}}^{*b}), b = 1, 2, \ldots, B$.

3. Calculate the lower, and upper $\alpha/2$ quantiles of $\{v^{*b}, b = 1, 2, \ldots, B\}$, denoted by $v_{\alpha/2}$, and $v_{1-\alpha/2}$, respectively.

4. Solve $\underset{\sim}{N}$ from $F_N(\underset{\sim}{N}, \widehat{\boldsymbol{\theta}}) = v_{\alpha/2}$, and $\widetilde{N}$ from $F_N(\widetilde{N}, \widehat{\boldsymbol{\theta}}) = v_{1-\alpha/2}$ to obtain the endpoints of the PI.

2.4.4 Prediction Results

In this section, we present the results of point predictions and PIs for the cumulative number of reported failures, based on the assumptions that both the failure-time distribution and the retirement-time distribution are Weibull. Figure 2.5 shows the point predictions for the cumulative number of reported failures, and the corresponding PIs based on $\mathbf{E}(R) = 98$ and $\beta_R = 1.5$. The cumulative number of reported failures is increasing rapidly until 100 months, and then tends to be stable after 150 months. The leveling-off is caused by the fact that more sample units are retired as time passes. Compared to the initial number of units not reported as failures (w = 120,889), the predicted cumulative number of reported failures (estimate of the expected number) is around 70, which is a small number of the units, relative to the number that had been put into service (i.e., less than 0.058%).

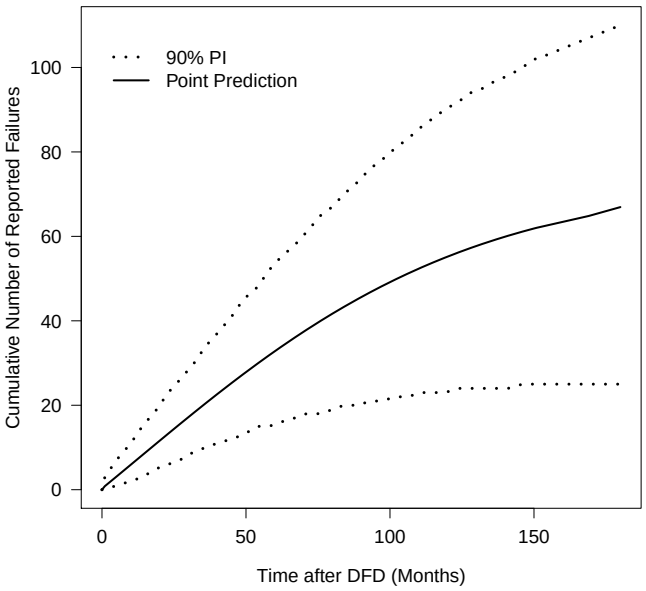


Figure 2.5: Plot of the 90% PI for cumulative number of reported failures based on a Weibull distribution assumption for retirement and failure distributions, with $\mathbf{E}(R) = 98$, and $\beta_R = 1.5$.

2.5 Sensitivity Analysis

The prediction of the cumulative number of reported failures is based on uncertain assumptions including the failure-time distribution as well as the parameters and the distribution for retirement times. Changes in these assumptions will affect the prediction results. Thus, it is necessary to do the sensitivity analysis to assess the effect of departures from the assumptions, and to understand which assumptions are conservative.

2.5.1 Parameter Assumptions

The prediction results shown in Figure 2.5 are based on the Weibull distribution for the retirement model with $\mathbf{E}(R) = 98$, and $\beta_R = 1.5$. Historical information suggests that the values of parameters in the retirement-time distribution are within a certain range. Thus, it is desirable to consider other retirement-time model parameters to assess the effect that deviations from the assumptions have on the prediction results. Table 2.5, and Figure 2.6

Table 2.5: ML estimates and standard errors for model parameters.

Assumed Retirement-time Distribution Parameters			Failure-time Distribution Estimates				
$\mathbf{E}(R)$	η_R	β_R	$\widehat{\eta}_T$	$\widehat{\beta}_T$	se$(\widehat{\eta}_T)$	se$(\widehat{\beta}_T)$	$-$Log likelihood
85	94.157	1.5	1390.523	2.928	555.691	0.436	436.927
90	99.696	1.5	1501.248	2.868	623.215	0.432	436.976
98	108.558	1.5	1670.901	2.788	730.451	0.428	437.047
85	95.912	2	1340.798	2.995	533.781	0.449	436.736
90	101.554	2	1486.736	2.908	622.556	0.443	436.805
98	110.581	2	1712.534	2.796	766.316	0.435	436.908

summarize the results of this sensitivity analysis. The standard errors of $\widehat{\eta}_T$, and $\widehat{\beta}_T$ are denoted as $se(\widehat{\eta}_T)$, and $se(\widehat{\beta}_T)$, respectively. From Table 2.5, we note that, as expected, the ML estimates $\widehat{\eta}_T$ and $\widehat{\beta}_T$ change under different assumptions for the retirement distribution parameters. There is more change in the estimates of $\widehat{\eta}_T$ because these estimates involve a substantial amount of extrapolation. The maximum log-likelihood values are close to each other, indicating there is little or no information about the retirement distribution parameters in the data. Figure 2.6 shows corresponding predictions for the cumulative number of reported failures. The graph indicates that, as the expected retirement time increases, or the Weibull shape parameter decreases (implying more spread in the retirement times), the predicted cumulative number of reported failures increases. The extension of retirement time leads to a higher risk of having failures, but the difference is not large given the same shape parameter. Compared to the change of expected retirement time, predictions are more sensitive to the assumption about the Weibull shape parameter. This difference is the result of the large change of the standard deviation in the retirement time. For example, the standard deviations for $\beta_R = 1.5$, and $\beta_R = 2$ given $\mathbf{E}(R) = 85$ are 57.57, and 44.4, respectively.

2.5.2 Distributional Assumptions

In the previous analysis, the retirement and the failure-time distributions were assumed to be Weibull. The data, however, do not provide much information to distinguish among competing distributions. Thus, it is useful to compute predictions with different retirement-time

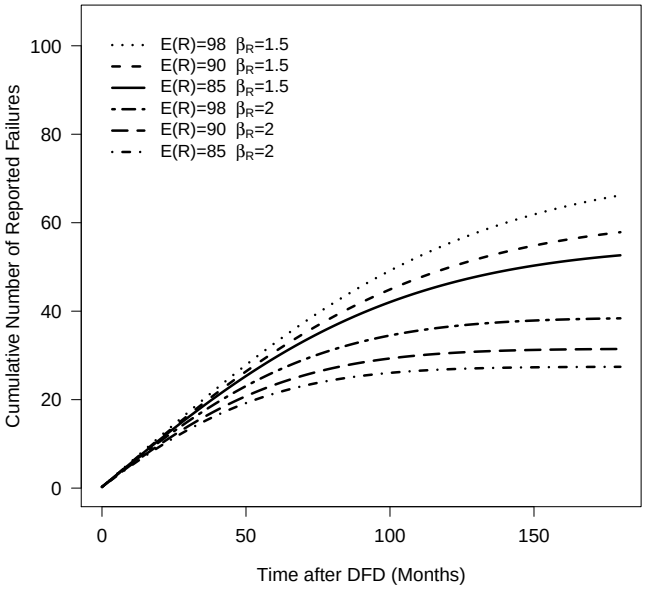


Figure 2.6: Weibull distribution point predictions for the cumulative number of reported failures with different values for the parameters in the retirement distributions.

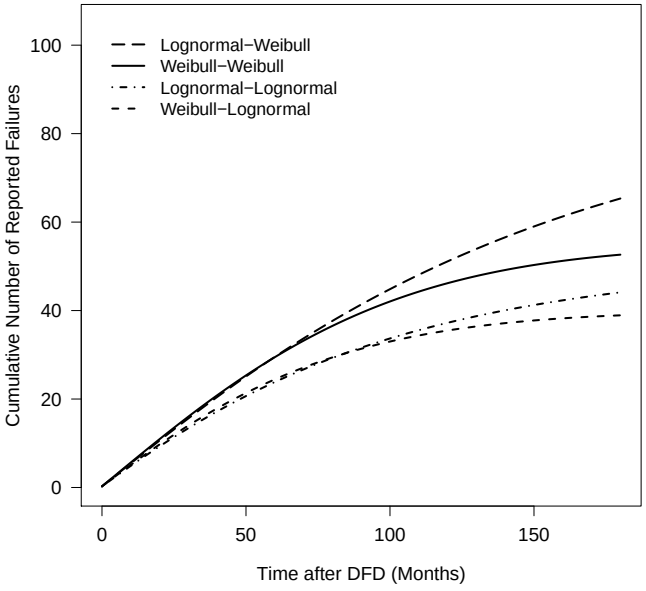


Figure 2.7: The predicted number of the reported failures based on different failure-time and retirement-time distributions when $\mathbf{E}(R) = 85$, and $\mathbf{SD}(R) = 57.7$. The legend indicates the distribution of retirement time, and the distribution of the failure time, respectively.

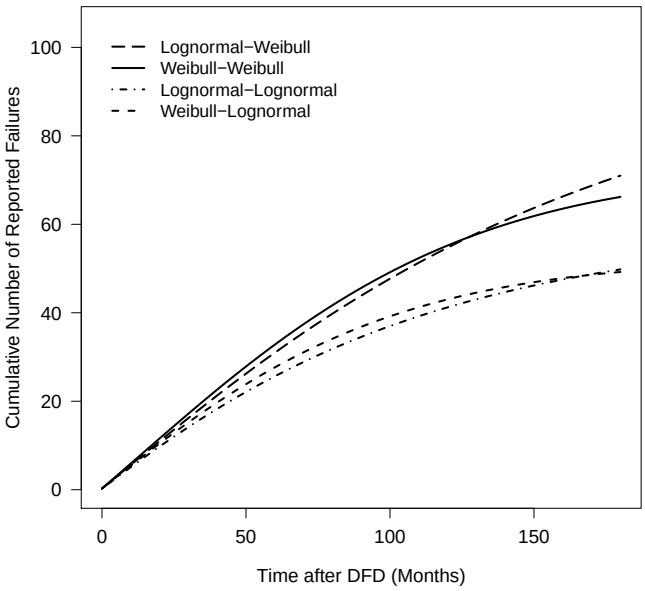


Figure 2.8: The predicted number of the reported failures based on different failure-time and retirement-time distributions when $\mathbf{E}(R) = 98$, and $\mathbf{SD}(R) = 66.5$. The legend indicates the distribution of retirement time, and the distribution of the failure time, respectively.

distribution and failure-time distribution assumptions. In this sensitivity analysis, we use the lognormal distribution as an alternative for the failure-time distribution and the retirement-time distribution. When the lognormal distribution is chosen as the retirement-time distribution, the assumed mean and standard deviation are specified to be the same as that assumed for the Weibull retirement-time distribution.

Table 2.6 shows the ML estimates for the failure-time distribution parameters when $\mathbf{E}(R) = 85$, and $\mathbf{SD}(R) = 57.7$, for the four combinations of the failure-time and retirement distributions. Instead of comparing values of $\widehat{\mu}_T$, we compare the values of $\log(\widehat{t}_{0.001})$ because with the adjustment of retirement distribution the expected lifetime is not meaningful in practice (most of the products have been retired by $\exp(\widehat{\mu}_T)$). From the results in Table 2.6, the maximum log-likelihood values are quite close for all four combinations. When the failure-time distribution is the same, there is no practical difference in the values of $\log(\widehat{t}_{0.001})$, and $\widehat{\sigma}_T$. Table 2.7 shows similar results when $\mathbf{E}(R) = 98$, and $\mathbf{SD}(R) = 66.5$.

Figure 2.7 shows the predicted cumulative number of reported failures with differen-

Table 2.6: ML estimates for parameters when $\mathbf{E}(R) = 85, \mathbf{SD}(R) = 57.7$.

Assumed Retirement-time Distribution	Assumed Failure-time Distribution	$\log(\widehat{t}_{0.001})$	$\widehat{\sigma}_T$	$-$Log likelihood
Weibull	Weibull	4.879	0.341	436.927
Weibull	Lognormal	4.899	1.292	436.640
Lognormal	Weibull	4.838	0.319	436.801
Lognormal	Lognormal	4.856	1.210	436.572

Table 2.7: ML estimates for parameters when $\mathbf{E}(R) = 98, \mathbf{SD}(R) = 66.5$.

Assumed Retirement-time Distribution	Assumed Failure-time Distribution	$\log(\widehat{t}_{0.001})$	$\widehat{\sigma}_T$	$-$Log likelihood
Weibull	Weibull	4.944	0.359	437.047
Weibull	Lognormal	4.969	1.364	436.735
Lognormal	Weibull	4.914	0.339	436.878
Lognormal	Lognormal	4.939	1.291	436.626

t retirement-time and failure-time distributions. The legend shows the failure-time and retirement-time distribution combinations. For example, Lognormal-Weibull indicates that retirement times follow a lognormal distribution, and failure times follow a Weibull distribution. Figure 2.8 shows similar comparisons under a different set of values for $\mathbf{E}(R)$, and $\mathbf{SD}(R)$. Given the data follow the same failure-time distribution, the predictions before 120 months do not depend strongly on assumptions about the retirement-time distribution. In other words, the effect of the failure-time distribution is much stronger than the effect of the retirement-time distribution before 120 months. Compared to other distribution combinations, a lognormal retirement distribution with a Weibull failure-time distribution provides the most conservative predictions (i.e., predicts more reported failures).

2.6 Conclusions, and Areas for Future Research

In this chapter, we provide general statistical methods to predict field failures, and conduct a risk assessment for products. To generate accurate predictions, the proposed method con-

siders the effect of retirement times and reporting delays when estimating the failure-time distribution, and when making predictions. Based on the failure-time model, we predict the cumulative number of reported failures, and construct a corresponding PI. We also conduct sensitivity analysis to assess the effect of different failure-time and retirement-time distributions. The proposed methods can also apply to other highly reliable products such as aircraft engine components, and medical devices, if the data have similar structures (i.e., with retirement or reporting delay).

Here are some possible areas for future research.

- In the product B application, the degradation process that causes the serious failures was related to product age (not the amount of use), and was unrelated to customer-perceivable performance degradation. Thus it was reasonable to assume that the retirement-time and failure-time random variables were s-independent. In other applications, a model that allows dependency could be used.

- For most products, it is not possible to track the retirement time for all units. For some applications, it would be possible and useful to track a representative subset of the product populations through a carefully designed field tracking study.

- Today, some products, even home appliances, can be connected to the internet, potentially providing detailed information about how each such unit is being used. See Hong and Meeker (2010), and Hong and Meeker (2013) for applications involving the prediction of future failures when a proportion of units in the product population are connected to the Internet. Having the additional information about which units are still in active use would reduce much of the uncertainty in predictions associated with a risk analysis.

- For the product B application, there was only limited information about the retirement-time distribution, based on a completely separate marketing study for a similar product. An alterative analysis could have taken that information, perhaps supplemented by expert opinion, to develop a joint prior distribution to describe unknown characteristics

(including the form and the parameters) of the retirement-time and failure-time distributions. This alternative would allow a fully Bayesian analysis to be performed.

Acknowledgments

The authors would like to thank the associate editor and two referees for their valuable comments that helped in improving this paper. The authors acknowledge Advanced Research Computing at Virginia Tech for providing computational resources. The work by Xu and Hong was partially supported by the National Science Foundation under Grant CMMI-1068933 to Virginia Tech and the 2011 DuPont Young Professor Grant.

Bibliography

M. J. Crowder. *Classical Competing Risks*. Chapman & Hall/CRC, Florida, Boca Raton, 2001.

M. Engehardt and L. J. Bain. Prediction intervals for the Weibull process. *Technometrics*, 20:167–169, 1978.

L. A. Escobar and W. Q. Meeker. Statistical prediction based on censored life data. *Technometrics*, 41:113–124, 1999.

S. Geisser. *Predictive Inference: An Introduction*. Chapman and Hall, New York, 1993.

Y. Hong. On computing the distribution function for the Poisson binomial distribution. *Computational Statistics and Data Analysis*, 59:41–51, 2013.

Y. Hong and W. Q. Meeker. Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics*, 52:148–159, 2010.

Y. Hong and W. Q. Meeker. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics*, 55(2):135–149, 2013.

Y. Hong, W. Q. Meeker, and L. A. Escobar. The relationship between confidence intervals for failure probabilities and life time quantile. *IEEE Transactions on Reliability*, 57:260–266, 2008.

Y. Hong, W. Q. Meeker, and J. D. McCalley. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics*, 3:857–879, 2009.

J. F. Lawless and M. Fredette. Frequentist prediction intervals and predictive distributions. *Biometrika*, 92:529–542, 2005.

R. W. Mee and D. Kushary. Prediction limits for the Weibull distribution utilizing simulation. *Computational Statistics & Data Analysis*, 17:327–336, 1994.

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

F. S. Menezes, M. J. F. Vivanco, and L. C. Sampaio. Determination of prediction intervals for a future number of failures: A statistical and Monte Carlo approach. *Brazilian Journal of Physics*, 36:690–699, 2006.

W. Nelson. *Applied Life Data Analysis*. John Wiley & Sons, New York, 1982.

W. Nelson. Weibull prediction of a future number of failures. *Quality and Reliability Engineering International*, 16:23–26, 2000.

M. A. Newton and A. E. Raftery. Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society: Series B*, 56:3–48, 1994.

D. J. Nordman and W. Q. Meeker. Weibull prediction intervals for a future number of failures. *Technometrics*, 44:15–23, 2002.

C. Park and K. B. Kulasekera. Parametric inference of incomplete data with competing risks among several groups. *IEEE Transactions on Reliability*, 53:11–21, 2004.

G. Tian, M. Tang, and J. Yu. Bayesian estimation and prediction for the power law process with left-truncated data. *Journal of Data Science*, 9:445–470, 2011.

G. Yang. Accelerated life test plans for predicting warranty cost. *IEEE Transactions on Reliability*, 59:628–634, 2010.

K. Zhao, D. Steffey, and J. Loud. Incorporating product retirement in field performance reliability analysis. *Reliability and Maintainability Symposium (RAMS), 2010 Proceedings - Annual*, pages 1–5, 2010.

Chapter 3 Nonlinear General Path Models for Degradation Data with Dynamic Covariates

Abstract

Degradation data have been widely used to estimate product reliability. Due to technology advancement, time-varying usage and environmental variables, called dynamic covariates, can be easily recorded, in addition to the traditional degradation measurements. The use of dynamic covariates is appealing because they have the potential to explain more variability in degradation paths. We propose a class of general path models to incorporate dynamic covariates for modeling of degradation paths. Physically-motivated nonlinear functions are used to describe the degradation paths and random effects are used to describe unit-to-unit variability. The covariate effects are modeled by shape-restricted splines. The estimation of unknown model parameters is challenging due to the involvement of nonlinear relationships, random effects and shaped-restricted splines. We develop an efficient procedure for parameter estimations. The performance of the proposed method is evaluated by simulations. An outdoor coating weathering dataset is used to illustrate the proposed method.

Key Words: Dynamic covariates, Organic coatings, Random effects, Reliability, Shape-restricted splines, Time to failure.

3.1 Introduction

3.1.1 Background

Reliability information is important for many business and industrial applications. Degradation data, which consist of repeated measurements of system deterioration over time, are widely used to obtain products and systems' reliability information (e.g., Meeker and Escobar, 1998). In degradation data analysis, degradation measurements are related to failures of product units. That is, when the degradation measurement of a unit reaches a threshold, the unit is defined as a failure. This type of failure is sometimes referred to as "soft failures" in literature. For example, the tread depth of an automobile tire decreases over time due to wear. When the tread depth of a tire reaches a certain threshold, a failure occurs and the tire needs to be replaced for reasons such as safety concerns.

Most products deployed in the field are exposed to dynamic environments, time-varying loads, and other controlling variables. Those dynamic covariates, such as temperature and usage history, can also affect the lifetime of units in the field. Thus, in addition to degradation measurements, dynamic covariates can also provide useful information for reliability analysis. Vaca-Trigo and Meeker (2009) studied the degradation of outdoor coatings, which is mainly caused by outdoor ultraviolet (UV) radiation, but is also affected by other covariates such as temperature (Temp.) and relative humidity (RH). For another example, Zhang et al. (2014) considered the degradation of heaters in ingot manufacturing. When the resistivity of a heater keeps increasing, the heater can not provide a stable and desired thermal field in the ingot growth process, leading to a large variation of ingot diameters. Thus, if the resistivity of a heater passes a certain threshold, the heater is defined as a failure and it needs to be replaced. The degradation of heaters is caused by accumulations of oxide inside furnaces, which is affected by covariates such as power, purity of gas and usage history. With the wide deployment of sensors and advancement of wireless technology, degradation data with dynamic covariates are readily available for modeling and analysis in both products and

industrial equipment in the field.

Most of the existing literature focuses on the modeling of degradation data without covariates or with time-invariant covariates. The use of dynamic covariates is appealing because they can potentially explain more variabilities in the degradation data. However, there are also challenges. First, physically-motivated and flexible models are needed to incorporate dynamic covariates into degradation paths. Second, dynamic covariates often show some specific effects on degradation processes. For example, a larger amount of UV exposure often leads to larger damage in the coating, indicating a monotonic increasing effect of UV on the degradation.

