ICSA Book Series in Statistics
Series Editors: Jiahua Chen · Ding-Geng (Din) Chen

Ding-Geng (Din) Chen
Yuhlong Lio
Hon Keung Tony Ng
Tzong-Ru Tsai *Editors*

Statistical Modeling for Degradation Data

Springer

ICSA Book Series in Statistics

More information about this series at http://www.springer.com/series/13402

Ding-Geng (Din) Chen • Yuhlong Lio
Hon Keung Tony Ng • Tzong-Ru Tsai
Editors

Statistical Modeling for Degradation Data

Springer

Editors
Ding-Geng (Din) Chen
School of Social Work & Department of Biostatistics
University of North Carolina
Chapel Hill, NC, USA

Department of Statistics
University of Pretoria
Pretoria, South Africa

Hon Keung Tony Ng
Department of Statistical Science
Southern Methodist University
Dallas, TX, USA

Yuhlong Lio
Department of Mathematical Sciences
University of South Dakota
Vermillion, SD, USA

Tzong-Ru Tsai
Department of Statistics
Tamkang University
New Taipei City, Taiwan

ISSN 2199-0980 ISSN 2199-0999 (electronic)
ICSA Book Series in Statistics
ISBN 978-981-10-5193-7 ISBN 978-981-10-5194-4 (eBook)
DOI 10.1007/978-981-10-5194-4

Library of Congress Control Number: 2017949531

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

In many realities, the lifetime of a given subject could be difficult to measure or obtain because of longevity and/or failures caused by unexpected cumulative, random shocks. In such situations, if a cumulative degradation characteristic related to lifetime can be observed, then the reliability or lifetime percentile can be modeled through these degradation measurements. In recent years, degradation data analysis has played an important role in widely varied disciplines such as reliability, public health sciences, and finance. For instance, reliability information for a highly reliable product can be obtained through analyzing degradation data. Statistical modeling and inference techniques have been developed based on different degradation measures.

Our aim in creating this book was to provide a venue for the timely dissemination of research on the statistical aspects of the analysis of degradation data and to promote further research and collaborative work in this area. The book strives to bring together experts engaged in statistical modeling and inference to present and discuss the most recent important advances in degradation data analysis and related applications. The authors have made their data and/or computer programs publicly available, thus making it possible for readers to replicate the model development or data analysis presented in each chapter as well as to readily apply these new methods in their own research.

The 17 chapters are organized into three parts. Part I includes six chapters that present a review of the theoretical framework of degradation data. Part II comprises four chapters with a common focus on modeling and experimental design in degradation. Part III is composed of seven chapters that outline applications in degradation data analysis. All chapters underwent a rigorous peer-review process.

Part 1: Review and Theoretical Framework (Chaps. 1, 2, 3, 4, 5, and 6)

Chapter 1 presents a general methodology for stochastic degradation models based on a generalized cumulative damage approach. In this chapter, Park uses a generalized cumulative damage approach with a stochastic process to describe degradation and provides stochastic accelerated degradation models that handle failure data analysis with both soft and hard failures.

In Chap. 2, Bae and Yuan propose a hierarchical Bayesian change-point regression model to fit two-phase nonlinear degradation data. The authors have developed a Gibbs sampling algorithm for the inference of the parameters in the change-point regression model.

In Chap. 3, Hu, Li, and Hu provide a comprehensive review of the different kinds of modeling and analytic approaches for residual life prediction. The authors illustrate their discussion using a practical field application example and discuss over related issues in model selection and initial product performance incorporation.

Chapter 4 focuses on shock models with Poisson and generalized Poisson shock processes. Finkelstein and Cha derive the corresponding survival probabilities and demonstrate relevant stochastic analysis of the conditional characteristics. In addition, the authors consider some applications of these models.

In Chap. 5, Peng and Ye review degradation-based reliability modeling and assessment of complex systems in dynamic environments using a progressive three-stage framework: Stage 1, one-dimensional degradation process under dynamic environments; Stage 2, multiple degradation processes under static environments; and Stage 3, multiple degradation processes under dynamic environments. Peng and Ye also present additional research from both theoretical and practical perspectives.