There are two main types of models for describing degradation data: the stochastic models and the general path models. The stochastic models assume the degradation follows a stochastic process such as the gamma process (e.g., Padgett and Tomlinson, 2002). The general degradation path models use a parametric form, usually motivated by the degradation mechanism, to describe the degradation path. In this chapter, we focus on general path models and propose a class of general path models to incorporate dynamic covariates for modeling of degradation paths. Physically-motivated nonlinear functions, which are more common in degradation mechanism (e.g., Meeker and LuValle, 1995, and Vaca-Trigo and Meeker, 2009), are used to describe the degradation paths. Random effects are used to describe unit-to-unit variability. Shape-restricted splines are applied to model the effects of dynamic covariates. We develop an efficient procedure to estimate the unknown model parameters and apply the proposed statistical method on the outdoor coating weathering data in Gu et al. (2009).

3.1.2 Related Literature and Contribution of This Work

Lu and Meeker (1993), and Meeker and Escobar (1998) described general path models for estimating failure-time distribution for degradation data. Degradation models based on stochastic processes were described in Whitmore (1995), Padgett and Tomlinson (2004), Park and Padgett (2005), and Ye and Chen (2013). To model the unit-to-unit variability, random effects were added into the degradation model (e.g., Lawless and Crowder, 2004, Bae and Kvam, 2004, and Yuan and Pandey, 2009). Wakefield (1996), Robinson and Crowder (2000),

and Shi and Meeker (2012) applied Bayesian methods to the degradation data. Bagdonavičius et al. (2004), and Bagdonavičius et al. (2005) introduced methods to model the multiple failure modes in degradation data. In the area of degradation with covariates, Bagdonavičius and Nikulin (2001), and Lawless and Crowder (2004) used stochastic processes to estimate the lifetime distribution.

Hong et al. (2014) incorporated dynamic covariates into the degradation using a linear random effects model. Here, we summarize the major advantages of the nonlinear random effects model in this chapter over the linear random effects model in Hong et al. (2014). Nonlinear degradation models are usually motivated by physical knowledge (e.g., kinetic theory, Gu et al., 2009). Thus the nonlinear model is closely related to the true degradation mechanism and has better interpretability. The linear model for the degradation path, however, lacks of physically interpretations in general. Degradation paths often show monotonicity. The linear model in Hong et al. (2014) can not guarantee the degradation path is always monotone increasing or decreasing. The nonlinear model can guarantee the monotonicity of the degradation path. As shown in the application, the nonlinear model can provide better model fitting to real data than the linear model. Thus, the nonlinear model is more flexible and can provide a better description of the degradation path.

The generalization from linear random effects model to nonlinear random effects model is non-trivial, especially for the parameter estimation. The estimation of unknown parameters in this chapter involves shape restrictions, random effects, and unknown parameters in the nonlinear function, which is technically challenging. We develop a novel estimation procedure, which integrates the linearization technique, the mixed primal-dual algorithm (Fraser and Massam, 1989), and the alternating algorithm (Lindstrom and Bates, 1990). The developed procedure can estimate unknown parameters efficiently and simulations show that its performance is good.

3.1.3 Overview

The rest this chapter is organized as follows. In Section 3.2, we introduce the notation for degradation data with dynamic covariates and the nonlinear random effects degradation model. In Section 3.3, we give the details of modeling cumulative exposure function based on shape-restricted splines. In Section 3.4, we present the likelihood function for the nonlinear random effects model and procedures for parameter estimations and statistical inferences. In Section 3.5, the proposed method is evaluated by simulations. In Section 3.6, we give an application of the proposed method on an outdoor weathering dataset. Section 3.7 contains conclusions and some areas for future work.

3.2 The Data and Model

3.2.1 Data

Let $y_i(t_{ij}), i = 1, \cdots, n,\ j = 1, \cdots, n_i$ be the observed degradation measurement for unit i at time t_{ij}, where n is the number of units and n_i is the number of measurement points for unit i. For the dynamic covariates, we define $X_i(t) = [X_{i1}(t), \cdots, X_{ip}(t)]'$ as a vector of dynamic covariate observations, p as the number of dynamic covariates, and $X_{ih}(t), h = 1, \cdots, p$ as the value of h^{th} dynamic covariate for unit i at time t. Define $\boldsymbol{X}_i(t) = \{X_i(s), 0 \leq s \leq t\}$ as the history of dynamic covariate processes for unit i from time 0 to time t. Let $\boldsymbol{X}(t) = \{\boldsymbol{X}_1(t), \cdots, \boldsymbol{X}_n(t)\}$ be the history of dynamic covariate processes from time 0 to time t for all units.

3.2.2 Nonlinear Degradation Path Model with Dynamic Covariates

In this chapter, we focus on nonlinear random effects degradation path models. The model for the degradation measurements is expressed as

$$y_i(t_{ij}) = f\big(D[\boldsymbol{X}_i(t_{ij}); \boldsymbol{\beta}_D]; \boldsymbol{\beta}_f, \boldsymbol{b}_i\big) + \epsilon_{ij}, \quad i = 1, \cdots, n,\ j = 1, \cdots, n_i, \tag{3.1}$$

where $f(\cdot;\boldsymbol{\beta}_f, \boldsymbol{b}_i)$ is a nonlinear degradation function with deterministic parameters $\boldsymbol{\beta}_f$ and unit specific parameter $\boldsymbol{b}_i$. The term $D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D]$ in (3.1) is the cumulative exposure function that depends on covariate history $\boldsymbol{X}_i(t_{ij})$, where $\boldsymbol{\beta}_D$ is a vector of parameters in $D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D]$. More details on cumulative exposure function will be discussed in Section 3.3. We denote $\boldsymbol{\beta} = (\boldsymbol{\beta}'_D, \boldsymbol{\beta}'_f)'$. The unit specific parameter $\boldsymbol{b}_i$, which is referred to as random effects, is used to describe unit-to-unit variability in degradation paths. We use a multivariate normal distribution $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma_b})$ to describe the random effects $\boldsymbol{b}_i$, where $\mathbf{0}$ is a zero vector and $\boldsymbol{\Sigma_b}$ is the variance-covariance matrix of $\boldsymbol{b}_i$. The error term ϵ_{ij} follows a normal distribution $\mathrm{N}(0, \sigma^2)$.

3.2.3 Commonly Used Functional Forms for Degradation Path

Motivated by physical or chemical degradation mechanisms, some functional forms are commonly used for the degradation path form $f(\cdot)$ in (3.1). Here we briefly describe some commonly used functional forms. The proposed procedure, however, can be applied to any forms. Meeker and LuValle (1995) suggested several statistical kinetic models to describe the growth of filaments in a circuit board. Based on the rate equation of first order reaction, the growth amount at time t is

$$a(t) = a(\infty) \times [1 - \exp(-kt)], \tag{3.2}$$

where $a(t)$ is the growth amount of products at time t, $a(\infty)$ is the maximum amount of products at $t = \infty$ and k is the reaction rate. The maximum amount of products $a(\infty)$ can be random from unit to unit, thus a random effect can be used to model $a(\infty)$.

To describe the damage from UV exposure, Temp., and RH, Vaca-Trigo and Meeker (2009) suggested an empirical model for the coating degradation process. The degradation at time t is

$$a(t) = \frac{a(\infty) - a(0)}{1 + \exp(z)}, \tag{3.3}$$

where $z = \{\log[d(t)] - \xi\}/\zeta$, $d(t)$ is the cumulative dosage effect of UV exposure at time t, ξ is the location parameter, and ζ is the degradation steepness level. Under this model, the initial

amount of degradation is $a(0)$ and the asymptotic degradation amount is $a(\infty)$. The range of responses from this model is between $a(0)$ and $a(\infty)$. To account for unit-to-unit variability, random effects can be used to describe the total amount of degradation $a(\infty) - a(0)$ and the degradation steepness ζ.

Inverse power relationship is also widely used in describing the relationship between degradation and voltage stress. The degradation model is $a(t) = (kt)^{1/\zeta}$, where k is a function of voltage and ζ is a parameter. Some other generalized acceleration models are available in Park and Padgett (2006). To further model the individual variability, random effects can be added into those nonlinear models.

Constant covariates (i.e., time-invariant covariates) can be incorporated into the degradation process. For example, temperature can be added into model (3.2) to describe the reaction rate k by using the Arrhenius equation (e.g., Escobar and Meeker, 2006). However, there is little work that considers degradation path model with dynamic covariates.

3.3 Modeling Cumulative Exposure Function

3.3.1 Cumulative Exposure Function for Dynamic Covariates

The cumulative exposure function $D[\boldsymbol{X}_i(t_{ij}); \boldsymbol{\beta}_D]$ in (3.1) describes the cumulative effect of dynamic covariates on unit i up to time t_{ij}. The function is a nondecreasing function of time due to accumulative effects. The cumulative exposure is closely related to the lifetime of the unit. When the cumulative exposure becomes large, the degradation level will pass the critical value, and a failure will occur. The cumulative exposure is obtained by accumulating instantaneous exposure, $g[X_i(t_{ij}); \boldsymbol{\beta}_D]$, from time 0 to time t_{ij}. That is,

$$D[\boldsymbol{X}_i(t_{ij}); \boldsymbol{\beta}_D] = \int_0^{t_{ij}} g[X_i(u); \boldsymbol{\beta}_D]\, du = \int_0^{t_{ij}} \sum_{h=1}^{p} g_h[X_{ih}(u); \boldsymbol{\beta}_D]\, du,$$

where $g_h[X_{ih}(t); \boldsymbol{\beta}_D], h = 1, \cdots, p$, is the instantaneous exposure function for the h^{th} dynamic covariate at time t. The cumulative exposure can be considered as an acceleration or deceleration of the usual time scale. For example, if a covariate $x(t)$ is time invariant (i.e., $x(t) = x$ for

all t), then $D[\boldsymbol{X}_i(t); \boldsymbol{\beta}_D] = g(x; \boldsymbol{\beta}_D)t$, where $g(x; \boldsymbol{\beta}_D)$ is the acceleration/deceleration factor of the usual time scale t.

Figure 3.1 illustrates the relationships among dynamic covariates, instantaneous exposure, cumulative exposure, and degradation. Figure 3.1(a) shows two different trends of dynamic covariate process: a constant one and an increasing one. Figure 3.1(b) shows a monotonic increasing covariate effect function $g(x) = 1.2\exp(x)$, where x is the value of dynamic covariate. At a particular time point t_e, the values of those covariates are x_0 and x_1 for the two covariate profiles, respectively. According to the effect function, the instantaneous effects are g_0 and g_1, as marked in Figure 3.1(b). Figure 3.1(c) shows the corresponding instantaneous exposure for the two covariate profiles as a function of time t. If we accumulate the instantaneous exposure functions over time, the cumulative exposure functions are obtained, which are displayed in Figure 3.1(d). The corresponding cumulative exposures at time t_e are d_0 and d_1, respectively. On the cumulative exposure scale, the degradation path is plotted on Figure 3.1(e), which is determined by the nonlinear degradation path function $f(D) = -3/[1 + \exp(1.7 - 0.008D)]$. Figure 3.1(f) shows the degradation paths which are functions of time. Due to the two different covariate profiles, the degradation paths are different on the time scale t. The unit exposed to the increasing covariate profile degrades much faster than the constant one, showing the effect of covariate process on the degradation process.

3.3.2 Computation of Cumulative Exposure

In practice, the functional form of exposure effect function $g(\cdot)$ is usually unknown. Physical or engineering knowledge, however, may suggest some particular shape of the effect function (e.g., monotone increasing effect function). Thus, we use shape-restricted splines to model the exposure effect function, which can model $g(\cdot)$ flexibly but also retain meaningful physical interpretations. Introduction of the shape-restricted splines is available in Meyer (2008). For example, an increasing function is obtained by a linear combination of I-spline basis functions with nonnegative coefficients and a constant function. A convex function is obtained by a linear combination of C-spline basis functions with nonnegative coefficients, a constant function, and

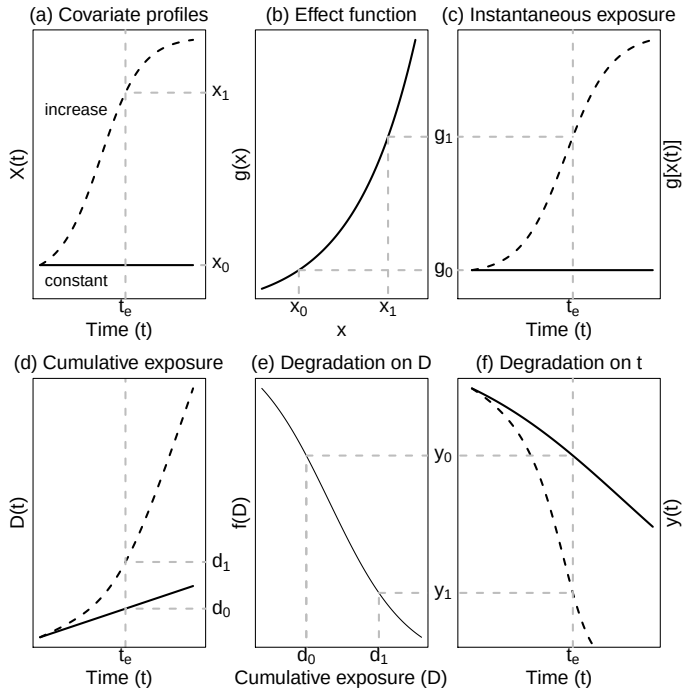


Figure 3.1: Illustration of relationships among dynamic covariates, instantaneous exposure, cumulative exposure and degradation.

an identity function.

The shape-restricted spline basis can be computed based on the dynamic covariates $\boldsymbol{X}(t)$ (e.g., see Figure 3.2). Define $w_{hv}[X_i(t_{ij})]$ as the value of v^{th} spline at time t_{ij} for h^{th} dynamic covariate, where $h = 1, \cdots, p$, and $v = 1, \cdots, V_h$. The number V_h depends on the shape of the effect function of the dynamic covariate, the order k, and the number of knots q of M-splines. If the effect function of the h^{th} dynamic covariate is monotone, then $V_h = 1 + q + k$, where "1" is for the constant function and "$q + k$" is the number of I-spline basis. If the effect function of the h^{th} dynamic covariate is convex, then $V_h = 1 + 1 + q + k$, where the first "1" is for the constant function, the second "1" is for the identity function and "$q + k$" is the number of C-spline basis. The coefficients of spline basis are defined as $\boldsymbol{\beta}_D$ which is a collection of β_{hv}. Some of the β_{hv} are restricted to be nonnegative due to the requirement of the shapes effect functions of dynamic covariates. Then, the instantaneous exposure function can be expressed as $g[X_i(t_{ij}); \boldsymbol{\beta}_D] = \sum_{h=1}^{p} \sum_{v=1}^{V_h} w_{hv}[X_i(t_{ij})]\beta_{hv}$. The cumulative exposure, $D[\boldsymbol{X}_i(t_{ij}); \boldsymbol{\beta}_D]$, can

be computed by

$$D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D] = \sum_{u\le t_{ij}} g[X_i(u);\boldsymbol{\beta}_D]\delta_{ij} = \sum_{u\le t_{ij}}\sum_{h=1}^{p}\sum_{v=1}^{V_h} w_{hv}[X_i(u)]\beta_{hv}\delta_{ij},$$

where $\delta_{ij} = t_{ij} - t_{i,j-1}$ and $u \in \{t_{i1}, \cdots, t_{ij}\}$. Here, $t_{i0} = 0$. To simplify the computation, define $W_{hv}(t_{ij}) = \sum_{u\le t_{ij}} w_{hv}[X_i(u)]\delta_{ij}$. Then, $D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D] = \sum_{h=1}^{p}\sum_{v=1}^{V_h} W_{hv}(t_{ij})\beta_{hv}$. Using matrix notation, the cumulative spline basis is expressed as

$$\boldsymbol{W}_i = \begin{bmatrix} W_{11}(t_{i1}) & \cdots & W_{1V_1}(t_{i1}) & \cdots & W_{p1}(t_{i1}) & \cdots & W_{pV_p}(t_{i1}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ W_{11}(t_{in_i}) & \cdots & W_{1V_1}(t_{in_i}) & \cdots & W_{p1}(t_{in_i}) & \cdots & W_{pV_p}(t_{in_i}) \end{bmatrix}.$$

Then, the cumulative exposure for unit i is $\boldsymbol{D}_i = \boldsymbol{W}_i\boldsymbol{\beta}_D$. The jth element of $\boldsymbol{D}_i$ is denoted by D_{ij} and note that $D_{ij} = D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D]$, $j = 1, \cdots, n_i$.

3.4 Parameter Estimation Procedures

3.4.1 The Likelihood Function and Challenges

Let $\boldsymbol{y}_i = [y_i(t_{i1}), \cdots, y_i(t_{in_i})]'$, $\boldsymbol{f}_i(\boldsymbol{W}_i;\boldsymbol{\beta},\boldsymbol{b}_i) = [f(D_{i1};\boldsymbol{\beta}_f,\boldsymbol{b}_i), \cdots, f(D_{in_i};\boldsymbol{\beta}_f,\boldsymbol{b}_i)]'$, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \cdots, \epsilon_{in_i})'$, $i = 1, 2, \cdots, n$. Note that $\boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Lambda}_i)$. Here, $\boldsymbol{\Lambda}_i$ is the variance-covariance matrix with $\sigma^2\mathbf{I}$ as a special case, where $\mathbf{I}$ is an identity matrix. The nonlinear random effects degradation path model can be re-expressed as

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{W};\boldsymbol{\beta},\boldsymbol{b}) + \boldsymbol{\epsilon}, \tag{3.4}$$

where

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_n \end{bmatrix}, \quad \boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1 \\ \vdots \\ \boldsymbol{W}_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}, \quad \boldsymbol{f}(\boldsymbol{W};\boldsymbol{\beta},\boldsymbol{b}) = \begin{bmatrix} \boldsymbol{f}_1(\boldsymbol{W}_1;\boldsymbol{\beta},\boldsymbol{b}_1) \\ \vdots \\ \boldsymbol{f}_n(\boldsymbol{W}_n;\boldsymbol{\beta},\boldsymbol{b}_n) \end{bmatrix},$$

and $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_1', \cdots, \boldsymbol{\epsilon}_n']'$. Here, $\boldsymbol{b} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Lambda})$, where $\boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\Sigma}_{\boldsymbol{b}}, \cdots, \boldsymbol{\Sigma}_{\boldsymbol{b}})$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{\Lambda}_1, \cdots, \boldsymbol{\Lambda}_n)$. Define $\boldsymbol{\sigma}_b$ and $\boldsymbol{\sigma}_e$ as the unique parameters in matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$, respectively.

Let $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\sigma}_e, \boldsymbol{\sigma}_b\}$ be the collection of unknown parameters in the nonlinear random effects model (3.4). The maximum likelihood (ML) method is used to estimate the unknown

model parameters. Given the observed covariate processes, the likelihood function is

$$\mathbf{L}(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{W}) \propto \prod_{i=1}^{n} \int_{\boldsymbol{b}_i} p(\boldsymbol{y}_i|\boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{W}_i) p(\boldsymbol{b}_i|\boldsymbol{\theta}) d\boldsymbol{b}_i, \tag{3.5}$$

where $p(\boldsymbol{y}_i|\boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{W}_i)$ is the probability density function (pdf) of $\mathrm{N}[\boldsymbol{f}_i(\boldsymbol{W}_i; \boldsymbol{\beta}, \boldsymbol{b}_i), \boldsymbol{\Lambda}_i]$, and $p(\boldsymbol{b}_i|\boldsymbol{\theta})$ is the pdf of $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{b}})$.

The ML estimates of parameters in the general nonlinear random effects model, however, are not easy to obtain because it often requires numerical integration in high dimension. Approximate loglikelihood procedures can provide easy-to-compute alternatives to estimate the unknown parameters, which includes the alternating algorithm (e.g., Lindstrom and Bates, 1990), Laplace's approximation (e.g., Wolfinger, 1993), and adaptive Gaussian quadrature approximation (e.g., Pinheiro and Bates, 1995). For the nonlinear random effects model in (3.4), some parameters are restricted to be nonnegative to guarantee the specific shape of the effect function, which increases the difficulty in parameter estimations. Thus, efficient methods are needed for estimating parameters with constraints.

3.4.2 The Proposed Parameter Estimation Procedure

Here we propose an efficient method to estimate unknown parameters. The procedure iterates between two main steps until it converges, which is called the outside iterations. Those two steps are described in Sections 3.4.2.1 and 3.4.2.2, respectively. In each step of the outside iteration, there are also inside iterations. To simplify the notation, $\boldsymbol{f}(\boldsymbol{W}; \boldsymbol{\beta}, \boldsymbol{b})$ is denoted as $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{b})$ in the rest of the chapter.

3.4.2.1 Estimation of Constrained Parameters

The main objective of this step is to update $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{b}}$, given the current estimates of $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Lambda}}$ (i.e., estimates at the j^{th} outside iteration). Note that some components of $\widehat{\boldsymbol{\beta}}$ are constrained. The mixed primal-dual bases algorithm (Fraser and Massam, 1989) is used for efficient computation.

Let $\widehat{\boldsymbol{\Sigma}}^{(j)}$ and $\widehat{\boldsymbol{\Lambda}}^{(j)}$ be the estimates of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ in j^{th} outside iteration, respectively. Given

$\widehat{\boldsymbol{\Sigma}}^{(j)}$ and $\widehat{\boldsymbol{\Lambda}}^{(j)}$, the updated estimates of $\boldsymbol{\beta}$ and $\boldsymbol{b}$, denoted as $\widehat{\boldsymbol{\beta}}^{(j)}$ and $\widehat{\boldsymbol{b}}^{(j)}$, respectively, are obtained by minimizing a penalized nonlinear least squares objective function. That is,

$$[\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})]'\widehat{\boldsymbol{\Lambda}}^{(j)^{-1}}[\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})]+\boldsymbol{b}'\widehat{\boldsymbol{\Sigma}}^{(j)^{-1}}\boldsymbol{b}. \tag{3.6}$$

Here $\|\boldsymbol{a}\|^2=\boldsymbol{a}'\boldsymbol{a}$ and $\boldsymbol{a}$ is a column vector. To simplify the computation, we introduce pseudo data. Specifically,

$$[\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})]'\widehat{\boldsymbol{\Lambda}}^{(j)^{-1}}[\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})]+\boldsymbol{b}'\widehat{\boldsymbol{\Sigma}}^{(j)^{-1}}\boldsymbol{b}=\|\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})\|^2, \tag{3.7}$$

where

$$\widetilde{\boldsymbol{y}}=\begin{pmatrix}\widehat{\boldsymbol{\Lambda}}^{(j)^{-1/2}}\boldsymbol{y}\\ \mathbf{0}\end{pmatrix}\quad\text{and}\quad\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})=\begin{pmatrix}\widehat{\boldsymbol{\Lambda}}^{(j)^{-1/2}}\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})\\ \widehat{\boldsymbol{\Sigma}}^{(j)^{-1/2}}\boldsymbol{b}\end{pmatrix}.$$

Here, $\widehat{\boldsymbol{\Sigma}}^{(j)^{-1/2}}=\text{diag}[\boldsymbol{M}'^{-1},\cdots,\boldsymbol{M}'^{-1}]$, $\boldsymbol{M}$ is the upper-triangle Cholesky factor of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}$ (i.e., $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}=\boldsymbol{M}'\boldsymbol{M}$), and $\widehat{\boldsymbol{\Lambda}}^{(j)^{-1/2}}$ is the square root matrix of $\widehat{\boldsymbol{\Lambda}}^{(j)^{-1}}$.

We use the linearization technique to handle nonlinear relationships. In particular, we apply the first order Taylor series expansion at the current estimates $\widehat{\boldsymbol{\beta}}^{(j,k)}$ and $\widehat{\boldsymbol{b}}^{(j,k)}$. Define $\widehat{\boldsymbol{\beta}}^{(j,k)}$ and $\widehat{\boldsymbol{b}}^{(j,k)}$ as the estimates in j^{th} outside iteration and k^{th} inside iteration for $\boldsymbol{\beta}$ and $\boldsymbol{b}$, respectively. Let $\widehat{\boldsymbol{\beta}}^{(j,0)}=\widehat{\boldsymbol{\beta}}^{(j-1)}$ and $\widehat{\boldsymbol{b}}^{(j,0)}=\widehat{\boldsymbol{b}}^{(j-1)}$. Define $\widehat{\boldsymbol{Q}}=\frac{\partial\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})}{\partial\boldsymbol{\beta}'}\big|_{\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)}}$, and $\widehat{\boldsymbol{Z}}=\frac{\partial\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})}{\partial\boldsymbol{b}'}\big|_{\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)}}$. After taking the Taylor series expansion at $\widehat{\boldsymbol{\beta}}^{(j,k)}$ and $\widehat{\boldsymbol{b}}^{(j,k)}$, (3.7) can be approximated by

$$\left\|\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{f}}\left(\widehat{\boldsymbol{\beta}}^{(j,k)},\,\widehat{\boldsymbol{b}}^{(j,k)}\right)-\widehat{\boldsymbol{Q}}\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}^{(j,k)}\right)-\widehat{\boldsymbol{Z}}\left(\boldsymbol{b}-\widehat{\boldsymbol{b}}^{(j,k)}\right)\right\|^2. \tag{3.8}$$

Because some β's in $\boldsymbol{\beta}$ are constrained (>0), we redefine $\boldsymbol{\beta}=(\boldsymbol{\beta}_r',\boldsymbol{\beta}_{nr}')'$, where $\boldsymbol{\beta}_r$ represents the parameters restricted to be nonnegative and $\boldsymbol{\beta}_{nr}$ represents parameters without restrictions in $\boldsymbol{\beta}$. The following uses an orthogonal decomposition to show that we can obtain the estimates of $\boldsymbol{\beta}_r$ first.

Let $\widehat{\boldsymbol{Q}}_r=\frac{\partial\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})}{\partial\boldsymbol{\beta}_r'}\big|_{\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)}}$ and $\widehat{\boldsymbol{Q}}_{nr}=\frac{\partial\widetilde{\boldsymbol{f}}(\boldsymbol{\beta},\boldsymbol{b})}{\partial\boldsymbol{\beta}_{nr}'}\big|_{\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)}}$. Equation (3.8) is re-expressed as

$$\begin{aligned}&\left\|\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{f}}(\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)})-\widehat{\boldsymbol{Q}}_r(\boldsymbol{\beta}_r-\widehat{\boldsymbol{\beta}}_r^{(j,k)})\right.\\ &\left.-\widehat{\boldsymbol{Q}}_{nr}\left(\boldsymbol{\beta}_{nr}-\widehat{\boldsymbol{\beta}}_{nr}^{(j,k)}\right)-\widehat{\boldsymbol{Z}}\left(\boldsymbol{b}-\widehat{\boldsymbol{b}}^{(j,k)}\right)\right\|^2.\end{aligned} \tag{3.9}$$

Define $\boldsymbol{X}^*=(\widehat{\boldsymbol{Q}}_{nr},\widehat{\boldsymbol{Z}})$, $\widetilde{\boldsymbol{y}}^*=\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{f}}(\widehat{\boldsymbol{\beta}}^{(j,k)},\widehat{\boldsymbol{b}}^{(j,k)})+\widehat{\boldsymbol{Q}}_r\widehat{\boldsymbol{\beta}}_r^{(j,k)}$, $\boldsymbol{\beta}^*=(\boldsymbol{\beta}_{nr}',\boldsymbol{b}')'$ and $\widehat{\boldsymbol{\beta}}^{*(j,k)}$ as its

corresponding estimate. Based on the new notation, equation (3.9) can be re-expressed to be

$$\begin{aligned}&\left\|\widetilde{\boldsymbol{y}}^{*}-\boldsymbol{X}^{*}\left(\boldsymbol{\beta}^{*}-\widehat{\boldsymbol{\beta}}^{*(j,k)}\right)-\widehat{\boldsymbol{Q}}_{r}\boldsymbol{\beta}_{r}\right\|^{2}\\ =&\left\|\widetilde{\boldsymbol{y}}^{*}-\boldsymbol{X}^{*}\left(\boldsymbol{\beta}^{*}-\widehat{\boldsymbol{\beta}}^{*(j,k)}\right)-\boldsymbol{P}_{\boldsymbol{X}^{*}}\widehat{\boldsymbol{Q}}_{r}\boldsymbol{\beta}_{r}-\left(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{X}^{*}}\right)\widehat{\boldsymbol{Q}}_{r}\boldsymbol{\beta}_{r}\right\|^{2},\end{aligned}\tag{3.10}$$

where $\boldsymbol{P}_{\boldsymbol{X}^{*}}=\boldsymbol{X}^{*}(\boldsymbol{X}^{*'}\boldsymbol{X}^{*})^{-1}\boldsymbol{X}^{*'}$, and $\boldsymbol{I}$ is an identity matrix. The estimation of unknown parameters in nonlinear random effects model is obtained by minimizing the loss function of projecting $\widetilde{\boldsymbol{y}}^{*}$ on two perpendicular spaces (column space of $\boldsymbol{X}^{*}$ and $\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{X}^{*}}$) shown in (3.10). Because the two spaces are orthogonal, parameters with restriction $\boldsymbol{\beta}_{r}$ can be estimated first.

To ensure the estimates of $\boldsymbol{\beta}_{r}$ nonnegative, the mixed primal-dual bases algorithm is applied, which is very efficient. The details of the mixed primal-dual bases algorithm are described in Fraser and Massam (1989). After obtaining the $\widehat{\boldsymbol{\beta}}_{r}^{(j,k+1)}$, we update the unconstrained component $\boldsymbol{\beta}^{*}$. Define $\widetilde{\boldsymbol{y}}^{**}=\widetilde{\boldsymbol{y}}^{*}-\widehat{\boldsymbol{Q}}_{r}\widehat{\boldsymbol{\beta}}_{r}^{(j,k+1)}$. Then, equation (3.7) is approximated by

$$\|\widetilde{\boldsymbol{y}}^{**}-\boldsymbol{X}^{*}(\boldsymbol{\beta}^{*}-\widehat{\boldsymbol{\beta}}^{*(j,k)})\|^{2}.\tag{3.11}$$

Let $\triangle=\boldsymbol{\beta}^{*}-\widehat{\boldsymbol{\beta}}^{*(j,k)}$. Then, the estimate of $\triangle$, denoted as $\widehat{\triangle}$, can be obtained by minimizing equation (3.11). That is, $\widehat{\triangle}=(\boldsymbol{X}^{*'}\boldsymbol{X}^{*})^{-1}\boldsymbol{X}^{*'}\widetilde{\boldsymbol{y}}^{**}$.

If the minimum of absolute values of $\widehat{\triangle}$ is smaller than the defined criteria (e.g., <0.0001), the inside iteration converges. We obtain the estimates $\widehat{\boldsymbol{\beta}}^{(j)}=\widehat{\boldsymbol{\beta}}^{(j,k)}$, and $\widehat{\boldsymbol{b}}^{(j)}=\widehat{\boldsymbol{b}}^{(j,k)}$. If not, let $\widehat{\boldsymbol{\beta}}^{*(j,k+1)}=\widehat{\triangle}+\widehat{\boldsymbol{\beta}}^{*(j,k)}$. The inside iteration continues with estimates $\widehat{\boldsymbol{\beta}}^{(j,k)}$ and $\widehat{\boldsymbol{b}}^{(j,k)}$ being updated to $\widehat{\boldsymbol{\beta}}^{(j,k+1)}$ and $\widehat{\boldsymbol{b}}^{(j,k+1)}$, respectively. Upon convergence, the restricted coefficients $\boldsymbol{\beta}_{r}^{(j)}=\widehat{\boldsymbol{\beta}}_{r}^{(j)}$ will be treated as known parameters, and other estimates $\widehat{\boldsymbol{\beta}}_{nr}^{(j)}$ and $\widehat{\boldsymbol{b}}^{(j)}$ will be used as the initial values in the step in Section 3.4.2.2 .

3.4.2.2 Estimation of $\boldsymbol{\sigma}_{e}$ and $\boldsymbol{\sigma}_{b}$

The main objective of this step is to update $\boldsymbol{\sigma}_{e}$ and $\boldsymbol{\sigma}_{b}$, given $\widehat{\boldsymbol{\beta}}_{r}^{(j)}$. Given $\boldsymbol{\beta}_{r}^{(j)}$, the estimates for $\boldsymbol{\sigma}_{e}$ and $\boldsymbol{\sigma}_{b}$ can be obtained by the alternating algorithm in Lindstrom and Bates (1990). Here, we briefly introduce the alternating algorithm which is realized by two steps.

With $\boldsymbol{\beta}_{r}^{(j)}$ being fixed, we can write $\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{b})$ as $\boldsymbol{f}(\boldsymbol{\beta}_{nr},\boldsymbol{b})$ where $\boldsymbol{\beta}_{nr}$ includes $\boldsymbol{\beta}_{f}$ and other

nonrestricted parameters in the cumulative effect function. Let $\widehat{\boldsymbol{\sigma}}_e^{(j+1)}$, and $\widehat{\boldsymbol{\sigma}}_b^{(j+1)}$ be the estimates for $\boldsymbol{\sigma}_e$ and $\boldsymbol{\sigma}_b$ in $(j+1)^{th}$ outside iteration.

Equation (3.8) now can be solved by using Gauss-Newton method directly. We can use $\widehat{\boldsymbol{\Sigma}}^{(j+1,0)} = \widehat{\boldsymbol{\Sigma}}^{(j)}$ and $\widehat{\boldsymbol{\Lambda}}^{(j+1,0)} = \widehat{\boldsymbol{\Lambda}}^{(j)}$ as the initial estimates for $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ at $(j+1)^{th}$ outside iteration. Given the information of $\widehat{\boldsymbol{\Sigma}}^{(j+1,m)}$ and $\widehat{\boldsymbol{\Lambda}}^{(j+1,m)}$, the estimates $\widehat{\boldsymbol{\beta}}_{nr}^{(j+1,m)}$ and $\widehat{\boldsymbol{b}}^{(j+1,m)}$ can be obtained. We define that $\widehat{\boldsymbol{\beta}}_{nr}^{(j+1,m)}$ and $\widehat{\boldsymbol{b}}^{(j+1,m)}$ are the estimates of $\boldsymbol{\beta}_{nr}$ and $\boldsymbol{b}$ at m^{th} inside iteration and $(j+1)^{th}$ outside iteration. Define $\widehat{\boldsymbol{L}}^{(j+1,m)} = \frac{\partial \boldsymbol{f}(\boldsymbol{\beta}_{nr},\boldsymbol{b})}{\partial \boldsymbol{b}'}\big|_{\widehat{\boldsymbol{\beta}}_{nr}^{(j+1,m)},\widehat{\boldsymbol{b}}^{(j+1,m)}}$.

With the estimates of $\boldsymbol{\beta}_{nr}$ and $\boldsymbol{b}$ updated, we now update the estimates of $\boldsymbol{\sigma}_e$ and $\boldsymbol{\sigma}_b$. Instead of using the exact likelihood function (3.5), the approximate likelihood is used. Using the Taylor expansion, the conditional distribution of $\boldsymbol{y}$ can be approximated by

$$\boldsymbol{y}\big|\boldsymbol{b} \dot{\sim} \mathrm{N}\left(\boldsymbol{f}(\boldsymbol{\beta}_{nr},\widehat{\boldsymbol{b}}^{(j+1,m)}) + \widehat{\boldsymbol{L}}^{(j+1,m)}\boldsymbol{b} - \widehat{\boldsymbol{L}}^{(j+1,m)}\widehat{\boldsymbol{b}}^{(j+1,m)}, \boldsymbol{\Lambda}\right).$$

Based on the assumption of $\boldsymbol{b} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, the approximate marginal distribution of $\boldsymbol{y}$ is

$$\boldsymbol{y} \dot{\sim} \mathrm{N}\left(\boldsymbol{f}(\boldsymbol{\beta}_{nr},\widehat{\boldsymbol{b}}^{(j+1,m)}) - \widehat{\boldsymbol{L}}^{(j+1,m)}\widehat{\boldsymbol{b}}^{(j+1,m)}, \boldsymbol{V}\right), \tag{3.12}$$

where $\boldsymbol{V} = \boldsymbol{\Lambda} + \widehat{\boldsymbol{L}}^{(j+1,m)}\boldsymbol{\Sigma}\widehat{\boldsymbol{L}}^{(j+1,m)'}$. Based on the approximate marginal distribution of $\boldsymbol{y}$, the corresponding approximate log-likelihood function can be obtained as follows. In particular,

$$\mathcal{L}_A\left(\boldsymbol{\beta}_{nr}, \boldsymbol{\sigma}_e, \boldsymbol{\sigma}_b\big|\widehat{\boldsymbol{b}}^{(j+1,m)}, \widehat{\boldsymbol{L}}^{(j+1,m)}, \boldsymbol{y}\right) = -\frac{1}{2}\log\left|\boldsymbol{V}\right| - \frac{1}{2}\boldsymbol{K}'\left(\boldsymbol{V}\right)^{-1}\boldsymbol{K}, \tag{3.13}$$

where $\boldsymbol{K} = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{\beta}_{nr},\widehat{\boldsymbol{b}}^{(j+1,m)}) + \widehat{\boldsymbol{L}}^{(j+1,m)}\widehat{\boldsymbol{b}}^{(j+1,m)}$. Using the Newton-Raphson method, $\boldsymbol{\beta}_{nr}^{(j+1,m+1)}$, $\widehat{\boldsymbol{\sigma}}_e^{(j+1,m+1)}$, and $\widehat{\boldsymbol{\sigma}}_b^{(j+1,m+1)}$ can be obtained.

The inside iterations of this step update the estimates of $\boldsymbol{\beta}_{nr}$, $\boldsymbol{b}$, $\boldsymbol{\sigma}_e$, and $\boldsymbol{\sigma}_b$ until they converge. Then, $\widehat{\boldsymbol{\sigma}}_e^{(j+1)}$, and $\widehat{\boldsymbol{\sigma}}_b^{(j+1)}$ are obtained from the results of last iteration (e.g., m^*). That is, $\widehat{\boldsymbol{\sigma}}_e^{(j+1)} = \widehat{\boldsymbol{\sigma}}_e^{(j+1,m^*)}$, and $\widehat{\boldsymbol{\sigma}}_b^{(j+1)} = \widehat{\boldsymbol{\sigma}}_b^{(j+1,m^*)}$.

We want to point out that it is important to choose good starting values for the proposed parameter estimation procedure. The procedure can converges well with good starting values. In Section 3.6.4, we provide a detailed discussion on choosing starting values for the proposed procedure. The performance of the proposed procedure will be studied in Section 3.5.

3.4.3 Confidence Interval Procedure for Parameters

Based on the proposed estimation procedure, one can obtain the ML estimates $\widehat{\boldsymbol{\theta}}$. The statistical inference is not straightforward due to the constrained estimate for $\boldsymbol{\beta}_r$. We use a parametric bootstrap method to quantify the uncertainty in parameter estimation. The bootstrap method is described in **Algorithm 1**.

Algorithm 1:

1. Generate residuals $\boldsymbol{\epsilon}^*$ from $\mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Lambda}})$;

2. Generate random effects $\boldsymbol{b}_i^*$ from $\mathrm{N}(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{b}}), i = 1, \cdots, n$;

3. Generate the degradation responses $\boldsymbol{y}_i^*$ with equation (3.4) based on $\boldsymbol{W}$, $\widehat{\boldsymbol{\beta}}$, $\boldsymbol{\epsilon}^*$ and $\boldsymbol{b}_i^*$;

4. Based on the new dataset $\{\boldsymbol{y}^*, \boldsymbol{W}\}$, all the estimates of unknown parameters $\boldsymbol{\theta}$ are obtained;

5. Repeat steps 1-4 a large number of times (B) and a series of estimated parameters are obtained (i.e., $\widehat{\boldsymbol{\theta}}^{*j}$, $j = 1, \cdots, B$);

6. The confidence interval (CI) for parameters can be computed based on the bias-corrected bootstrap method (e.g., Efron and Tibshirani, 1993). This CI is denoted as CI_{bc}.

3.4.4 Estimation of the Failure-time Distribution

One of the major objectives of degradation analysis is to estimate the failure-time distribution. For a decreasing degradation path, the failure time is defined as the time at which the degradation level first passes a critical value c_0. Let $f\big(D[\boldsymbol{X}_i(t); \boldsymbol{\beta}_D]; \boldsymbol{\beta}_f, \boldsymbol{b}_i\big)$ be a decreasing degradation path. Then, the failure time for unit i can be expressed as:

$$t_i^* = \inf\{t : f\big(D[\boldsymbol{X}_i(t); \boldsymbol{\beta}_D]; \boldsymbol{\beta}_f, \boldsymbol{b}_i\big) \leq c_0\}, \quad i = 1, \cdots, n. \tag{3.14}$$

Note that t_i^* is a function of $\boldsymbol{\beta}$, $\boldsymbol{b}_i$, and $\boldsymbol{X}_i(\infty)$. The failure time of product, denoted as $T[c_0, \boldsymbol{X}(\infty), \boldsymbol{b}]$, is a random variable due to the fact that both $\boldsymbol{X}(\infty)$ and $\boldsymbol{b}$ are random. The cumulative distribution function (cdf) of $T[c_0, \boldsymbol{X}(\infty), \boldsymbol{b}]$ can be expressed as

$$F(t; \boldsymbol{\theta}) = \mathrm{E}_{\boldsymbol{X}(\infty)}\left[\mathrm{E}_{\boldsymbol{b}}\left(\Pr\{T[c_0, \boldsymbol{X}(\infty), \boldsymbol{b}] \leq t\}\right)\right], \quad t > 0. \tag{3.15}$$

Because it is difficult, if not impossible, to obtain an explicit form of $F(t;\boldsymbol{\theta})$, numerical methods or Monte Carlo simulation are often used to compute the cdf.

3.5 Simulation Studies

In this section, we use simulation to study the performance of the proposed method. The simulated dataset is similar to the organic coating data in Gu et al. (2009). To simplify the setting, we consider only one dynamic covariate (i.e., the temperature).