Chapter 6 provides a general survey of the modeling and applications on non-destructive and destructive degradation tests. Authors Tsai, Lin, and Balakrishnan discuss the practical applications of burn-in and acceptance tests and provide other examples from the literature.

Part 2: Modeling and Experimental Designs (Chaps. 7, 8, 9, and 10)

Chapter 7 presents a degradation test plan for a nonlinear random coefficient model. In this chapter, Kim and Bae develop a design for a cost-efficient degradation test plan in the context of a nonlinear random coefficient model while satisfying precision constraints for the failure-time distribution. The proposed method is applied to degradation data of plasma display panels, and a sensitivity analysis is provided to show the robustness of the proposed test plan.

In Chap. 8, Tsai, Lio, Jiang, Ng, and Chen overview a statistical inference of the lifetime percentiles of light-emitting diodes (LED) based on lumen degradation

measurements collected by the constant-stress accelerated degradation testing method with the stress loadings of ambient temperature and drive current. The discussion focuses on the process of reaching a compromise between the experimental budget and estimation precision of reliability inference when implementing a constant-stress accelerated degradation test.

In Chap. 9, Balakrishnan, Tsai, and Lin discuss the inferences and optimal design of gamma degradation models. Additionally, they review the corresponding optimal designs for efficiently conducting degradation experiments on burn-in tests, which are used for classifying a unit as either *typical* or *weak* before products reach consumers.

Chapter 10 presents model misspecification analysis of inverse Gaussian and gamma degradation processes. In this chapter, Tseng and Yao first review results of model misspecification analysis. Then, motivated by stress relaxation data, the authors investigate the effect of misspecification on the prediction of a product's mean lifetime, when an increasing degradation model is not correctly specified.

Part 3: Applications (Chaps. 11, 12, 13, 14, 15, 16, and 17)

Chapter 11 addresses the practical application of Fréchet Shock-degradation models for system failures. Authors Lee and Whitmore present a wide variety of data structures for which this model is suitable and illustrate the estimation methods through three case applications.

In Chap. 12, Xie and Hong present the statistical methods for thermal index (TI) estimation based on accelerated destructive degradation test data. The authors provide a comprehensive review of the three methods and illustrate ways of estimating TI based on different models. The chapter includes comprehensive simulation studies that demonstrate the properties of different methods.

Chapter 13 addresses the remaining useful life estimation, a core concept in health management and research in the field of reliability. Ling, Ng, and Tsui present the inference on the remaining useful life under gamma degradation models with random effects. A Monte Carlo simulation study is carried out to evaluate the performance of the developed method using real data on the light intensity of LED lights.

Chapter 14 illustrates the R package "ADDT" for the analysis of accelerated destructive degradation test data. In this chapter, Xie and Hong provide a detailed description of the uses and functions of the R package ADDT for performing least squares, parametric, and semiparametric analyses of accelerated destructive degradation test data.

Chapter 15 centers on CD4 data, a critical laboratory indicator of the strength of the human immune system and a predictor of the progress of HIV. In this chapter, He, Liu, and Wang focus on modeling and inference of CD4 data and illustrate functional data analysis using a smoothing technique to extract the form of the underlying random function.

Chapter 16 discusses the remaining useful life prediction for rechargeable batteries. In this chapter, Wang and Tsui propose several state space models based on prognostic methods to predict battery remaining useful life. The authors then use these models to design a method for predicting battery life at different discharge rates.

Chapter 17 examines system identification for accelerated destructive degradation testing of nonlinear dynamic systems. In this chapter, Crous, Wilke, Kok, Chen, and Heyns develop three methods and conduct tests using simulated data. The authors demonstrate the superiority of a prototype bootstrapping strategy; using 400,000 data points generated by this strategy, the input signals were predicted with mean square errors of 5.08e-4.

As a general note, the references for each chapter immediately follow the chapter text. We have organized the chapters as self-contained units so readers can more easily and readily refer to the cited sources for each chapter. To facilitate readers' understanding of the methods presented in this book, the corresponding data and computing program can be requested from the chapter authors.

The editors are deeply grateful to the many who have supported the creation of this book. We first thank the authors of each chapter for their contributions and their generous sharing of their knowledge, time, and expertise. Second, our sincere gratitude goes to all the chapter reviewers for their expert reviews of each chapter of this book, which substantially improved the quality of this book. We gratefully acknowledge the professional support of Hannah Qiu (Springer/ICSA Book Series coordinator) and Wei Zhao (associate editor) from Springer Beijing that made publishing this book by Springer a reality.

We welcome readers' comments, including notes on typos or other errors, and look forward to receiving suggestions for improvements to future editions. Please send comments and suggestions to any of the editors listed below.

Chapel Hill, NC, USA — Ding-Geng (Din) Chen
Vermillion, SD, USA — Yuhlong Lio
Dallas, TX, USA — Hon Keung Tony Ng
New Taipei City, Taiwan — Tzong-Ru Tsai

List of Chapter Reviewers

Dr. Yili Hong, Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
Dr. Qingpei Hu, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
Dr. Man Ho Ling, Department of Mathematics and Information Technology, The Hong Kong Institute of Education, Hong Kong SAR, China
Dr. Chanseok Park, Department of Industrial Engineering, Pusan National University, South Korea
Dr. Ho Yiu So, Department of Mathematics and Statistics, McMaster University, Hamilton, Canada
Dr. Chih-Chun Tsai, Department of Mathematics, Tamkang University, Taiwan
Dr. Xiao Wang, Department of Statistics, Purdue University, West Lafayette, Indiana, USA
Dr. Zhisheng Ye, Department of Industrial & Systems Engineering, National University of Singapore, Singapore

Contents

Contributors

Suk Joo Bae Department of Industrial Engineering, Hanyang University, Seoul, South Korea

N. Balakrishnan Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada

Ji Hwan Cha Department of Statistics, Ewha Womans University, Seoul, Republic of Korea

Ding-Geng (Din) Chen School of Social Work & Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA

Department of Statistics, University of Pretoria, Pretoria, South Africa

Jacq Crous Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa

Maxim Finkelstein Department of Mathematical Statistics, University of the Free State, Bloemfontein, South Africa

Shuang He Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN, USA

Stephan Heyns Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa

Lanqing Hong Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, Singapore

Yili Hong Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Lirong Hu Center of Quality and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

Qingpei Hu Center of Quality and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

Nan Jiang Department of Mathematical Sciences, University of South Dakota, Vermillion, SD, USA

Zhongnan Jin Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Seong-Joon Kim Chosun University, Gwangju, Seoul, South Korea

Schalk Kok Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa

Mei-Ling T. Lee School of Public Health, University of Maryland, College Park, MD, USA

Lingjiang Li Search and Big Data Division, Bejing Jingdong Shangke Information Technology Co., LTD, Beijing, China

Chien-Tai Lin Department of Mathematics, Tamkang University, Tamsui, Taiwan

Man Ho Ling Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China

Yuhlong Lio Department of Mathematical Sciences, University of South Dakota, Vermillion, SD, USA

Chuanhai Liu Department of Statistics, Purdue University, West Lafayette, IN, USA

Hon Keung Tony Ng Department of Statistical Science, Southern Methodist University, Dallas, TX, USA

Chanseok Park Department of Industrial Engineering, Pusan National University, Busan, South Korea

Weiwen Peng School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, China

Chih-Chun Tsai Department of Mathematics, Tamkang University, Tamsui, Taiwan

Tzong-Ru Tsai Department of Statistics, Tamkang University, New Taipei City, Taiwan

Sheng-Tsaing Tseng Institute of Statistics, National Tsing Hua University, Taiwan, Republic of China

Kwok-Leung Tsui Department of Systems Engineering and Engineering Management, The City University of Hong Kong, Hong Kong, China

Jennifer H. Van Mullekom Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Dong Wang Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong, China

Xiao Wang Department of Statistics, Purdue University, West Lafayette, IN, USA

Daniel Nicolas Wilke Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa

G.A. Whitmore McGill University, Montreal, QC, Canada

Yimeng Xie Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Yu-Cheng Yao Institute of Statistics National Tsing Hua University, Taiwan, Republic of China