3.5.1 Simulation Setup

The temperature (Temp.) process is simulated based on the following model:

$$\text{Temp.}(t_{ij}) = \mu + \kappa \sin\left[\frac{2\pi}{365}(t_{ij}-\eta)\right] + \left(1+\upsilon\left\{1+\sin\left[\frac{2\pi}{365}(t-\varsigma)\right]\right\}\right)\epsilon(t_{ij}),$$

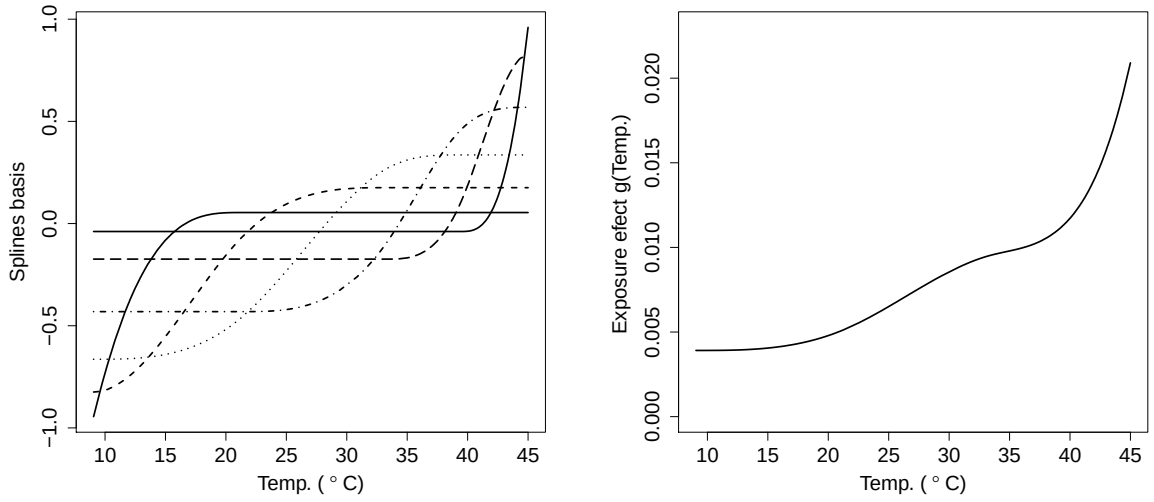
where $\epsilon(t_{ij}) \sim \text{N}(0, 2^2)$ and $(\mu, \kappa, \eta, \upsilon, \varsigma)' = (25, 16, 103, 0.31, 33)'$. The specific degradation path model is defined based on the model introduced by Vaca-Trigo and Meeker (2009). That is,

$$\begin{aligned} y_i(t_{ij}) &= f(D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D];\boldsymbol{\beta}_f,\boldsymbol{b}_i) + \epsilon_i(t_{ij}) \\ &= -\frac{G\times\exp(b_{i1})}{1+\exp\left(-\frac{\log\{D[\boldsymbol{X}_i(t_{ij});\boldsymbol{\beta}_D]\}}{H\times\exp(b_{i2})}\right)} + \epsilon_i(t_{ij}), \end{aligned} \tag{3.16}$$

where $i = 1, \cdots, n$ and $j = 1, \cdots, n_i$. Note, $y_i(t_{ij})$ is the degradation value and $\boldsymbol{X}_i(t_{ij})$ is the history information of dynamic covariate (i.e., Temp.) for individual i from time 0 to time t_{ij}. Regarding the physical interpretation, G is the average total asymptotic degradation amount of product and H is the average steepness level of degradation. The random effects $\boldsymbol{b}_i = (b_{i1}, b_{i2})'$ account for the unit-to-unit variability for G and H caused by unobservable factors. We assume $\boldsymbol{b}_i$ follows a multivariate normal distribution $\text{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{b}})$, where

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{\boldsymbol{b}} = \begin{bmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{bmatrix}.$$

The random effects are independent from the error terms $\epsilon_i(t_{ij})$'s which are modeled to be independently and identically distributed with a normal distribution, i.e., $\epsilon_i(t_{ij}) \overset{\text{iid}}{\sim} \text{N}(0, \sigma^2)$.



(a) Splines basis

(b) Temp. effect function

Figure 3.2: Plots of spline basis functions (a) and effect function of the Temp. covariate (b) that used in the simulation study.

The cumulative exposure $D[\boldsymbol{X}_i(t_{ij}); \boldsymbol{\beta}_D]$ is computed based on the shape-restricted spline basis derived from the dynamic covariates $\boldsymbol{X}_i(t_{ij})$, and it is a nondecreasing function of time. Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_D', \boldsymbol{\beta}_f')'$, where $\boldsymbol{\beta}_D$ are the unknown coefficients for the shape-restricted spline basis and $\boldsymbol{\beta}_f = (G, H)'$ are the rest unknown parameters in model (3.16).

The dynamic covariate Temp. has a monotonic increasing effect on the coating degradation. The number of knots q and the spline order k in the construction of shape-restricted spline basis are set to be $q = 3$ and $k = 3$, respectively. The true value of $\boldsymbol{\beta}_D = (0.009, 0, 0, 0.006, 0, 0.005, 0.006)'$, $G = 2$, $H = 1.1$, $\sigma_0 = 0.06$, $\sigma_1 = 0.02$, $\rho = -0.5$, and $\sigma = 0.02$. There are 20 observations in each unit. Based on those parameter settings, a simulated dataset $\{y_i(t_{ij}), \boldsymbol{X}_i(t_{ij})\},\ i = 1, \cdots, n,\ j = 1, \cdots, 20$, can be obtained. Figure 3.2(a) shows the spline basis functions (not including constant function) and Figure 3.2(b) shows the temperature effect function. To validate the proposed estimation procedure, we choose the sample size $n = 20, 50$, and 100.

Table 3.1: Estimated mean, bias, variance, and MSE of the parameters based on 1000 repeats for parameters in the degradation path model $(G, H, \sigma_0, \sigma_1, \rho, \sigma)'$.

# of units	Parameter	Value	Mean	Bias	Variance	MSE
$n = 20$	G	2.000	2.021	-0.021	6.770×10^{-3}	7.228×10^{-3}
	H	1.100	1.104	-0.004	0.701×10^{-3}	0.714×10^{-3}
	σ_0	0.060	0.058	0.002	0.087×10^{-3}	0.091×10^{-3}
	σ_1	0.020	0.018	0.002	0.039×10^{-3}	0.042×10^{-3}
	ρ	-0.500	-0.540	0.040	0.113965	0.115526
	σ	0.020	0.020	< 0.001	$< 0.001 \times 10^{-3}$	$< 0.001 \times 10^{-3}$
$n = 50$	G	2.000	2.020	-0.020	2.656×10^{-3}	3.037×10^{-3}
	H	1.100	1.105	-0.005	0.299×10^{-3}	0.324×10^{-3}
	σ_0	0.060	0.059	0.001	0.038×10^{-3}	0.038×10^{-3}
	σ_1	0.020	0.019	0.001	0.018×10^{-3}	0.019×10^{-3}
	ρ	-0.500	-0.518	0.018	0.036648	0.036975
	σ	0.020	0.020	< 0.001	$< 0.001 \times 10^{-3}$	$< 0.001 \times 10^{-3}$
$n = 100$	G	2.000	2.016	-0.016	1.245×10^{-3}	1.498×10^{-3}
	H	1.100	1.104	-0.004	0.137×10^{-3}	0.157×10^{-3}
	σ_0	0.060	0.059	0.001	0.019×10^{-3}	0.019×10^{-3}
	σ_1	0.020	0.020	< 0.001	0.008×10^{-3}	0.009×10^{-3}
	ρ	-0.500	-0.525	0.025	0.016186	0.016796
	σ	0.020	0.020	< 0.001	$< 0.001 \times 10^{-3}$	$< 0.001 \times 10^{-3}$

3.5.2 Simulation Results and Discussions

By applying the proposed estimation procedure and model (3.16) to simulated datasets, the results of parameter estimates (not including $\boldsymbol{\beta}_D$) based on 1000 repeats are shown in Table 3.1. It shows that the estimates converge to the true parameters well, although there exist small biases in the estimates of G, H, and ρ. Following Hong and Meeker (2011), we study the mean squared error (MSE) of parameter estimators. The values of MSE in Table 3.1 and Figure 3.3 show decreasing trends as the sample size increases. Using the proposed procedure, the estimates converge with less than five outside iterations. Overall, the simulation studies show that the performance of the proposed estimation procedure is satisfactory, even when the sample size is small (e.g., $n = 20$).

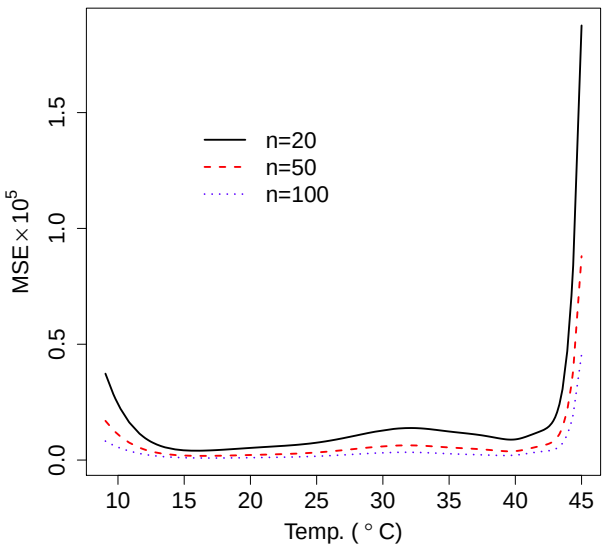
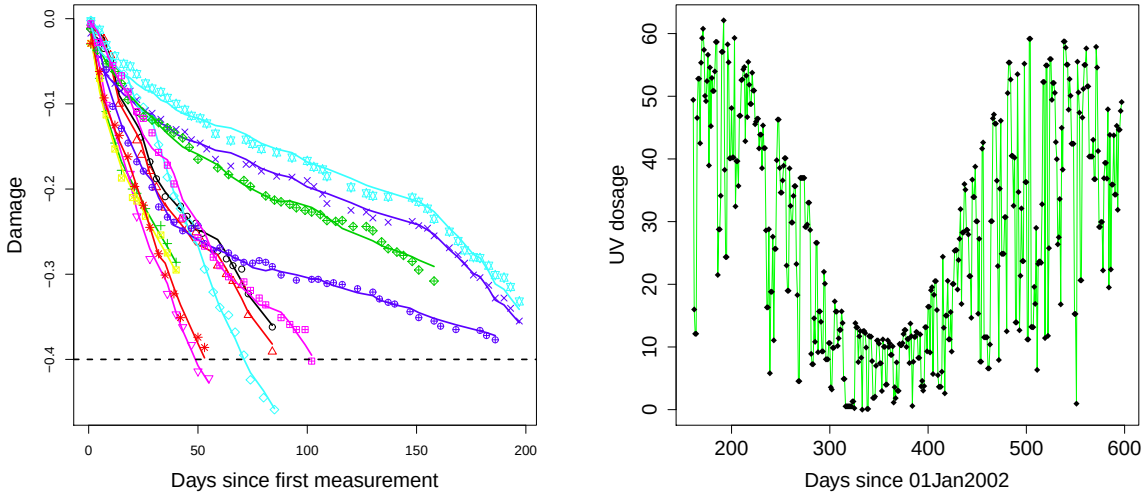


Figure 3.3: Plot of the MSE of the estimator of the effect function.

3.6 Application

3.6.1 Outdoor Weathering Data

To illustrate the proposed method, the outdoor weathering data in Gu et al. (2009) is used. The dataset contains degradation measurements and environmental covariates for units tested during the period from 2002 to 2006. The outdoor UV exposure, Temp., and RH. were recorded automatically with the sensors for each unit. Fourier transform infrared spectroscopy (FTIR) is used to measure the degradation level based on the changes of signals around the wavenumber 1250 cm^{-1}. There are $n = 36$ specimens tested with staggered entries. A failure is defined when the degradation measurement of a unit is lower than -0.4. More details are available in Gu et al. (2009). A subset of degradation data and the UV covariate are shown in Figure 3.4.



(a) Degradation paths (b) UV process

Figure 3.4: Plots showing a subset of the degradation paths (a) and a subset of the daily values of the UV dosage.

3.6.2 Model Fitting and Comparisons

Model (3.16) is used to fit the outdoor weathering data with three dynamic covariates (UV, Temp., and RH). Based on the previous study in Gu et al. (2009), the shapes of the effect functions for UV dosage and temperature are nondecreasing, while the shape of the effect function of RH is convex. Thus, we use shape-restricted splines to estimate those covariate effect functions.

To construct the cumulative shape-restricted spline basis matrix, the number of knots for each dynamic covariate and the spline order in the M-spline functions (e.g., Meyer, 2008) needed to be determined. We use the same number of knots (i.e., q) for all dynamic covariates. Those knots are evenly located among the sorted covariate values. We use the Bayesian information criterion (BIC) for knots selection, which puts a greater penalty on the number of parameters. We use the approximate likelihood function in (3.13) to calculate the BIC. Figure 3.5 shows that the optimized choice: the number of knots $q = 3$ and the spline order $k = 3$, which has the minimum BIC.

Using the proposed method, we obtain the estimated model parameters. The parameter

estimates (not including $\boldsymbol{\beta}_D$) with 95% confidence intervals with bias correction (CI_{bc}) are shown in Table 3.2. Figure 3.4(a) shows the fitted degradation paths for twelve randomly selected units, which shows that the proposed model fits the degradation data well. Based on Table 3.2, we know that the population average asymptotic degradation amount is around 1.006 and the population average value of the steepness of the degradation is around 1.072. Those values can help the researcher understand what the maximum amount of degradation is and how fast the degradation is in population level.

Figure 3.6 shows the estimated covariate effect functions and the corresponding 95% pointwise CI_{bc} for the three dynamic covariates. The figure shows that the UV dosage has a strong effect on the degradation of the coating, which is consistent with the physical knowledge because the degradation is mainly caused by photodegradation. The temperature effect is relatively small compared to the UV effect. The RH effect function has the expected convex shape, and it reaches the lowest point during $15 \sim 20\%$ RH.

To make a comparison, we fit the linear random effects model in Hong et al. (2014) and calculate the BIC based on their model. The BIC for the nonlinear random effects model in this chapter is -5463.03, which is much smaller than the value from the linear random effects model -4640.19. The smaller BIC value in the nonlinear random effects model indicates a better overall fitting. Figure 3.7 shows the residuals for all units in the outdoor weathering data, which indicates that both models meet the constant variance assumption, but the nonlinear random effects model have a smaller variance for the error term. To see the improvement at the individual level, we pick several units to compare the degradation path fitting. Figure 3.8 indicates that the nonlinear random effects model fitting is better than the linear random effects model fitting at the individual level.

3.6.3 Failure-time Distribution Estimation

Here, we apply a Monte Carlo simulation approach to estimate the failure-time distribution. A model for the covariate processes is needed, and we use the parametric model for covariates in Hong et al. (2014). In general, one first needs to simulate the covariate processes and the

Table 3.2: Parameter estimates and approximate 95% CI_{bc}'s for $(G, H, \sigma_0, \sigma_1, \rho, \sigma)'$ in the degradation path model.

Parameter	Estimate	Standard Error	95% CI_{bc}	
			Lower	Upper
G	1.006	0.062	0.889	1.101
H	1.072	0.032	1.007	1.128
σ_0	0.105	0.015	0.078	0.137
σ_1	0.146	0.018	0.116	0.188
ρ	-0.656	0.111	-0.807	-0.305
σ	0.010	0.00026	0.010	0.011

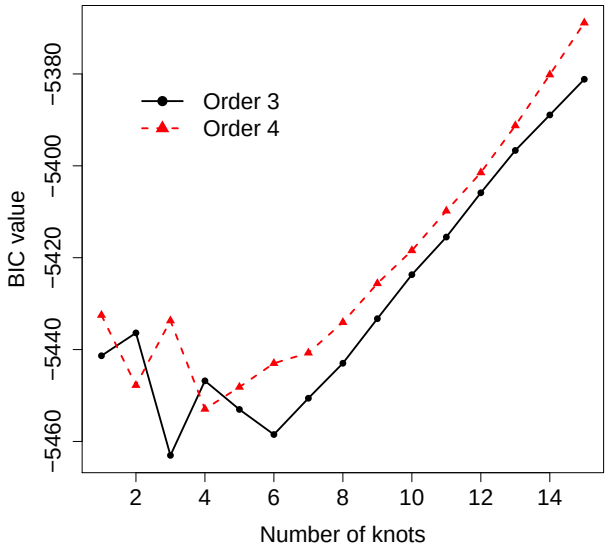
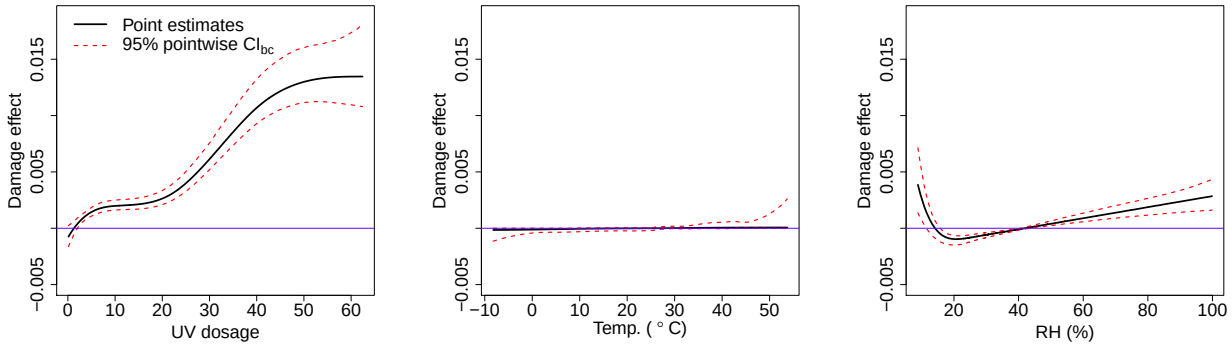
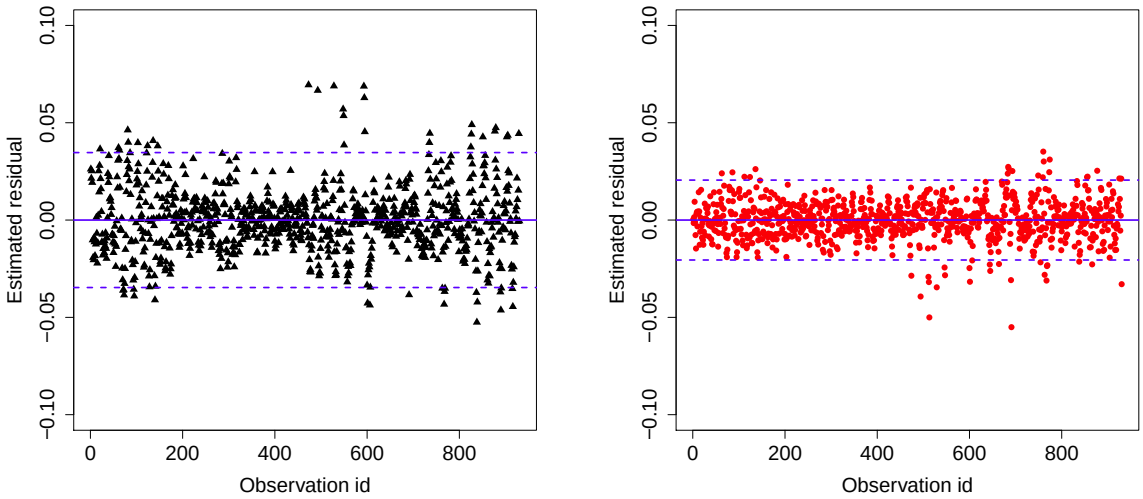


Figure 3.5: Plot of BIC values for selection of knots and spline order.



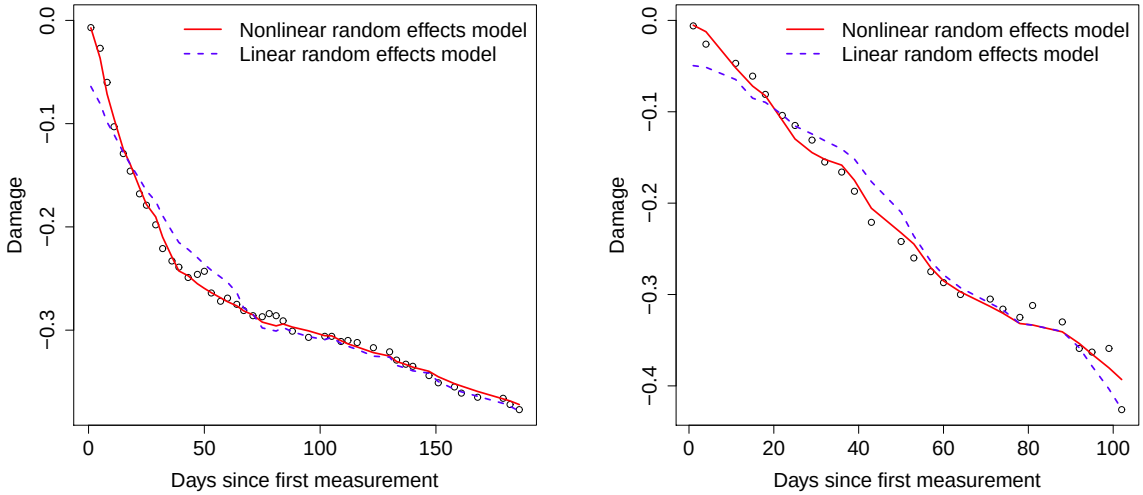
(a) UV (b) Temp. (c) RH

Figure 3.6: Estimated effect functions and the corresponding approximate 95% pointwise CI_{bc} for the three covariates.



(a) Linear random effects model (b) Nonlinear random effects model

Figure 3.7: Comparisons of residuals based on the linear random effects model and the nonlinear random effects model (the dashed lines show the ±1.96 standard deviations).