Zhisheng Ye Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, Singapore

Tao Yuan Department of Industrial and Systems Engineering, Ohio University, Athens, OH, USA

About the Editors

Professor Ding-Geng (Din) Chen is a fellow of the American Statistical Association and currently the Wallace H. Kuralt distinguished professor at the University of North Carolina at Chapel Hill, USA, and an extraordinary professor at the University of Pretoria, South Africa. He was a professor at the University of Rochester and the Karl E. Peace endowed eminent scholar chair in biostatistics at Georgia Southern University. He is also a senior consultant for biopharmaceuticals and government agencies with extensive expertise in clinical trial biostatistics and public health statistics. Professor Chen has written more than 150 refereed publications and coauthored/coedited 12 books on clinical trial methodology, meta-analysis, causal inference, and public health statistics.

Professor Yuhlong Lio is a professor in the Department of Mathematical Sciences at the University of South Dakota, Vermillion, SD, USA. He is an associate editor of professional journals, including the *Journal of Statistical Computation and Simulation* and the *Electronic Journal of Applied Statistical Analysis*. His research interests include reliability, quality control, censoring methodology, kernel smoothing estimate, and accelerated degradation data modeling. Dr. Lio has written more than 70 refereed publications.

Professor Hon Keung Tony Ng is currently a professor of statistical science with the Southern Methodist University, Dallas, TX, USA. He is an associate editor of *Communications in Statistics*, *Computational Statistics*, the *Journal of Statistical Computation and Simulation*, and *Statistics & Probability Letters*. His research interests include reliability, censoring methodology, ordered data analysis, nonparametric methods, and statistical inference. He has published more than 100 research papers in refereed journals. He is the coauthor of the book *Precedence-Type Tests and Applications* and coeditor of *Ordered Data Analysis, Modeling and Health Research Methods*. Professor Ng is a fellow of the American Statistical Association, an elected senior member of the IEEE, and an elected member of the International Statistical Institute.

Professor Tzong-Ru Tsai is a professor in the Department of Statistics at Tamkang University in New Taipei City, Taiwan. His main research interests include quality control and reliability analysis. In 2010 and 2011, he served as a senior consultant for the Electronic and Optoelectronic System Research Laboratories of the Industrial Technology Research Institute, focusing on accelerated degradation testing. Currently, he is a consultant for electronic companies with extensive expertise in statistical quality control and experimental design. He also serves as an executive editor and associate editor for academic journals. Dr. Tsai has written more than 70 refereed publications.

Part I
Review and Theoretical Framework

Chapter 1
Stochastic Accelerated Degradation Models Based on a Generalized Cumulative Damage Approach

Chanseok Park

Abstract A general methodology for stochastic degradation models is introduced that allows for both hard and soft failures to be taken into account when conducting parametric inference on lifetimes. Due to the development of engineering and science technology, modern products have longer lifetimes and greater reliability than ever before. Thus, it often takes more time to observe failures under normal-use conditions. Accelerated tests have been developed in order to deal with this lifetime-to-failure increase. Accelerated tests decrease the strength or lifetime to failure by exposing the specimens or products to harsh conditions. This exposure results in earlier breakdowns. Modelling these accelerated tests requires the use of stochastic degradation models with accelerating explanatory variables. By using a generalized cumulative damage approach with a stochastic process describing degradation, we develop stochastic accelerated degradation models which handle failure data consisting of both hard and soft failures.

Keywords Accelerated life test • Cumulative damage • Degradation process • First passage time • Stochastic process

1.1 Introduction

Many systems, products and materials age, wear or degrade over time before they experience breakdown or failure. Degradation measurements can provide additional valuable information that can be used to supplement the actual failure observations. Thus, both degradation measurements (soft failures) and actual failure observations (hard failures) should be considered in statistical inference procedures. By using a generalized cumulative damage approach with a stochastic process describing *degradation*, a general methodology that incorporates both hard and soft failures is developed.