(a) Sample G3-10 (b) Sample G9-9

Figure 3.8: Comparisons of fitted paths based on the linear random effects model and the nonlinear random effects model for units G3-10 and G9-9.

random effects. Based on the estimated parameters, one can simulate the degradation paths. Given the simulated degradation paths, the failure times $\widehat{t}_i^*, i = 1, \cdots, B$ are obtained. The empirical failure-time distribution is computed by $\widehat{F}(t) = \sum_{i=1}^{B} \mathbf{1}(\widehat{t}_i^* \leq t)/B$, where $\mathbf{1}(\cdot)$ is an indicator function. The CI for the cdf of failure-time can be obtained by using **Algorithm 1**. Figure 3.9 shows the estimated cdf for a population with units starting randomly from day 161 to day 190.

3.6.4 Discussion on Starting Values

The selection of starting values is important in optimizations. Knowing the background of the data will be helpful. In the outdoor weathering application, because G is the total asymptotic degradation amount of the coating, it can be estimated by a value that is larger than the maximum of the degradation measurements in the data. The steepness level of the degradation path H can be initialized as one. The random effects are chosen as the means of their distributions (i.e., $\mathbf{0}$). Given the initial setting of G and H, the initial estimates of cumulative exposure can be obtained from equation (3.16) by treating the error terms as

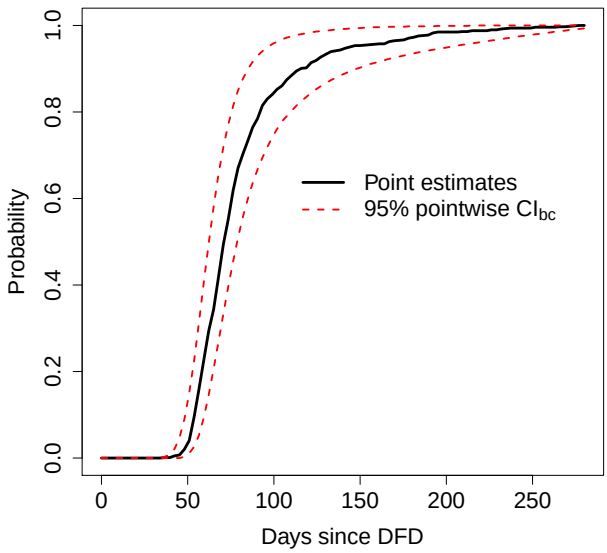


Figure 3.9: The estimated cdf and 95% pointwise CIs for a population with units starting randomly from day 161 to day 190.

0. The initial coefficients for the spline basis can be estimated by the mixed primal-dual basis algorithm given the initial estimates of cumulative exposure. Once we have the initial estimates of the cumulative exposure function, we can treat it as a fixed term in the model and use R package *nls* to estimate fixed-effect parameters (i.e., G and H) individually. Once we obtained the estimates of G and H for each unit, the initial estimates of $\boldsymbol{\Sigma_b}$ can be obtained as the sample variance-covariance matrix of those estimates of G and H. Similar strategy can be used for $\boldsymbol{\Lambda}$.

An alternative way of obtaining starting values is to fit a nonlinear (without random effects) model with constraints on some parameters ($\boldsymbol{\beta}_r$). Here we briefly explain the idea. Without random effects, the model in (3.4) reduces to $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{W};\boldsymbol{\beta}) + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Lambda})$. We need to minimize the nonlinear least squares function

$$\left\|\boldsymbol{\Lambda}^{-1/2}\boldsymbol{y} - \boldsymbol{\Lambda}^{-1/2}\boldsymbol{f}(\boldsymbol{W};\boldsymbol{\beta})\right\|^2. \tag{3.17}$$

First we estimate $\boldsymbol{\beta}_r$ and $\boldsymbol{\beta}_{nr}$, given $\boldsymbol{\Lambda}$. We linearize the nonlinear function $f(\cdot)$ using the Taylor series expansion at the current estimates. Based on the mixed primal-dual bases

algorithm, the estimate of $\boldsymbol{\beta}_r$ can be updated. Then, estimate of $\boldsymbol{\beta}_{nr}$ can be updated using the Gauss-Newton method. Compared to the step in Section 3.4.2.1, this step will be much simpler because random effects are not considered. Second we update $\boldsymbol{\Lambda}$, given $\boldsymbol{\beta}_r$. This step can be obtained by the maximization of the likelihood function using the Newton-Raphson method or other optimization methods. We do the iterations of those two steps until all parameter estimates converge.

3.7 Conclusion and Areas for Future Research

In this chapter, we propose a class of nonlinear general degradation path models with random effects to describe degradation data with dynamic covariates. The effects of dynamic covariates are incorporated into the degradation model via shape-restricted splines. We develop an efficient procedure to estimate the unknown parameters. The outdoor weathering data are used to illustrate the method. The proposed method can also be applied to nonlinear (without random effects) models with shape-restricted covariate effects, as described in Section 3.6.4, in which case the estimation will be simpler because the absence of random effects.

Many degradation mechanisms show nonlinear degradation trends in practice. Comparing to linear models, nonlinear models are more flexible and have broader applications with complicated data. Besides the outdoor weathering data, the proposed method can also be applied on many products used in highly variable conditions. For example, the degradation of batteries is affected by temperature and use frequency, and the degradation of the solar panel is affected by outdoor environments (i.e., UV dosage, Temp., and RH). For another example, the method can be applied to the degradation of the heater in ingot manufacturing in Section 3.1.1, where the degradation is affected by the power and other process variables (e.g., temperature and chemical component of raw materials).

For degradation modeling, we want to point out that it is important to consider physical/chemical knowledge. For our motivating application (i.e., the outdoor weathering data), the degradation is a result of loss of chemical structures. Thus, the cumulative damage model

can provide a good description for the failure mechanism. For other failure mechanisms, the cumulative damage model may not be appropriate. Also, nonparametric methods are used to model the effects of dynamic covariates. In the motivating application, we have data points for a wide range of the covariates. Thus, extrapolation is not used in predictions. If extrapolation is needed in some applications, the nonparametric method should be used with caution, although the shape-restricted covariate effect function allows one to do some extrapolation.

The proposed method can be applied to the health care problems and pharmacokinetics research. The proposed method can also be useful in maintenance scheduling. Based on the distribution of lifetime, one can also develop a better strategy to make maintenance and logistical decisions. Our current method only considers soft failures, it will also be interesting to consider both soft and hard failures in the future. Heterogeneities in environments (e.g., Ye et al., 2013) can also be considered in degradation modeling.

Bayesian methods can be applied for parameter estimations for nonlinear models with random effects (e.g., Davidian and Giltinan, 2003). The Monte Carlo Markov chain (MCMC) techniques will be needed to obtain the posteriors for inferences in most cases. The challenge in applying the Bayesian method to the model used in this chapter could be the constraints on some of the parameters, which will be an interesting topic for future research. Functional varying coefficient regression techniques are also available in literature by using regression splines (e.g., Huang and Shen, 2004) and penalized splines (e.g., Cao et al., 2010). Applying functional regression approach to degradation data with time-varying covariates can also be an interesting topic for future research.

Bibliography

S. J. Bae and P. H. Kvam. A nonlinear random-coefficients model for degradation testing. *Technometrics*, 46:460–469, 2004.

V. Bagdonavičius and M. S. Nikulin. Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, 7:85–103, 2001.

V. Bagdonavičius, A. Bikelis, and V. Kazakevičius. Statistical analysis of linear degradation and failure time data with multiple failure modes. *Lifetime Data Analysis*, 10:65–81, 2004.

V. Bagdonavičius, F. Haghighi, and M. Nikulin. Statistical analysis of general degradation path model and failure time data with multiple failure modes. *Communications in Statistics - Theory and Methods*, 34:1771–1791, 2005.

Y. Cao, H. Lin, T. Wu, and Y. Yu. Penalized spline estimation for functional coefficient regression models. *Computational Statistics and Data Analysis*, 54:891–905, 2010.

M. Davidian and D. M. Giltinan. Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, 8: 387C419, 2003.

B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, FL: Boca Raton, 1993.

L. A. Escobar and W. Q. Meeker. A review of accelerated test models. *Statistical Science*, 21: 552–577, 2006.

D. A. S. Fraser and H. Massam. A mixed primal-dual bases algorithm for regression under inequality constraints. Application to concave regression. *Scandinavian Journal of Statistics*, 16:65–74, 1989.

X. Gu, D. Stanley, W. E. Byrd, B. Dickens, I. Vaca-Trigo, W. Q. Meeker, T. Nguyen, J. W. Chin, and J. W. Martin. Linking accelerate laboratory test with outdoor performance results for a model epoxy coating system. In J. Martin, R. A. Ryntz, J. Chin, and R. A. Dickie, editors, *Service Life Prediction of Polymeric Materials*. Springer, NY: New York, 2009.

Y. Hong and W. Q. Meeker. The importance of identifying different components of a mixture distribution in the prediction of field returns. *Applied Stochastic Models in Business and Industry*, 2011.

Y. Hong, Y. Duan, W. Q. Meeker, D. L. Stanley, and X. Gu. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics*, 2014. in press, DOI:10.1080/00401706.2014.915891.

J. Z. Huang and H. Shen. Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scandinavian Journal of Statistics*, 31:515–534, 2004.

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

M. J. Lindstrom and D. M. Bates. Nonlinear mixed effects models for repeated-measures data. *Biometrics*, 46:673–687, 1990.

C. J. Lu and W. Q. Meeker. Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 34:161–174, 1993.

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

W. Q. Meeker and M. J. LuValle. An accelerated life test model based on reliability kinetics. *Technometrics*, 37:133–146, 1995.

M. C. Meyer. Inference using shape-restricted regression splines. *The Annals of Applied Statistics*, 2:1013–1033, 2008.

W. J. Padgett and M. A. Tomlinson. A cumulative damage model for strength of materials when initial damage is a gamma process. *Journal of Statistical Theory and Applications*, 1: 1–14, 2002.

W. J. Padgett and M. A. Tomlinson. Inference from accelerated degradation and failure data based on Gaussian process models. *Lifetime Data Analysis*, 10:191–206, 2004.

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

C. Park and W. J. Padgett. Stochastic degradation models with several accelerating variables. *IEEE Transactions on Reliability*, 55:379–390, 2006.

J. C. Pinheiro and D. M. Bates. Approximations to the log-likelihood function in the non-linear and mixed-effects moel. *Journal of Computational and Graphical Statistics*, 4:12–35, 1995.

M. Robinson and M. Crowder. Bayesian methods for a growth-curve degradation model with repeated measures. *Lifetime Data Analysis*, 6:357–374, 2000.

Y. Shi and W. Q. Meeker. Bayesian methods for accelerated destructive degradation test planning. *IEEE Transactions On Reliability*, 61:245–253, 2012.

I. Vaca-Trigo and W. Q. Meeker. A statistical model for linking field and laboratory exposure results for a model coating. In J. Martin, R. A. Ryntz, J. Chin, and R. A. Dickie, editors, *Service Life Prediction of Polymeric Materials*. Springer, NY: New York, 2009.

J. Wakefield. The Bayesian analysis of population pharmacokinetic models. *Journal of the American Statistical Association*, 91:62–75, 1996.

G. A. Whitmore. Estimation degradation by a Wiener diffusion process subject to measurement error. *Lifetime Data Analysis*, 1:307–319, 1995.

R. Wolfinger. Laplace's approximation for nonlinear mixed models. *Biometrika*, 80:791–795, 1993.

Z.-S. Ye and N. Chen. The inverse Gaussian process as a degradation model. *Technometrics*, 56:302–311, 2013.

Z.-S. Ye, Y. Hong, and Y. Xie. How do heterogeneities in operational environments affect field failures? *The Annals of Applied Statistics*, 7:2249–2271, 2013.

X. X. Yuan and M. D. Pandey. A nonlinear mixed-effects model for degradation data obtained from in-service inspections. *Reliability Engineering and System Safety*, 94:509–519, 2009.

J. Zhang, W. Li, K. Wang, and R. Jin. Process adjustment with an asymmetric quality loss function. *Journal of Manufacturing Systems*, 33:159–165, 2014.

Chapter 4 A Multi-level Trend-renewal Process for Modeling Systems with Recurrence Data

Abstract

A repairable system is a system that can be restored to an operational state after a repair event. The system may experience multiple events over time, which are called recurrent events. To model the recurrent event data, the renewal process (RP), the nonhomogeneous Poisson process (NHPP), and the trend-renewal process (TRP) are often used. Compared to the RP and NHPP, the TRP is more flexible for modeling, because it includes both RP and NHPP as special cases. However, for a multi-level system (e.g., system, subsystem, and component levels), the current TRP model may not be adequate if the repair is in the form of replacement and subsystem-level replacement events affect the rate of occurrence of the component-level replacement events. In this chapter, we propose a general class of models to describe replacement events in a multi-level repairable system by extending the TRP model. We also develop procedures for estimation of model parameters and prediction of future events based on historical data. The proposed model and method are validated by simulation studies and are illustrated by an simulated application.

Key Words: Multi-level repairable systems, Nonhomogeneous Poisson processes, Renewal process, Random effect, Time-dependent covariate, Trend-renewal process.

4.1 Introduction

4.1.1 Background

A repairable system is defined as a system that can be restored to an operational state after a repair. In practise, a repairable system may experience multiple replacement events at different levels over time. For example, we consider a repairable vehicle with three levels: system (e.g., a truck), subsystem (e.g., the truck engine), and component (e.g., the oil pump). The replacement events can be the replacement of the oil pump (may be new or refurbished) or the replacement of the entire engine (may be new or refurbished). For some other examples, the failure of a computer motherboard can be repaired by replacing the whole motherboard or just by replacing the failed capacitor on the motherboard. The failure of a gearbox can be repaired by replacing the whole gearbox or just by replacing the failed gear.

In this chapter, we consider a two-level repairable system where events can occur at the subsystem level, or the component (within a subsystem) level. We focus on a specific subsystem (e.g., the engine) in a vehicle and a particular component within that subsystem (e.g., the oil pump), although a subsystem may have many components. In particular,

- The replacement of a subsystem is called a subsystem event. In this case, the system can only be fixed by a subsystem replacement.

- The replacement of a component is called a component event. In this case, the system is fixed by the component replacement.

Often, the failed subsystems or components are replaced with refurbished units that are not as good as new units. When the subsystem is replaced, of course, the components inside such subsystem will be replaced at the same time leading to a change of the risk of having a failure at the component level. This repair information at multiple levels is available through maintenance records. In addition to replacement event times, dynamic covariates, such as system usage information, loading, and shocks may also be available.

One important goal of the modeling of the replacement events is to do field failure prediction, which is useful for purposes such as prognostics, maintenance scheduling, and spare parts provisioning. Prediction of component replacement events is difficult when there are also subsystem replacements. The objective of this chapter is to use replacement information at multiple levels and system usage information to make field failure prediction for a critical component in a subsystem. We need a model that can incorporate the effects of system usage information and other possibly unobservable factors, and the effects that replacements at different levels have on components. The model can also handle that replacements may not be perfect and there are possible system-to-system differences.

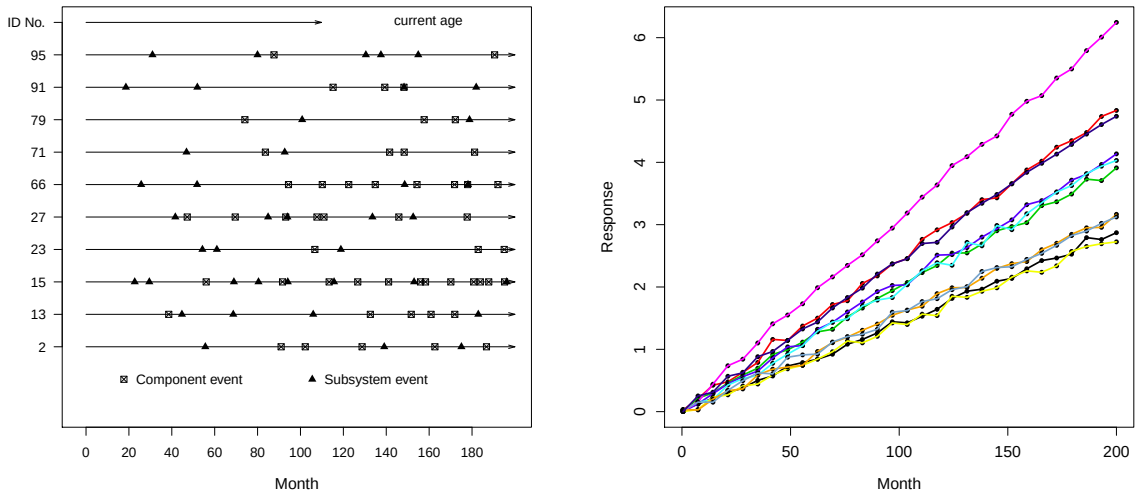
4.1.2 Motivating Application

This chapter is motivated by the need to model recurrent events from a fleet of industrial systems, which we call Vehicle B. Vehicle B is a two-level repairable system, and it may experience subsystem (engine) and component (oil pump) events over time. To protect sensitive and proprietary information, we use simulated data with rescaled time units to illustrate our method. In the simulated Vehicle B data, there are $n = 100$ systems, and the data freeze date (DFD) is 200 months. The total number of component events and subsystem events are 1176 and 392, respectively. The event histories of ten randomly selected units from the Vehicle B fleet are shown in Figure 4.1(a).

Besides the event histories, a time-dependent covariate called cumulative usage is also simulated. Figure 4.1(b) shows the cumulative usage for ten randomly selected units from the Vehicle B fleet. Compared to the length of running times of the systems, the time needed to effect a repair is ignorable and is assumed to be zero. The prediction of the total number of component events in a future time point is needed in this application.

4.1.3 Related Literature and This Work

The nonhomogeneous Poisson process (NHPP) and the renewal process (RP) are the two most commonly used models in the analysis of recurrent event data (e.g., Zhao and Liu,



(a) Event processes (b) Cumulative usage processes

Figure 4.1: Plots of event processes and cumulative usage processes for ten randomly selected units in the Vehicle B fleet.

2003, Leemis, 2004, and Hong et al., 2013) with the assumption that the effect of repair is perfect or minimal, respectively. For general repairs, Brown and Proschan (1983) proposed an imperfect repair model. Kijima (1989) introduced two types of virtual age models by reducing the age of the system after each repair. Lawless and Thiagarajah (1996) used a proportional intensity model to incorporate renewals and time trends. Wang and Pham (1996) proposed a quasi-renewal process with the consideration of maintenance cost. Doyen and Gaudoin (2004) proposed two new classes of imperfect repair models.

Lindqvist et al. (2003) and Lindqvist (2006) introduced a trend-renewal process (TRP) which includes the NHPP and RP as special cases, which has been widely used in literature (e.g., Yang et al., 2012, and Pietzner and Wienke, 2013). Heggland and Lindqvist (2007) derived a non-parametric maximum likelihood estimator of the intensity function for the TRP. Franz et al. (2013b) proposed methods for point prediction and interval prediction for the first time to failure using simulation. For virtual age models, Yañez et al. (2002) and Yu et al. (2013) proposed the methods to estimate the expected number of failures using Monte Carlo simulation and an analytic approach, respectively.

For multi-level repairable system analysis, some papers have focused on the reliability analysis of a system by combining information from different levels using Bayesian methods. Examples include Johnson et al. (2005), Wilson et al. (2006), and Liu et al. (2011). We know of no previous work that has been done for reliability estimation and prediction for multi-level repairable systems with the consideration of the effect of subsystem events on component events.

Motivated by the Vehicle B data, we propose a multi-level trend renewal process (MTRP) with time-dependent covariates in the modeling of component events. Based on the MTRP model, we also develop procedures for obtaining point predictions and prediction intervals for the number of component events in a future time. To incorporate system-to-system variability, random effects are introduced in the MTRP model and the parameters are estimated by the Metropolis-within-Gibbs algorithm. The methods of estimation and predictions are validated by simulation studies and illustrated by the Vehicle B application.

4.1.4 Overview

The rest of the chapter is organized as follows. In Section 4.2, we introduce some existing models and the proposed MTRP model. In Section 4.3, we develop estimation methods for unknown parameters in the proposed MTRP model. In Section 4.4, we develop procedures for point predictions and prediction intervals (PI) based on Monte Carlo simulation. In Section 4.5, we validate the proposed methods by simulations. In Section 4.6, the methods of modeling and predictions are illustrated based on a simulated application. Section 4.7 contains a summary and some future research topics.

4.2 Repairable System Models

4.2.1 Existing Models

Let $0 < T_1 \cdots < T_i < \cdots$ be the event times from a repairable system. Denote $N(t)$ as the counting process, which counts the number of events that occur in time interval $(0, t]$, and let

$\mathcal{F}_t$ be the event history up to time t. The event intensity for the counting process is

$$\lambda(t|\mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\{N(t+\Delta t) - N(t) = 1|\mathcal{F}_{t^-}\}}{\Delta t},$$

where $\mathcal{F}_{t^-}$ is the event history immediately prior to time t. The cumulative event intensity function is defined as $\Lambda(t) = \int_0^t \lambda(u|\mathcal{F}_{u^-})\, du$.

The RP, denoted by RP(F), corresponds to a perfect repair (i.e., replacement with a new unit), and the gaps between event times are independently and identically distributed (iid) with F. Here F is a cumulative distribution function (cdf). That is, $T_{i+1} - T_i \overset{\text{iid}}{\sim} F,\ i = 1, 2, \cdots$. Let $h(z)$ be the hazard function of F. The event intensity function of the RP is $\lambda(t|\mathcal{F}_{t^-}) = h[t - T_{N(t^-)}]$, where $T_{N(t^-)}$ is the last event time before time t. The NHPP corresponds to a minimal repair (e.g., adjustment or replacement of a small part of a large unit). The intensity function is $\lambda(t|\mathcal{F}_{t^-}) = \lambda(t)$, which does not depend on the event history. For the NHPP, the transformed event times $\Lambda(T_i)$ follows a homogeneous Poisson process (HPP) with mean of one. The gaps between the transformed times are iid with an exponential distribution with mean one. That is, $\Lambda(T_{i+1}) - \Lambda(T_i) \overset{\text{iid}}{\sim} \text{Exp}(1),\ i = 1, 2, \cdots$.

The TRP model describes situations that are in-between NHPP and RP, and contains NHPP and RP as special cases. The gaps between the transformed event times are iid with RP(F). That is, $\Lambda(T_{i+1}) - \Lambda(T_i) \overset{\text{iid}}{\sim} F,\ i = 1, 2, \cdots$. We denote the TRP by TRP(F, λ), where $\lambda(t) = d\Lambda(t)/dt$ is called the trend function and F is called the renewal distribution function. Figure 4.2 illustrates the definition of the TRP. The event intensity function is $\lambda(t|\mathcal{F}_{t^-}) = h\{\Lambda(t) - \Lambda[T_{N(t^-)}]\}\lambda(t)$. The trend function $\lambda(t)$ reflects system deterioration (or improvement) overtime, independent of replacement events or other repair-related events. The factor $h\{\Lambda(t) - \Lambda[T_{N(t^-)}]\}$ reflects the effect of the most recent repair at time $T_{N(t^-)}$. After each repair, there is a change in the event intensity function. The behavior of the change is determined by the hazard function of the renewal distribution function F.

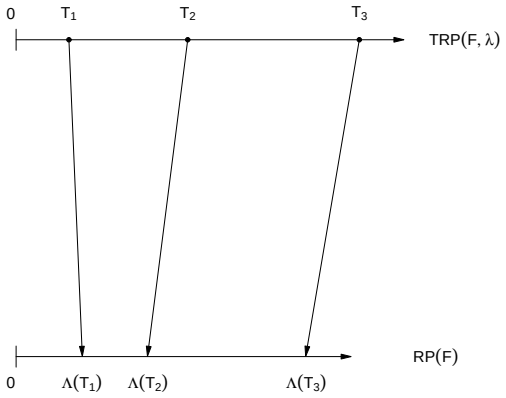


Figure 4.2: Illustration of the TRP model.

4.2.2 Notation for Data

We consider a single multi-level repairable system, which is under observation in time interval $(0, \tau_i]$, where $i = 1, \cdots, n$. The subsystem consists of many components and we focus on the replacement of a critical component. Let $N_i(t) = N_{is}(t) + N_{ic}(t)$ be the total number of replacement events up to time t, where $N_{is}(t)$ and $N_{ic}(t)$ are the number of subsystem events and component events up to time t for system i, respectively. Let $0 < t^s_{i1} < \cdots < t^s_{i,N_{is}(\tau_i)} < \tau_i$ be the times for subsystem events, and let $0 < t^c_{i1} < \cdots < t^c_{i,N_{ic}(\tau_i)} < \tau_i$ be the times for component events. The replacement event times, regardless of the types, are denoted by $0 < t_{i1} < \cdots < t_{i,N_i(\tau_i)} < \tau_i$.

In the Vehicle B data, the time-dependent covariate (i.e., cumulative usage) at time t is denoted by $X_i(t)$ for system i, where $i = 1, \cdots, n$. The time-dependent covariate process for system i is denoted by $\boldsymbol{X}_i(t)$, where $\boldsymbol{X}_i(t) = \{X_i(u) : 0 < u \leq t\}$. Covariate process $\boldsymbol{X}_i(t)$ is recorded at time t_{ik}, where $k = 1, \cdots, n_i$, and n_i is the number of time points where the covariate information is available for system i before the end of observation τ_i.

With the consideration of time-dependent covariate, the replacement events history up to time t is $\mathcal{F}_t = \{N_{ic}(u), N_{is}(u), X_i(u) : 0 < u \leq t\}$, and the history of subsystem events up to

time t is $\mathcal{F}_t^s = \{N_{is}(u), X_i(u) : 0 < u \leq t\}$.

4.2.3 The Proposed Multi-level Trend-renewal Process

For a two-level repairable system unit i, the intensity functions for the subsystem and component level events are modeled as follows:

$$\text{Subsystem level: } \lambda_i^{s*}(t|\mathcal{F}_{t^-}^s;\boldsymbol{\theta}^s) = h^{s*}\{\Lambda_i^*(t) - \Lambda_i^*[t_{i,N_{is}(t^-)}^s];\boldsymbol{\theta}^s\}\lambda_i^*(t;\boldsymbol{\theta}^s), \tag{4.1}$$

$$\text{Component level: } \lambda_i^c(t|\mathcal{F}_{t^-};\boldsymbol{\theta}^c) = h^c\left\{\Lambda_i^s(t|\mathcal{F}_{t^-}^s) - \Lambda_i^s\left[t_{i,N_i(t^-)} \mid \mathcal{F}^s_{t^-_{i,N_i(t^-)}}\right];\boldsymbol{\theta}^c\right\}\lambda_i^s(t|\mathcal{F}_{t^-}^s;\boldsymbol{\theta}^c). \tag{4.2}$$

In (4.1), we use a TRP model, TRP(F^{s*}, λ_i^*), without random effects, to describe the subsystem level events. The unknown model parameters in (4.1) are denoted by $\boldsymbol{\theta}^s$. Here, F^{s*} is the renewal distribution for subsystem event process, $h^{s*}(\cdot)$ is the corresponding hazard function, and $\Lambda_i^*(t) = \int_0^t \lambda_i^*(u;\boldsymbol{\theta}^s)du$ is the cumulative intensity function. The function $\lambda_i^*(t;\boldsymbol{\theta}^s) = \lambda_b^*(t)\exp\{\gamma^* g[X_i(t)]\}$ is the intensity trend function for system unit i with $\lambda_b^*(t)$ as the baseline intensity function and γ^* as the coefficient of a transformed function of time-dependent covariate (i.e., $g[X_i(t)]$). In the rest of this chapter, we use $g[X_i(t)] = \log[X_i(t)]$ as the function of time-dependent covariate.

The proposed MTRP model for component events in (4.2) is an extension of the TRP in the sense that we use a new trend function $\lambda_i^s(t|\mathcal{F}_{t^-}^s;\boldsymbol{\theta}^c)$ for the component event process that can incorporate the effect of subsystem events on the component events, because the intensity of component events may be affected by the subsystem events. In particular,

$$\lambda_i^s(t|\mathcal{F}_{t^-}^s;\boldsymbol{\theta}^c) = h^s\{\Lambda_i(t) - \Lambda_i[t_{i,N_{is}(t^-)}^s]\}\lambda_i(t;\boldsymbol{\theta}^c), \tag{4.3}$$

where $\boldsymbol{\theta}^c = (\theta_1^c, \cdots, \theta_z^c)'$ denotes a vector of unknown parameters with length of z in (4.2). The function $h^s(\cdot)$ in (4.3) describes the effects of subsystem events on the intensity of component events. Like in (4.1), the form of $h^s(\cdot)$ can be taken as a hazard function , and its corresponding cdf form is $F^s(\cdot)$. The function $\lambda_i(t;\boldsymbol{\theta}^c)$ describes the effect of covariate and other unknown factors on the component intensity function. Here, $\Lambda_i(t) = \int_0^t \lambda_i(u;\boldsymbol{\theta}^c)\ du$.

The renewal distribution function of component events model in (4.2) is denoted by $F^c(\cdot)$.

Let $f^c(t)$, $S^c(t) = 1 - F^c(t)$, and $h^c(t)$ be the probability density function (pdf), survival function, and hazard function corresponding to F^c, respectively. Also, let $\Lambda_i^s(t|\mathcal{F}_{t^-}^s) = \int_0^t \lambda_i^s(u|\mathcal{F}_{u^-}^s; \boldsymbol{\theta}^c)\, du$, and $\Lambda_i(t) = \int_0^t \lambda_i(u; \boldsymbol{\theta}^c)\, du$.

Note that (4.3) has the same form of the TRP intensity model for subsystem events but with a different set of parameter $\boldsymbol{\theta}^c$. The function in (4.3) can reflect effect that subsystem events have on the component events intensity. For example, when $h^s(\cdot)$ is a constant function, subsystem events have no effect on components [e.g., Figure 4.3(a)]; when $\lambda_i(t; \boldsymbol{\theta}^c)$ in (4.3) is a constant function over time and $h^s(\cdot)$ is not a constant function, a subsystem event corresponds to a perfect repair effect and result in an immediate reduction of the intensity function [e.g., Figure 4.3(b)]; when neither $\lambda_i(t; \boldsymbol{\theta}^c)$ nor $h^s(\cdot)$ is a constant function, the subsystem events are imperfect repairs [e.g., Figure 4.3(c)]. Thus, (4.3) is a flexible trend function for describing the effect of subsystem events on component events intensity.

Similar to the subsystem event process, the incorporation of time-dependent covariate can be achieved by

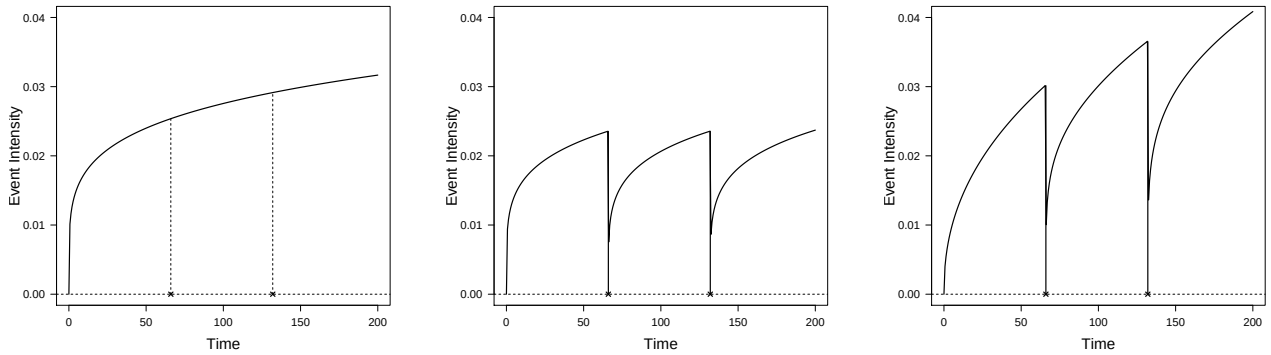
$$\lambda_i(t; \boldsymbol{\theta}^c) = \lambda_b(t) \exp\{\gamma \log[X_i(t)]\} \quad i = 1, \cdots, n. \tag{4.4}$$

Here, $\lambda_b(t)$ denotes the baseline intensity trend function if no component/subsystem replacement event occurs. The MTRP with a time-dependent covariate can be denoted by MTRP(F^c, F^s, λ_i).

To incorporate unit-to-unit variability in component event process, we use random effects in the intensity function (4.4) as follows,

$$\lambda_i(t; \boldsymbol{\theta}^c) = \lambda_b(t) \exp\{\gamma \log[X_i(t)] + w_i\} \quad i = 1, \cdots, n, \tag{4.5}$$

where the random effect for system i, w_i, is iid with $\mathrm{N}(0, \sigma_r^2)$. Here, σ_r^2 is the variance of the normal distribution, and is not contained in $\boldsymbol{\theta}^c$. Define $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)'$. The heterogeneous MTRP for component events process is denoted by HMTRP(F^c, F^s, λ_i).



(a) Minimal repair effects (b) Perfect repair effects (c) Imperfect repair effects

Figure 4.3: Different cases of the trend function in (4.3). The symbol of "$*$" indicates the occurrence of subsystem event.

4.2.4 Properties and Special Cases of MTRP

From another perspective, the MTRP model in (4.2) includes two TRP models in a hierarchical structure. The higher-level TRP model is used to describe the effect of subsystem events on the intensity of component events. The lower-level TRP model is used to model the component events with the higher-level TRP intensity function as its trend function. Thus, the MTRP can be defined as,

$$\Lambda_i(T^s_{i,j+1}) - \Lambda_i(T^s_{ij}) \overset{\text{iid}}{\sim} F^s(\cdot),$$

$$\Lambda^s_i(T^c_{ij}|\mathcal{F}^s_{T^{c}_{ij}-}) - \Lambda^s_i\left[T_{i,N_i(T^c_{ij}-)} \,|\, \mathcal{F}^s_{T^-_{i,N_i(T^c_{ij}-)}}\right] \overset{\text{iid}}{\sim} F^c(\cdot), \quad i = 1, \cdots, n, \quad j = 1, 2, \cdots.$$

For system i, T^c_{ij} is the jth time of component event, and $T_{N_i(T^c_{ij}-)}$ denotes the most recent event time before T^c_{ij}. The cumulative intensity function for component events is computed as

$$\begin{aligned}\Lambda^c_i(t|\mathcal{F}_{t^-}) =& \int_0^t \lambda^c_i(u|\mathcal{F}_{u^-};\boldsymbol{\theta}^c)\, du \\ =& H^c\left\{\Lambda^s_i(t|\mathcal{F}^s_{t^-}) - \Lambda^s_i[t_{i,N_i(t^-)}|\mathcal{F}^s_{t^-_{i,N(t^-)}}]\right\} + \sum_{j=1}^{N_i(t^-)} H^c\left[\Lambda^s_i(t_{ij}|\mathcal{F}^s_{t^-_{ij}}) - \Lambda^s_i(t_{i,j-1}|\mathcal{F}^s_{t^-_{i,j-1}})\right],\end{aligned}$$

where $H^c(t) = \int_0^t h^c(u)\, du$ is the cumulative hazard function corresponding to F^c.

The MTRP model is a general model that includes the TRP, RP, NHPP and HPP models

as special cases. The intensity function (4.2) of the MTRP model reduces to the TRP model if $h^s(\cdot)$ is a constant function indicating no subsystem repair effect (i.e., only minimal repair effects from subsystem events). When both $h^s(\cdot)$ and $\lambda_i(t;\boldsymbol{\theta}^c)$ are constant functions, the MTRP model reduces to the RP model, indicating that subsystem events have no effect on component events, and component events are perfect repairs. When $h^c(\cdot)$, $h^s(\cdot)$ are constant functions and $\lambda_i(t;\boldsymbol{\theta}^c)$ is a function of t, the MTRP model reduces to an NHPP model, indicating component events are minimal repairs and there are no subsystem event effects. When $F^c(\cdot)$ is an Exp(1) distribution and both $h^s(\cdot)$ and $\lambda_i(t;\boldsymbol{\theta}^c)$ are constant functions, the MTRP model reduces to the HPP model. The likelihood ratio test can be used in the comparison and selection of these nested models. More details on the subject of model selection can be found in Lindqvist et al. (2003). Here, the application of the special cases, TRP, RP, NHPP, and HPP models, are slightly different from the usual application of these models because subsystem events induce censoring during the process operation. For example, if a subsystem is replaced, then observation is terminated on the component in that subsystem.

4.3 Parameter Estimation

The estimate of parameter $\boldsymbol{\theta}^s$ in (4.1) can be obtained by using the method of Lindqvist et al. (2003). So we focus on the estimation of the parameters in the component event model in (4.2).

4.3.1 The Likelihood Function

Note that component events for system i follow MTRP(F^c, F^s, λ_i). For the convenience in the expression of likelihood function, the component events are denoted by $\{t_{ij}, \delta_{ij}^c\}$, where $i = 1, \cdots, n$, and $j = 1, \cdots, N_i(\tau_i)$. Here, t_{ij} is the event time for system i, and δ_{ij}^c is the component event indicator. If the replacement is for a component, then indicator equals to one, otherwise zero. Let $t_{i0} = 0$, $t_{i,N_i(\tau_i)+1} = \tau_i$, $\Lambda_i^c(0|\mathcal{F}_0) = \Lambda_i^s(0|\mathcal{F}_0^s) = 0$, $\delta_{i,N_i(\tau_i)+1}^c = 0$, and $\mathcal{F} = \{N_{ic}(u), N_{is}(u), X_i(u) : 0 < u \leq \tau_i,\ i = 1, \cdots, n\}$. The likelihood function for component

events in MTRP model can be expressed as

$$
\begin{aligned}
L(\boldsymbol{\theta}^c; \mathcal{F}, \boldsymbol{w}) =& \prod_{i=1}^{n} \left(\left\{ \prod_{j=1}^{N_i(\tau_i)+1} [\lambda_i^c(t_{ij}|\mathcal{F}_{t_{ij}^-}; \boldsymbol{\theta}^c)]^{\delta_{ij}^c} \right\} \times \exp[-\Lambda_i^c(\tau_i|\mathcal{F}_{\tau_i^-})] \right) \\
=& \prod_{i=1}^{n} \prod_{j=1}^{N_i(\tau_i)+1} \left(\left\{ f^c[\Lambda_i^s(t_{ij}|\mathcal{F}_{t_{ij}^-}^s) - \Lambda_i^s(t_{i,j-1}|\mathcal{F}_{t_{i,j-1}^-}^s)] \lambda_i^s(t_{ij}|\mathcal{F}_{t_{ij}^-}^s; \boldsymbol{\theta}^c) \right\}^{\delta_{ij}^c} \right. \\
& \left. \times \left\{ S^c[\Lambda_i^s(t_{ij}|\mathcal{F}_{t_{ij}^-}^s) - \Lambda_i^s(t_{i,j-1}|\mathcal{F}_{t_{i,j-1}^-}^s)] \right\}^{1-\delta_{ij}^c} \right). \qquad (4.6)
\end{aligned}
$$

4.3.2 Estimation Procedure

We first discuss the estimation procedure for the unknown parameters in the model with random effects (i.e., HMTRP). Bayesian methods using diffuse prior distributions provide a convenient method to obtain the estimates of the unknown parameters. We suggest a Metropolis-within-Gibbs algorithm for this purpose. Define $\upsilon = 1/\sigma_r^2$ as the precision of the random effects distribution, and let $\upsilon \sim \text{Gamma}(a_1, a_2)$ be the conjugate prior distribution for the random effects. Here, a_1 is the shape parameter and a_2 is the rate parameter of a gamma distribution. Gelman (2006) and DePalma (2013) suggested using 0.001 for both a_1 and a_2. Then the mean and variance of the prior distribution of υ are 1 and 1000, respectively. We use a uniform distribution to describe prior information on $\boldsymbol{\theta}^c$. Define $L_i(\boldsymbol{\theta}^c|\mathcal{F}_{\tau_i}, w_i)$ as the conditional likelihood function of system i given random effect w_i, and $L(\boldsymbol{\theta}^c|\mathcal{F}, \boldsymbol{w}) = \prod_{i=1}^{n} L_i(\boldsymbol{\theta}^c|\mathcal{F}_{\tau_i}, w_i)$ as the conditional likelihood function for all n systems. Then, the pdf of full joint distribution of parameters in the MTRP model is

$$
P(\boldsymbol{\theta}^c, \boldsymbol{w}, \upsilon|\mathcal{F}) \propto L(\boldsymbol{\theta}^c|\mathcal{F}, \boldsymbol{w}) P(\boldsymbol{w}|\upsilon) P(\upsilon), \qquad (4.7)
$$

where $P(\boldsymbol{w}|\upsilon)$ is the pdf of a multivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}_r$. That is $\boldsymbol{w} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_r = \boldsymbol{I}/\upsilon)$, where $\boldsymbol{I}$ is an $n \times n$ identity matrix. The pdf of Gamma(a_1, a_2) is denoted by $P(\upsilon)$. Based on the full joint distribution, we can obtain

the joint posterior distribution of parameters in the model as follows,

$$w_i|\boldsymbol{\theta}^c, \upsilon \propto L_i(\boldsymbol{\theta}^c|\mathcal{F}_{\tau_i}, w_i)\upsilon^{1/2}\exp\left(-\frac{\upsilon w_i^2}{2}\right) \tag{4.8}$$

$$\upsilon|\boldsymbol{w}, \boldsymbol{\theta}^c \propto \text{Gamma}\left(\frac{n}{2}+a_1, \frac{\boldsymbol{w}'\boldsymbol{w}}{2}+a_2\right) \tag{4.9}$$

$$\boldsymbol{\theta}^c|\boldsymbol{w}, \upsilon \propto L(\boldsymbol{\theta}^c|\mathcal{F}, \boldsymbol{w}). \tag{4.10}$$

Algorithm 1:

1. Initialize all the parameters $\boldsymbol{w}^{(0)}$, $\upsilon^{(0)}$ and $\boldsymbol{\theta}^{c(0)}$.

2. Update $w_i^{(j)},\ i=1,\cdots,n$ using Metropolis algorithm at jth step:

a) Sample $w_i^* \sim \text{N}(w_i^{(j-1)}, \sigma_{w_i}^2)$, where $\sigma_{w_i}^2$ is the variance of the proposal distribution for unit i.

b) Accept w_i^* as $w_i^{(j)}$ with the probability
$\min\left\{\frac{L_i\left(\boldsymbol{\theta}^{c(j-1)}|\mathcal{F}_{\tau_i}, w_i^*\right)\exp\left(-\frac{\upsilon^{(j-1)}w_i^{*2}}{2}\right)}{L_i\left(\boldsymbol{\theta}^{c(j-1)}|\mathcal{F}_{\tau_i}, w_i^{(j-1)}\right)\exp\left(-\frac{\upsilon^{(j-1)}\left[w_i^{(j-1)}\right]^2}{2}\right)}, 1\right\}$, otherwise, set $w_i^{(j)} = w_i^{(j-1)}$.