C. Park (✉)
Department of Industrial Engineering, Pusan National University, 46241, Busan, South Korea
e-mail: cpark2@gmail.com; cp@pusan.ac.kr

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_1

Statistical inference about the reliability of a product or system is generally based on lifetime testing. With the development of engineering and science technology, many products have longer lifetimes and greater reliability than ever before. Thus, degradation measurements and actual failure observations under normal-use conditions take more time than they used to. Also, the increase in measurements and observations is costly. Since it is both costly and time-consuming to obtain hard or soft failure observations under normal-use conditions, the use of accelerated tests in order to obtain observations in a more timely fashion is becoming much more important in practice.

Accelerated life tests decrease the strength or lifetime to failure by exposing the products or specimens to higher levels of stress conditions than normal-use conditions. The exposure increase results in earlier breakdowns or failures. In the afore-mentioned situations, it is necessary to develop useful and appropriate statistical models for statistical inference on the lifetime of the products or specimens under study. Such statistical models should appropriately incorporate the accelerating variables and degradation measurements as well as any actual failures observed.

Traditional acceleration rules such as the power rule or Arrhenius reaction rate rule are widely used in engineering in order to incorporate an accelerating variable. However, these rules are based on the assumption that only a single accelerating variable exists. In the case of models with two accelerating variables, specific link functions have been developed such as the generalized Eyring relationship model [1]. Unfortunately, the restriction to two accelerating variables limits the use of these models in practice. In many situations, the use of more than two accelerating variables can be useful in fitting accelerated test models because this often leads to improve prediction accuracy of device reliability under normal-use conditions.

Park and Padgett [2, 3] also extended existing results by developing more general models based on the use of a hyper-cuboidal volume approach as an overall acceleration measure. This approach, which will be discussed in depth in Sect. 1.4, can incorporate several accelerating variables and, at the same time, allow for the use of both hard and soft failures in statistical inference procedures.

Several authors have studied degradation models and accelerated test models for statistical inference in reliability. For example, Nelson [4–6] references results on degradation models and accelerated tests. The books by Bagdonavicius and Nikulin [7] and Meeker and Escobar [8] provide a good overview of degradation models and accelerated tests. Degradation models based on Brownian (or also known as Wiener) processes and related processes have been proposed by several authors including Doksum and Normand [9], Lu [10], Whitmore [11], Whitmore and Schenkelberg [12], Whitmore et al. [13], Pettit and Young [14], and Padgett and Tomlinson [15]. Degradation models based on gamma processes have been investigated by Bagdonavicius and Nikulin [16], Lawless and Crowder [17], and Lin et al. [18].

Specifically with respect to stochastic accelerated degradation models, Park and Padgett [2, 3] constructed a general class of accelerated testing models which include other existing models as special cases. One notable special case are the Brownian-motion-based models proposed by Lu [10] Padgett and Tomlinson [15],

and Li et al. [19]. Other cases include the geometric Brownian motion process models and gamma motion process models considered by Bagdonavicius and Nikulin [16] and Lawless and Crowder [17]. The methodology in Park and Padgett [2, 3] is general in that it easily allows for various assumptions about the stochastic processes underlying the degradation process.

Regression-type methods for modelling general degradation path also have been studied by Meeker and Escobar [8], Lu and Meeker [20], Boulanger and Escobar [21], Hamada [22], and Meeker et al. [23] among others. Meeker and Escobar [8], Carey and Koenig [24], and Yanagisawa [25] illustrate the use of degradation models with accelerated testing in practical engineering applications. Recently, Bayesian methods are also incorporated into accelerated degradation problems by several authors including Wang et al. [26], Guan et al. [27], and Fan and Chen [28]. Recent works by Ye et al. [29] and Ye and Xie [30] also provide a good overview of stochastic degradation modelling and analysis.

1.2 Basic Properties for Stochastic Cumulative Damage Process

Suppose that the lifetime of a system or the strength of a material can be described by a stochastic model. The stochastic model is advocated by many authors including Doksum and Normand [9], Lu [10], Whitmore [11], Whitmore and Schenkelberg [12], Pettit and Young [14], Padgett and Tomlinson [15], Bagdonavicius and Nikulin [16], Lawless and Crowder [17], Bhattacharyya and Fries [31], Desmond [32], Padgett [33], Durham and Padgett [34], and Park and Padgett [2, 3, 35–38] among others.