3. Sample $\upsilon^{(j)} \sim \text{Gamma}\left(\frac{n}{2}+a_1, \frac{\boldsymbol{w}^{(j)\prime}\boldsymbol{w}^{(j)}}{2}+a_2\right)$.

4. Update values of elements of $\boldsymbol{\theta}^{c(j)} = (\theta_1^{c(j)}, \cdots, \theta_i^{c(j)}, \cdots, \theta_z^{c(j)})$ successively at the jth step. Let $\boldsymbol{\theta}_{i-1}^{c(j)} = (\theta_1^{c(j)}, \cdots, \theta_{i-1}^{c(j)}, \theta_i^{c(j-1)}, \cdots, \theta_z^{c(j-1)})$. Note that $\boldsymbol{\theta}_0^{c(j)} = \boldsymbol{\theta}^{c(j-1)}$.

a) Sample $\theta_i^{c*} \sim \text{N}(\theta_i^{c(j-1)}, \sigma_{\theta_i^c}^2)$, where $\sigma_{\theta_i^c}^2$ is the variance of the proposal distribution. Let $\boldsymbol{\theta}_i^{c*} = (\theta_1^{c(j)}, \cdots, \theta_{i-1}^{c(j)}, \theta_i^{c*}, \theta_{i+1}^{c(j-1)}, \cdots, \theta_z^{c(j-1)})$.

b) Accept θ_i^{c*} as $\theta_i^{c(j)}$ in $\boldsymbol{\theta}_i^{c(j)}$ with the probability
$\min\left\{\frac{L[\boldsymbol{\theta}_i^{c*}|\mathcal{F}, \boldsymbol{w}^{(j)}]}{L[\boldsymbol{\theta}_{i-1}^{c(j)}|\mathcal{F}, \boldsymbol{w}^{(j)}]}, 1\right\}$, otherwise, set $\theta_i^{c(j)} = \theta_i^{c(j-1)}$ in $\boldsymbol{\theta}_i^{c(j)}$.

c) Repeat steps a) and b) for $i=1,\cdots,z$ at the given j. Let $\boldsymbol{\theta}^{c(j)} = \boldsymbol{\theta}_z^{c(j)}$.

5. Repeat steps 2-4 for a large number of times until a sufficient number of draws from the joint posterior distribution have been obtained.

To obtain optimal acceptance rates (around .44 according to Gelman et al., 1997 and Roberts and Rosenthal, 2001), the tuning parameters σ_{w_i} and $\sigma_{\theta_i^c},\ i = 1, \cdots, z$ can be well adjusted by applying the method given in Roberts and Rosenthal (2009).

For model without random effects (MTRP), **Algorithm 1** can still be used by omitting steps 1-3. An alternative way to estimate the unknown parameters is the maximum likelihood (ML) method based on the likelihood function (4.6).

Once the parameters estimates in the MTRP/HMTRP model are obtained, the residuals of the model can be estimated by using the cumulative hazard function. That is, the residuals $R_{ij} = H^c[\Lambda_i^s(t_{ij}|\mathcal{F}^s_{t_{ij}^-}) - \Lambda_i^s(t_{i,j-1}|\mathcal{F}^s_{t_{i,j-1}^-})]$. The residuals $(R_{ij}, \delta_{ij}^c = 1)$ are expected to behave as samples from the Exp(1) distribution, which can be used to evaluate the goodness of fit of the model.

4.4 Prediction for Component Events

4.4.1 Point Prediction

Accurate prediction of future events is important to the manufacturer of products, or the operators of fleets of systems, for purposes of controlling operating costs, optimizing the number of spare components, and assessing the risk of excessive warranty returns. Here, we focus on the prediction of events at component level.

Let $\boldsymbol{\theta}_x$ as the parameters in the model of time-dependent covariate, and $\boldsymbol{X}_i(t_1, t_2) = \{X_i(t);\ t_1 < t \leq t_2\}$. The predicted cumulative number of component events up to time $m > 0$ for a fleet of n units can be obtained by:

$$\begin{aligned} N_c(m; \boldsymbol{\theta}^c, \boldsymbol{\theta}^s, \boldsymbol{\theta}_x) &= \sum_{i=1}^{n} N_{ic}(m; \boldsymbol{\theta}^c, \boldsymbol{\theta}^s, \boldsymbol{\theta}_x) \\ &= \sum_{i=1}^{n} \mathrm{E}_{\boldsymbol{X}_i(\tau_i, \tau_i+m)|\boldsymbol{X}(\tau_i)} \mathrm{E}_{w_i} \big\{ N_{ic}[m, \boldsymbol{X}_i(\tau_i, \tau_i + m), w_i; \boldsymbol{\theta}^c, \boldsymbol{\theta}^s, \boldsymbol{\theta}_x] \big\}, \end{aligned} \tag{4.11}$$

where $N_{ic}(m; \boldsymbol{\theta}^c, \boldsymbol{\theta}^s, \boldsymbol{\theta}_x)$ is the predicted cumulative number of component events in system i up to time m after the DFD. Parameter vector $\boldsymbol{\theta}^s$ is needed in the prediction of component events, because the model for component events depends on the history of subsystem events.

Because the closed-form expression for (4.11) is not available, numerical methods or Monte Carlo simulation can be used. For prediction in the general recurrent process, Monte Carlo simulation is more common and easier for computation as illustrated in Yańez et al. (2002) and Franz et al. (2013a).

4.4.2 Prediction for the Time-Dependent Covariate

To predict future recurrent events for a system with a time-dependent covariate, it is necessary to have a parametric model for the covariate process. Based on the covariate pattern shown in Figure 4.1(b), we use a linear mixed-effects model to fit the time-dependent covariate data. In particular,

$$X_i(t_{ik}) = t_{ik}(\beta_x + \nu_i) + \epsilon_i(t_{ik}) \quad i = 1, \cdots, n, \quad k = 1, \cdots, n_i, \tag{4.12}$$

where β_x is the coefficient of time, ν_i is the random effect, and $\epsilon_i(t_{ik})$ is the error term. We assume that $\nu_i \overset{\text{iid}}{\sim} \text{N}(0, \sigma_\nu^2)$, and $\epsilon_i(t_{ik}) \overset{\text{iid}}{\sim} \text{N}(0, \sigma_x^2)$ which is independent of ν_i. The parameters in (4.12) are denoted by $\boldsymbol{\theta}_x = (\beta_x, \sigma_\nu, \sigma_x)'$. The estimation of $\boldsymbol{\theta}_x$ in the covariate model can be accomplished by using existing software packages (e.g., using the R function `lme`).

We use an approach that is similar to Hong and Meeker (2013) for the covariate prediction. Let $\boldsymbol{t}_i = (t_{i1}, \cdots, t_{in_i})'$, $\boldsymbol{t}_{im} = (t_{i,n_{i+1}}, \cdots, t_{i,n_i+m_i})'$ be the observed time points before τ_i and the predicted time points during $(\tau_i, \tau_i + m]$, respectively. Let $\boldsymbol{X}_i(\boldsymbol{t}_i) = [X_i(t_{i1}), \cdots, X_i(t_{in_i})]'$, $\boldsymbol{X}_i(\boldsymbol{t}_{im}) = [X_i(t_{i,n_{i+1}}), \cdots, X_i(t_{i,n_i+m_i})]'$ be the corresponding time-dependent covariate processes. Here, m_i is the number of predicted time points for system i. The joint distribution of $\boldsymbol{X}_i(\boldsymbol{t}_i)$ and $\boldsymbol{X}_i(\boldsymbol{t}_{im})$ can be expressed as

$$\text{N}\left[\begin{pmatrix} \boldsymbol{t}_i \\ \boldsymbol{t}_{im} \end{pmatrix} \beta_x, \quad \begin{pmatrix} \boldsymbol{\Sigma}_{i11} & \boldsymbol{\Sigma}_{i12} \\ \boldsymbol{\Sigma}_{i21} & \boldsymbol{\Sigma}_{i22} \end{pmatrix}\right],$$

where $\boldsymbol{\Sigma}_{i11} = \sigma_\nu^2 \boldsymbol{t}_i \boldsymbol{t}_i' + \sigma_x^2 \boldsymbol{I}_{n_i}$, $\boldsymbol{\Sigma}_{i22} = \sigma_\nu^2 \boldsymbol{t}_{im} \boldsymbol{t}_{im}' + \sigma_x^2 \boldsymbol{I}_{m_i}$, and $\boldsymbol{\Sigma}_{i12} = \sigma_\nu^2 \boldsymbol{t}_i \boldsymbol{t}_{im}'$. Here, $\boldsymbol{I}_{n_i}$ and $\boldsymbol{I}_{m_i}$ are $n_i \times n_i$ and $m_i \times m_i$ identity matrices. The conditional distribution of $\boldsymbol{X}_i(\boldsymbol{t}_i) | \boldsymbol{X}_i(\boldsymbol{t}_{im})$ is

$$\text{N}\Big(\boldsymbol{t}_{im}\beta_x + \boldsymbol{\Sigma}_{i21}\boldsymbol{\Sigma}_{i11}^{-1}[\boldsymbol{X}_i(\boldsymbol{t}_i) - \boldsymbol{t}_i\beta_x], \ \boldsymbol{\Sigma}_{i22} - \boldsymbol{\Sigma}_{i21}\boldsymbol{\Sigma}_{i11}^{-1}\boldsymbol{\Sigma}_{i12}\Big). \tag{4.13}$$

Based on (4.13), the future values of the time-dependent covariate processes can be predicted.

4.4.3 Subsystem Event Simulations

Because the model for component events depends on the history of subsystem events, the simulation of subsystem events is needed in the prediction of component events. Let $\varsigma_i = \tau_i + m$ be the end of the prediction time for system i, $\widehat{F}^{s*}$ be the estimate of renewal distribution function F^{s*}, $\widehat{\Lambda}_i^*$ be the estimate of Λ_i^*, and $\widehat{\Lambda}_i^{*-1}(\cdot)$ be the corresponding inverse function given $\widehat{\boldsymbol{\theta}}^s$ and $\widehat{\boldsymbol{\theta}}_x$. Here, $\widehat{\boldsymbol{\theta}}^s$ and $\widehat{\boldsymbol{\theta}}_x$ are ML estimates of $\boldsymbol{\theta}^s$ and $\boldsymbol{\theta}_x$, respectively. Based on the definition of the TRP model, the gaps between two consecutive transformed subsystem event times follow distribution F^{s*}. That is, $\Lambda_i^*(t_{i,j+1}^s) - \Lambda_i^*(t_{ij}^s) \overset{\text{iid}}{\sim} F^{s*}$, where $i = 1, \cdots, n$ and $j = 1, 2, \cdots$. The subsystem events can be simulated as follows.

Algorithm 2

1. Generate the ith predicted time-dependent covariate process based on $\widehat{\boldsymbol{\theta}}_x$ using the conditional distribution (4.13).

2. Compute $\widehat{\Lambda}_i^*(\varsigma_i)$ as the length of study in the transformed scale.

3. Generate the random variable U_{ij} from distribution $\widehat{F}^{s*}$ and obtain the sequence of times $T_{ij}^* = \widehat{\Lambda}_i^*[t_{i,N_{is}(\tau_i)}^s] + \sum_{k=1}^j U_{ik},\ j = 1, \cdots, C_i^s$, until $T_{i,C_i^s+1}^* > \widehat{\Lambda}_i^*(\varsigma_i)$. Then, C_i^s is the number of simulated subsystem events for unit i.

4. Compute the simulated event times $T_{ij}^s = \widehat{\Lambda}_i^{*-1}(T_{ij}^*),\ j = 1, \cdots, C_i^s$.

5. Repeat steps 1-4 for each system i, where $i = 1, \cdots, n$.

Note that in step 3 the time of the first simulated event should be larger than τ_i, because the simulation is conditional on the history. Otherwise it needs to be re-simulated.

4.4.4 Point Prediction Computing

According to the definition of the MTRP model, the gaps (i.e., $d_1, d_2, \cdots$) between the component event and the most recent event (either component event or subsystem event) in the transformed time scale [i.e., $\Lambda^s(t)$] follow the F^c distribution as shown in Figure 4.4. Because

the component event process is censored by the subsystem events, the component events can be simulated in the intervals (i.e., $I_1, I_2, ...$ in Figure 4.4) of the subsystem events under the transformed time [i.e., $\Lambda^s(t)$] scale. In particular, for each system,

$$\Lambda_i^s(T_{ij}^c|\mathcal{F}_{T_{ij}^{c}-}^s) - \Lambda_i^s\left[T_{i,N_i(T_{ij}^{c}-)} \,|\, \mathcal{F}^s_{T^-_{i,N_i(T_{ij}^{c}-)}}\right] \overset{\text{iid}}{\sim} F^c(\cdot),$$

where $i = 1, \cdots, n$ and $j = 1, 2, \cdots$. Let $\widehat{\boldsymbol{\theta}}^c$ and $\widehat{\upsilon}$ be the maximum a posterior (MAP) estimates of $\boldsymbol{\theta}^c$ and υ, respectively. The prediction of the cumulative number of component events, $\widehat{N}_c(m; \widehat{\boldsymbol{\theta}}^c, \widehat{\boldsymbol{\theta}}^s, \widehat{\boldsymbol{\theta}}_x)$, can be computed by using the following algorithm.

Algorithm 3

1. Repeat **Algorithm 2** steps 1-4. Then the predicted covariate process and simulated times of subsystem events for system i are obtained. The simulated subsystem events times are denoted by $t^s_{i,N_{is}(\tau_i)+1}, \cdots, t^s_{i,N_{is}(\tau_i)+C_i^s}$. Here C_i^s is the simulated number of subsystem events. Set $t^s_{i,N_{is}(\tau_i)+0} = t^s_{i,N_{is}(\tau_i)}$ and $t^s_{i,N_{is}(\tau_i)+C_i^s+1} = \varsigma_i$.

2. The random effect w_i can be obtained by the Metropolis algorithm. The conditional pdf of w_i is proportional to the product of $L_i\left(\widehat{\boldsymbol{\theta}}^c|\mathcal{F}_t, w_i\right)$ and $P(w_i|\widehat{\boldsymbol{\theta}}^c, \widehat{\upsilon})$. For the MTRP model, this step can be skipped.

3. Calculate the simulated time intervals which are separated by the simulated subsystem events for system i in the transformed time scale. That is $I_{ik} = \widehat{\Lambda}_i^s\left[\widehat{\Lambda}_i(t^s_{i,N_{is}(\tau_i)+k})\right] - \widehat{\Lambda}_i^s\left[\widehat{\Lambda}_i(t^s_{i,N_{is}(\tau_i)+k-1})\right]$, $k = 1, \cdots, C_i^s + 1$, where $\widehat{\Lambda}_i^s(\cdot)$ and $\widehat{\Lambda}_i(\cdot)$ are the estimates of $\Lambda_i^s(\cdot)$ and $\Lambda_i(\cdot)$ based on the estimates $\widehat{\boldsymbol{\theta}}^c$ and $\widehat{\boldsymbol{\theta}}_x$, respectively.

4. For each simulated time interval, generate random variables U_{ikl} from distribution $\widehat{F}^c$, while $\sum_l U_{ikl} \leq I_{ik}$. The number of generated U_{ikl} is recorded as C^c_{ik}.

5. Calculate the number of simulated component events in system i. That is $C_i^c = \sum_{k=1}^{C_i+1} C^c_{ik}$.

6. Repeat steps 1-5 for each system i, where $i = 1, \cdots, n$.

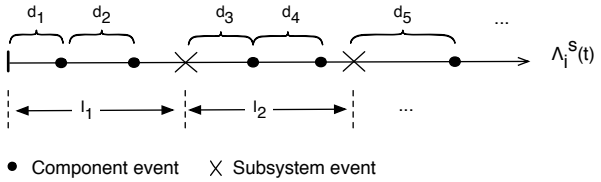


Figure 4.4: Illustration of component event simulation.

7. Repeat steps 1-6 B times and a series of $C_i^{c(b)},\ b = 1, \cdots, B,\ i = 1, \cdots, n$ are obtained. Then, $\widehat{N}_c(m; \widehat{\boldsymbol{\theta}}^c, \widehat{\boldsymbol{\theta}}^s, \widehat{\boldsymbol{\theta}}_x) = \sum_{i=1}^{n} \sum_{b=1}^{B} C_i^{c(b)}/B$.

Note that in step 4 the time of the first simulated event should be larger than τ_i. Otherwise it need to be re-simulated.

4.4.5 Prediction Interval Computing

In order to obtain PIs for the cumulative number of events, we also need to take into account the distribution of estimated parameters as well as the uncertainty in the future values of the time-dependent covariate. Instead of sampling the posterior distributions, we use multivariate normal distribution to approximate the distributions of parameter estimators for fast computing. We focus on the PI for the cumulative number of component events. The algorithm is described as follows.

Algorithm 4

1. Simulate $\widehat{\boldsymbol{\theta}}_x^*$, $\widehat{\boldsymbol{\theta}}^{s*}$, $\widehat{\boldsymbol{\theta}}^{c*}$, and $\widehat{\upsilon}^*$ from $\mathrm{N}(\widehat{\boldsymbol{\theta}}_x, \widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\theta}}_x})$, $\mathrm{N}(\widehat{\boldsymbol{\theta}}^s, \widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\theta}}^s})$, $\mathrm{N}(\widehat{\boldsymbol{\theta}}^c, \widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\theta}}^c})$ and $\mathrm{N}(\widehat{\upsilon}, \widehat{\sigma}_{\widehat{\upsilon}}^2)$, respectively.

2. Replace $\widehat{\boldsymbol{\theta}}_x$ by $\widehat{\boldsymbol{\theta}}_x^*$, $\widehat{\boldsymbol{\theta}}^s$ by $\widehat{\boldsymbol{\theta}}^{s*}$, $\widehat{\boldsymbol{\theta}}^c$ by $\widehat{\boldsymbol{\theta}}^{c*}$, and $\widehat{\upsilon}$ by $\widehat{\upsilon}^*$, and repeat steps 1-7 in **Algorithm 3** to obtain $\widehat{N}_c^*(m; \widehat{\boldsymbol{\theta}}^{c*}, \widehat{\boldsymbol{\theta}}^{s*}, \widehat{\boldsymbol{\theta}}_x^*)$.

3. Repeat steps 1-2 B times to obtain $\widehat{N}_c^{*(b)}(m; \widehat{\boldsymbol{\theta}}^{c*}, \widehat{\boldsymbol{\theta}}^{s*}, \widehat{\boldsymbol{\theta}}_x^*)$, where $b = 1, \cdots, B$.

4. The $100(1-\alpha)\%$ PI for N_c is the $(\alpha/2, 1-\alpha/2)$ quantile of the B ordered values of $\widehat{N}_c^{*(b)}(m; \widehat{\boldsymbol{\theta}}^{c*}, \widehat{\boldsymbol{\theta}}^{s*}, \widehat{\boldsymbol{\theta}}_x^*)$.

4.5 Finite-Sample Performance of Estimation Methods

In this section, we use simulation to study the effect of sample size and number of events performance of the estimation methods.

4.5.1 Design of Simulations

In the simulation, the time is the calendar time, and the time for repair is ignorable. Only one time-dependent covariate is considered with the form of (4.12) and 30 observed time points per system. The parameter settings are $\boldsymbol{\theta}_x = (\beta_x, \sigma_\nu, \sigma_x)' = (0.02, 0.004, 0.05)'$.

The subsystem events follow a TRP model with trend function $\lambda_i^*(t; \boldsymbol{\theta}^s) = at^{a-1}\exp\{\kappa \log[X_i(t)]\}$ and the renewal distribution function F^{s*}. We set F^{s*} to be a Weibull distribution and the corresponding hazard function is $h^*(t) = (\beta/\eta)(t/\eta)^{\beta-1}$, where η is the scale parameter (also the approximate .63 quantile) and β is the shape parameter. Because the mean of renewal function is restricted to one, the corresponding hazard function can be expressed as

$$h^*(t) = \Gamma(1+\sigma)^{\frac{1}{\sigma}} t^{\frac{1}{\sigma}-1} \left(\frac{1}{\sigma}\right), \tag{4.14}$$

where $\sigma = 1/\beta$. The parameters for the subsystem are $\boldsymbol{\theta}^s = (a, \sigma, \kappa)' = (0.3, 0.8, 0.8)'$.

Given the above simulated subsystem events, the simulation of component events is based on the HMTRP model with the consideration of time-dependent covariate and random effects. Let $\lambda_b = (\alpha/\varphi)(t/\varphi)^{\alpha-1}$. The trend function of system i in (4.5) can be denoted by $\lambda_i(t; \boldsymbol{\theta}^c) = (\alpha/\varphi)(t/\varphi)^{\alpha-1}\exp\{\gamma \log[X_i(t)] + w_i\}$, where $w_i \sim \mathrm{N}(0, \sigma_r^2)$. Here, φ is set to be one in the HMTRP model. The renewal distributions for subsystem F^s and component F^c are both Weibull distributions with mean one. The mean-one restriction is used so that all model parameters are identifiable (Lindqvist et al., 2003). Similar to (4.14), the hazard functions can be expressed as $h^c(t) = \Gamma(1+\sigma_0)^{1/\sigma_0} t^{1/\sigma_0 - 1}(1/\sigma_0)$, and $h^s(t) = \Gamma(1+\sigma_1)^{1/\sigma_1} t^{1/\sigma_1 - 1}(1/\sigma_1)$,

respectively. Then the intensity function of HMTRP can be expressed as

$$\lambda_i^c(t|\mathcal{F}_{t^-}) = h^c\left\{\Lambda_i^s(t|\mathcal{F}^s_{t^-}) - \Lambda_i^s\left[t_{i,N_i(t^-)} \mid \mathcal{F}^s_{t^-_{i,N_i(t^-)}}\right]\right\}h^s\left\{\Lambda_i(t) - \Lambda_i[t^s_{i,N_{is}(t^-)}]\right\}\lambda_i(t;\boldsymbol{\theta}^c), \tag{4.15}$$

where

$$\lambda_i(t;\boldsymbol{\theta}^c) = (\alpha/\varphi)(t/\varphi)^{\alpha-1}\exp\{\gamma\log[X_i(t)] + w_i\}.$$

The parameters for the component event intensity are $\boldsymbol{\theta}^c = (\alpha, \sigma_0, \sigma_1, \gamma)' = (0.4, 0.6, 0.75, 0.9)'$ and $\sigma_r = 0.5$.