An *additive* damage model for failure of a system or a material was originally developed by Desmond [32]. The basic idea of this model is as follows. As the tensile load on the specimen is increased, the amount of cumulative damage after $n+1$ increments of stress is

$$X_{n+1} = X_n + \Delta_n, \tag{1.1}$$

where Δ_j for $j = 0, 1, 2 \ldots$ are independent and identically distributed (iid) damages to the system at each increment. A *multiplicative* damage model also has been proposed by Padgett [33], where the cumulative damage is given by

$$X_{n+1} = X_n + \Delta_n X_n.$$

Alternatively, one can assume that the lifetime of a system or the strength of a material is under an increasing stress level or tensile load and the cumulative damage X_{n+1} after $(n+1)$ increments of stress can then be described by a stochastic cumulative damage model. Thus, it is appropriate to represent this damage by

$$X_{n+1} = X_n + \Delta_n\, h(X_n), \tag{1.2}$$

where Δ_j for $j = 0, 1, 2 \ldots$ are iid damages and $h(\cdot)$ is the *damage model function*. This cumulative damage model was originally proposed in Durham and Padgett [34]. Damage models above can be generalized to an even greater degree by using a *damage accumulation function* which was proposed by Park and Padgett [35]

$$c(X_{n+1}) = c(X_n) + \Delta_n \, h(X_n), \tag{1.3}$$

where $c(\cdot)$ is the damage accumulation function. A continuous version of the generalized cumulative damage model in (1.3) can be expressed in a differential form

$$d\, c(X_t) = h(X_t)\, d\, D_t, \tag{1.4}$$

which leads to the following stochastic integral equation

$$\int_0^t \frac{1}{h(X_t)} d\, c(X_t) = D_t - D_0. \tag{1.5}$$

For more details on the stochastic integral, the reader is referred to Jacod and Shiryaev [39]. Since we assume that Δ_j for $j = 0, 1, 2 \ldots$ are iid in (1.2) and (1.3), it is reasonable to assume that the stochastic process $\{D_t, t \geq 0\}$ should satisfy the basic properties of random processes:

(i) $D_0 = 0$.
(ii) $\{D_t, t \geq 0\}$ has stationary increments (time-homogeneity).
(iii) $\{D_t, t \geq 0\}$ has independent increments.

By selecting a damage accumulation function $c(\cdot)$ along with an accompanying damage model function $h(\cdot)$ with the stochastic process $\{D_t, t \geq 0\}$ above, it becomes straightforward to develop general models which include previously well-known models [2, 10, 14, 15] as special cases. Using this generalized approach, a degradation model based on a Brownian motion process is illustrated in [10, 14, 15] and degradation models based on geometric Brownian and Gamma processes are developed in [2]. We briefly summarize these models in Table 1.1.

A well-known difficulty with the Brownian motion process model [10, 14, 15] is that it is not strictly increasing and not strictly positive. Therefore, at some point in time, the degradation measurement can take on a negative value which clearly

Table 1.1 Developed models with damage accumulation and model functions with stochastic process

$c(\cdot)$	$h(\cdot)$	$\{D_t, t \geq 0\}$	Degradation model
$c(u) = u$	$h(u) = 1$	Brownian	Brownian motion process model [10, 14, 15]
$c(u) = \ln u$	$h(u) = 1$	Brownian	Geometric Brownian motion process model [2]
$c(u) = u$	$h(u) = 1$	Shifted gamma	Gamma motion process model [2]

lacks a physical interpretation. The geometric Brownian motion process is always positive but not strictly increasing. Thus, at some point in time before degradation is measured, the degradation value can become larger than the measurement itself. Although this can be interpreted as a healing process, it is often an unrealistic physical assumption in certain situations. On the other hand, the gamma motion process is strictly positive and strictly increasing and avoids the problems associated with Brownian motion and geometric Brownian motion. This is a simple example of how one can select the appropriate stochastic model by considering the physical phenomenon underlying the application.