The number of systems n was selected to be 50, 100, and 200. For each value of n, we simulated data 1000 times based on the parameter settings. By selecting different DFDs, the expected number of subsystem events (m_1) and component events (m_2) in each sample size (i.e., $n = 50, 100$, and 200) were controlled to the following three combinations (m_1, m_2): (0.7, 1.3), (1.7, 4.3), and (3.9, 12.7).

4.5.2 Simulation Results

Based on the **Algorithm 1**, the estimated parameters and corresponding standard errors (SE) for the component events model were obtained and are shown in Tables 4.1-4.3. In the results, the estimates are close to the true values of the parameters, and they are approximating to the true settings as the sample size (n) and the length of study time (DFD) increase. Also, the coverage probability (CP) of the confidence interval procedure for each parameter is close to the nominal value of .95. The results show that our proposed estimation procedure can estimate the parameters well.

4.6 Application in Vehicle B Data

In this section, we use the simulated Vehicle B data to illustrate our proposed method in the estimation and prediction of component events. The Vehicle B data set is simulated based on the models and parameters setting in Section 4.5 with $n = 100$. The event plot and covariate plot have been displayed in Figure 4.1.

Table 4.1: Summary of the simulation studies of the HMTRP given average number of subsystem events $m_1 = 0.7$ and component events $m_2 = 1.3$.

Sample size	Parameter	Value	Mean	Bias	SE	MSE$\times 10^3$	CP
$n = 50$	α	.40	.3999	.0001	.0285	.8132	.940
	σ_0	.60	.6274	.0274	.0953	9.831	.938
	σ_1	.75	.7969	.0469	.1475	23.96	.927
	γ	.90	1.003	.1030	.2267	61.99	.930
	σ_r	.50	.5142	.0142	.1709	29.43	.911
$n = 100$	α	.40	.4008	.0008	.0187	.3510	.937
	σ_0	.60	.6136	.0136	.0627	4.123	.925
	σ_1	.75	.7697	.0197	.0855	7.700	.934
	γ	.90	.9462	.0462	.1181	16.08	.942
	σ_r	.50	.5103	.0103	.1015	10.42	.926
$n = 200$	α	.40	.4003	.0003	.0134	.1793	.930
	σ_0	.60	.6068	.0068	.0418	1.794	.938
	σ_1	.75	.7572	.0072	.0561	3.204	.942
	γ	.90	.9167	.0167	.0795	6.606	.933
	σ_r	.50	.5041	.0041	.0594	3.549	.951

Table 4.2: Summary of the simulation studies of the HMTRP given average number of subsystem events $m_1 = 1.7$ and component events $m_2 = 4.3$.

Sample size	Parameter	Value	Mean	Bias	SE	MSE$\times 10^3$	CP
$n = 50$	α	.40	.3989	.0011	.0142	.2027	.935
	σ_0	.60	.6023	.0023	.0380	1.447	.940
	σ_1	.75	.7547	.0047	.0429	1.866	.953
	γ	.90	.9102	.0102	.0701	5.014	.920
	σ_r	.50	.5105	.0105	.0774	6.095	.930
$n = 100$	α	.40	.3995	.0005	.0097	.0947	.942
	σ_0	.60	.6013	.0013	.0258	.6676	.947
	σ_1	.75	.7523	.0023	.0295	.8748	.945
	γ	.90	.9059	.0059	.0443	2.000	.946
	σ_r	.50	.5042	.0042	.0544	2.981	.922
$n = 200$	α	.40	.3995	.0005	.0066	.0440	.951
	σ_0	.60	.6001	.0001	.0181	.3277	.943
	σ_1	.75	.7510	.0010	.0212	.4523	.938
	γ	.90	.9025	.0025	.0319	1.026	.951
	σ_r	.50	.5030	.0030	.0361	1.310	.953

Table 4.3: Summary of the simulation studies of the HMTRP given average number of subsystem events $m_1 = 3.9$ and component events $m_2 = 12.7$.

Sample size	Parameter	Value	Mean	Bias	SE	MSE$\times 10^3$	CP
$n = 50$	α	.40	.3996	.0004	.0125	.1554	.931
	σ_0	.60	.6008	.0008	.0202	.4074	.940
	σ_1	.75	.7514	.0014	.0198	.3955	.926
	γ	.90	.9033	.0033	.0350	1.237	.931
	σ_r	.50	.5073	.0073	.0615	3.836	.934
$n = 100$	α	.40	.3998	.0002	.0087	.0762	.933
	σ_0	.60	.6003	.0003	.0140	.1957	.935
	σ_1	.75	.7503	.0003	.0139	.1936	.934
	γ	.90	.9015	.0015	.0250	.6289	.942
	σ_r	.50	.5041	.0041	.0421	1.790	.944
$n = 200$	α	.40	.3999	.0001	.0060	.0366	.939
	σ_0	.60	.6003	.0003	.0096	.0925	.942
	σ_1	.75	.7503	.0003	.0092	.0842	.951
	γ	.90	.9008	.0008	.0176	.3108	.932
	σ_r	.50	.5010	.0010	.0296	.8791	.940

4.6.1 Parameter Estimation

Table 4.4 lists the estimates and standard errors of parameters in the covariate model (4.12), subsystem events model and component events models. The Vehicle B data are also fitted by the sub-models of HMTRP (i.e., HTRP, HRP and HNHPP). All the estimates in the Table 4.4 are close to the true settings. In the TRP model of subsystem events, the estimate of the Weibull shape parameter $1/\sigma$ is greater than one indicating the trend of the hazard function corresponding to the renewal function is increasing for subsystem events.

For the component events, several sub-models are compared to the HMTRP model, and we use the deviance information criterion (DIC) as a criterion for Bayesian model selection. Similar to the Akaike information criterion (AIC), it considers the model adequacy and model complexity. Define the deviance as $D(\boldsymbol{\theta}) = -2\log[f(y|\boldsymbol{\theta})] + 2\log[g(y)]$, where $\boldsymbol{\theta}$ denotes the vector of unknown parameters, $f(y|\boldsymbol{\theta})$ is the likelihood function and $g(y)$ is a standardizing term. Then, the DIC can be expressed as $\text{DIC} = \bar{D} + p_D$, where $\bar{D}$ indicates the goodness of fit with the form of $\bar{D} = \text{E}_{\boldsymbol{\theta}|y}[D(\boldsymbol{\theta})] = \text{E}_{\boldsymbol{\theta}|y}[-2\ln f(y|\boldsymbol{\theta})]$ by setting $g(y) = 1$, and p_D indicates

Table 4.4: Parameter estimates and standard errors component event, subsystem, and covariate models, based on Vehicle B data.

Component Event Models									Subsystem Event Model		
Models		α	φ	γ	σ_0	σ_1	σ_r	DIC	a	σ	κ
HMTRP	est.	.40	1	.93	.60	.77	.53	7538.5	.31	.76	.74
	SE	.01	–	.03	.01	.01	.05		.01	.03	.05
HTRP	est.	.41	1	1.1	.60	1	.71	7722.3	Covariate Model		
	SE	.01	–	.03	.01	–	.05		β_x	σ_ν	σ_x
HRP	est.	1	19.8	0	.70	1	.72	8167.9	.0202	.0042	.0494
	SE	–	1.72	–	.02	–	.06		.0004	.0003	.0006
HNHPP	est.	1.66	54.3	0	1	1	.78	8160.1			
	SE	.04	3.4	–	–	–	.08				

the penalty for model complexity with the form of $p_D = \bar{D} - D(\bar{\boldsymbol{\theta}})$. Here, $\bar{\boldsymbol{\theta}}$ is the posterior mean of $\boldsymbol{\theta}$. By simple transformation, DIC can be re-expressed as DIC $= D(\bar{\boldsymbol{\theta}}) + 2p_D$, which is similar to the form of AIC. More detail about DIC can be found in Spiegelhalter et al. (2002), and Berg et al. (2004).

The DIC is easy to be computed via MCMC method. The estimate of $\bar{D}$ and p_D can be computed by $\bar{D} = \sum_{b=1}^{B}[-2L(\boldsymbol{\theta}^{c(b)}|\mathcal{F}_t, \boldsymbol{w}^{(b)})]/B$ and $p_D = \bar{D} - [-2L(\widehat{\boldsymbol{\theta}}^c|\mathcal{F}_t, \widehat{\boldsymbol{w}})]$ with $\widehat{\boldsymbol{\theta}}^c = \sum_{b=1}^{B}\boldsymbol{\theta}^{c(b)}/B$ and $\widehat{\boldsymbol{w}} = \sum_{b=1}^{B}\boldsymbol{w}^{(b)}/B$, respectively. Here, $\boldsymbol{\theta}^{c(b)}$ and $\boldsymbol{w}^{(b)}$ are the simulated posterior estimates after burn in. The results of DIC in Table 4.4 show that the HMTRP model fits the data better than other sub-models because of the smallest value of DIC. Both the estimates of $1/\sigma_0$ and $1/\sigma_1$ are larger than one indicating that the $h^c(\cdot)$ and $h^s(\cdot)$ are increasing functions, and subsystem/component events have effect on the intensity trend of component events.

To check the goodness of fit of the model, we use the Cox-Snell residuals plot. The estimated residuals $\widehat{R}_{ij}$ (with δ_{ij}^c as the censoring indicator) are expected to behave like a censored sample from an Exp(1) distribution. Figure 4.5 is a Cox-Snell residual plot which shows that the HMTRP model provides a good fit to the data, as most of the points align well with the 45° line.

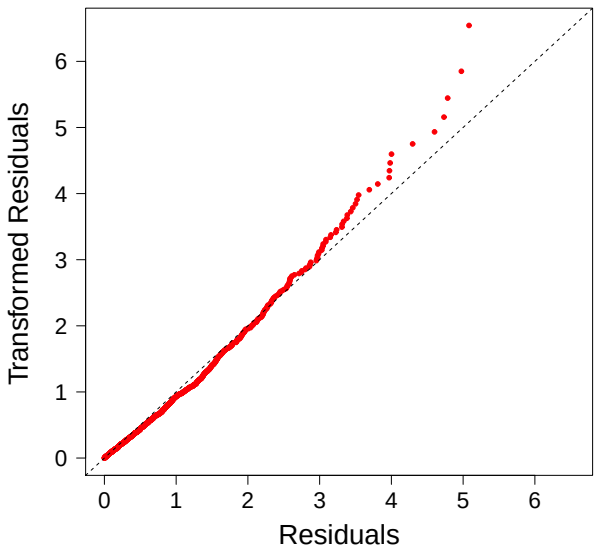


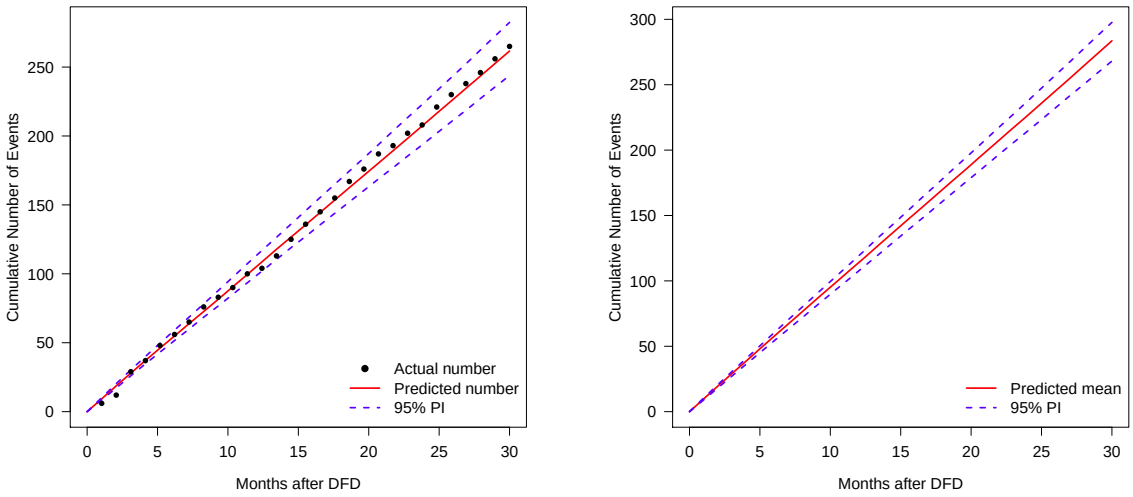
Figure 4.5: Residual plot for the HMTRP model in Vehicle B data.

4.6.2 Prediction Results

We validate the prediction procedure based on the Vehicle B data by setting τ_i back 30 months. The prediction results are shown in Figure 4.6(a). The actual cumulative numbers of component events are all contained in the 95% PIs indicating that our prediction procedure performs quite well. After the validation of the prediction method, we predict the cumulative number of component events in the next 30 months after the DFD, shown in Figure 4.6(b).

4.7 Conclusion and Discussion

In this chapter, we propose an MTRP to model component events in multi-level repairable systems by extending the TRP method. Based on the MTRP, we also give Monte Carlo based procedures to provide point predictions and PIs for the cumulative number of future replacement events. The proposed MTRP model is a general recurrence process which includes the TRP, RP, and NHPP models as special cases. Using likelihood ratio tests or other criterions (e.g., AIC, DIC, etal.), we can select the appropriate sub-model and determine the existence



(a) Back test (b) Prediction

Figure 4.6: Plots of the predicted cumulative number of component events in Vehicle B data.

of effects from subsystem replacement events and component replacement events as well as the shape of the respective intensity functions. To explain more system-to-system variability, time-dependent covariates as well as random effects are introduced into the heterogeneous MTRP model (i.e., HMTRP). A Metropolis-within-Gibbs algorithm is suggested to estimate the unknown parameters in the HMTRP model. Performance of the estimation and prediction were checked with simulation studies. An industrial application is also used to explain the proposed method. Although one time-dependent covariate is considered in this chapter, the extension to multiple covariates is straightforward.

In the future related research, several possible areas can be continued.

- Some misspecification models can be considered in the simulation studies and application.
- The proposed model and methods can apply to the system with more levels, although we just consider two levels in this chapter.
- A more complex system with multiple types of events (e.g., different failure modes) at the component level and at the subsystem level can be considered.

- Current model could also be extended to consider events that occur in many subsystems, with the possibility of interaction between subsystems.
- Relating physical models for failure to the empirical replacement data and models for cumulative damage.

Bibliography

A. Berg, R. Meyer, and J. Yu. Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics*, 22:107–120, 2004.

M. Brown and F. Proschan. Imperfect repair. *Journal of Applied Probability*, 20:851–859, 1983.

G. DePalma. Bugs Bayesian inference using Gibbs sampling. 2013. URL http://www.stat.purdue.edu/ gdepalma/Sec

L. Doyen and O. Gaudoin. Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering and System Safety*, 84:45–56, 2004.

J. Franz, A. Jokiel-Rokita, and R. Magiera. Prediction in trend-renewal processes for repairable systems. *Statistics and Computing*, pages DOI 10.1007/s11222–013–9393–5, 2013a.

J. Franz, A. Jokiel-Rokita, and R. Magiera. Prediction in trend-renewal processes for repairable systems. *Statistics and Computing*, 2013b. Published on line.

A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1:515–533, 2006.

A. Gelman, W. R. Gilks, and G. O. Roberts. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Annals of Applied Probability*, 7:110–120, 1997.

K. Heggland and B. Lindqvist. A non-parametric monotone maximum likelihood estimator of time trend for repairable system data. *Reliability Engineering & System Safety*, 92:575–584, 2007.

Y. Hong and W. Q. Meeker. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics*, in press, DOI:10.1080/00401706.2013.765324, 2013.

Y. Hong, M. Li, and B. Osborn. System unavailability analysis based on window-observed recurrent event data. *Applied Stochastic Models in Business and Industry*, 2013. doi: 10.1002/asmb.1984.

V. E. Johnson, A. Moosman, and P. Cotter. A hierarchical model for estimating the early reliability of complex systems. *IEEE Transactions on Reliability*, 54:224–231, 2005.

M. Kijima. Some results for repairable systems with general repair. *Journal of Applied Probability*, 26:89–102, 1989.

J. Lawless and K. Thiagarajah. A point-process model incorporating renewals and time trends, with application to repairable systems. *Technometrics*, 38:131–138, 1996.

L. M. Leemis. Technical note: Nonparametric estimation and variate generation for a nonhomogeneous Poisson process from event count data. *IIE Transactions*, 36:1155–1160, 2004.

B. Lindqvist. On the statistical modeling and analysis of repairable systems. *Statistical Science*, 21:532–551, 2006.

B. Lindqvist, G. Elvebakk, and K. Heggland. The trend-renewal process for statistical analysis of repairable systems. *Technometrics*, 45:31–44, 2003.

J. Liu, J. Li, and B. U. Kim. Bayesian reliability modeling of multi-level system with interdependent subsystems and components. *IEEE International Conference on Intelligence and Security Informatics*, pages 252–257, 2011.

D. Pietzner and A. Wienke. The trend-renewal process: a useful model for medical recurrence data. *Statistics in Medicine*, 32:142–152, 2013.

G. O. Roberts and J. S. Rosenthal. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science*, 16:351–367, 2001.

G. O. Roberts and J. S. Rosenthal. Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18:349–367, 2009.

D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistifcal Society: Series B*, 64:583–639, 2002.

H. Wang and H. Pham. A quasi renewal process and its applications in imperfect maintenance. *International Journal of Systems Science*, 27:1055–1062, 1996.

A. G. Wilson, T. L. Graves, M. S. Hamada, and C. S. Reese. Advances in data combination, analysis and collection for system reliability assessment. *Statistical Science*, 21:514–531, 2006.

M. Yańez, F. Joglar, and M. Modarres. Generalized renewal process for analysis of repairable systems with limited failure experience. *Reliability Engineering and System Safety*, 77: 167–180, 2002.

Q. Yang, Y. Hong, Y. Chen, and J. Shi. Failure profile analysis of complex repairable systems with multiple failure modes. *IEEE Transactions on Reliability*, 61:180–191, 2012.

Q. Yu, H. Guo, and H. Liao. An analytical approach to failure prediction for systems subject to general repairs. *IEEE Transactions on Reliability*, 62:714–721, 2013.

R. Zhao and B. Liu. Renewal process with fuzzy interarrival times and rewards. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 11:573–586, 2003.

Chapter 5 General Conclusions and Areas for Future Work

5.1 Conclusions

For highly reliable products, the number of failures is usually small, however, the prediction of the failures is still important to the manufacturers. In most of the literature, researchers just consider the limited time-to-failure events to fit the model and do the prediction. However, the field data in practice are often complicated and may contain some situations such as the retirement of products, leading to the prediction of the failures far away from the reality. In Chapter 2, we develop a statistical prediction procedure that considers the impact of the product retirements and reporting delays. Although we only consider such two related situations in the prediction, the method proposed can also apply to other practical situations.

Because highly reliable products usually have relative long lifetimes, it is difficult to collect enough failure events to do the prediction. In Chapter 3, we develop a nonlinear mixed-effects general path model to predict the lifetime of the coating product based on the degradation analysis. In this chapter, we develop a method to combine the information of dynamic covariates such as the ultraviolet dosage, temperature, and relative humidity, into the general path model. Considering the specific effect shapes of the dynamic covariates, the shape-restricted splines are used which also bring some challenges in the parameter estimation in such nonlinear mixed-effect model. To solve this problem, an efficient algorithm is developed.

In Chapter 5, we propose a multi-level trend-renewal process to model and predict the component replacement events with the consideration of the effect of subsystem replacement events. Besides the subsystem replacement effects, we also consider the time-dependent covariates as well as the unit-to-unit variability in the modeling and prediction of component replacement events. Because a closed-form expression for the prediction is not available, Monte Carlo simulation is used to compute the point predictions and prediction intervals for the cumulative number of component events. The proposed model includes renewal process,

nonhomogeneous Poisson process, and trend-renewal process as special cases.

5.2 Areas for Future Work

The model and prediction procedures in Chapter 2 are based on the assumption that the failure-time distribution and retirement-time distribution are independent because of the nature of the motivating data. In other cases, this independent relationship may not be true, and the dependence structure can be considered. We only use limited information about the retirement distribution (i.e., mean and variance) due to the limited information we have. For some applications, it would be possible and useful to track some representative samples through a carefully designed field tracking study.

In Chapter 3, we use a nonlinear mixed-effect general path model to do the degradation analysis. In the future, we can consider other stochastic models to incorporate the dynamic covariates information. Variable selection for dynamic covariates also can be considered as a future research topic.

In Chapter 4, we consider the modeling and prediction of one particular component in a two-level system. It will be more interesting and useful to study more components and subsystems with more levels, which, however, will be more challenging.