1.3 The Distribution of the Failure Time and the Degradation

1.3.1 Degradation Model Based on Brownian Motion Process

First, we consider the degradation model based on the Brownian motion process. Suppose that we have the following cumulative damage function of the system at time t,

$$X_t - X_0 = D_t,$$

where $\{D_t, t \geq 0\}$ is a Brownian motion process with mean αt and variance $\beta^2 t$. We assume that the process started at $X_0 = x_0$. Then it is straightforward to show that X_t is normally distributed with mean $x_0 + \alpha t$ and variance $\beta^2 t$. The probability density function (pdf) of X_t is given by

$$f(x|x_0, t) = \frac{1}{\beta\sqrt{t}}\,\phi\Big(\frac{x - x_0 - \alpha t}{\beta\sqrt{t}}\Big), \tag{1.6}$$

where $\phi(\cdot)$ is the pdf of the standard normal distribution.

Next, suppose that we assume that the damage threshold level for failure is a known positive constant, $C > 0$. Then the failure time of the system is the first passage time to the threshold C which is given by $S = \inf\{t > 0 : X_t \geq C\}$. It is well known that under the Brownian motion process, the first passage time is given by the inverse Gaussian distribution [40] with pdf

$$g(s|x_0, C) = \frac{\sqrt{\lambda}}{\sqrt{2\pi s^3}} \exp\Big(-\frac{\lambda(s-\mu)^2}{2\mu^2 s}\Big), \tag{1.7}$$

where $s > 0$, $\mu = (C - x_0)/\alpha$ and $\lambda = (C - x_0)^2/\beta^2$. Then the pdf of the degradation measurements can be obtained as follows. As described in Fig. 1.1, any sample path over the interval $(0, t)$ with the initial value $X_0 = x_0$ and terminal value $X_t = x$

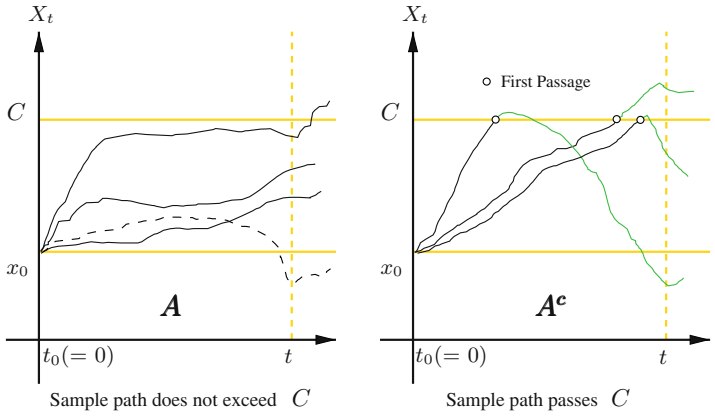

Fig. 1.1 Degradation path and the first passage under the Brownian motion process

either *never exceeds* the threshold at C (we denote this event by A), or *only passes once through* the threshold at C (we denote this by A^c).

It is immediate from the law of total probability that the pdf $f(x|x_0, t)$ in (1.6) can be written in terms of the events A and A^c as

$$f(x|x_0, t) = f_A(x|x_0, t) + f_{A^c}(x|x_0, t). \tag{1.8}$$

Note that the support of the sub-density $f_A(\cdot)$ is $(-\infty, C)$ and that of the sub-density $f_{A^c}(\cdot)$ is $(-\infty, \infty)$. We need to derive the pdf of the degradation level which is given by the conditional pdf of X_t given the event A

$$f(x|A, x_0, t) = \frac{f_A(x|x_0, t)}{P(A)}. \tag{1.9}$$

We first derive $f_{A^c}(x|x_0, t)$ since it is easier to derive $f_{A^c}(x|x_0, t)$ than $f_A(x|x_0, t)$. The sub-density $f_{A^c}(x|x_0, t)$ is obtained by deriving the marginal of the joint pdf of S and X_t whose pdf's are given in (1.6) and (1.7). Thus we have

$$\begin{aligned} f_{A^c}(x|x_0, t) &= \int_0^t g(s|x_0, C) f(x|C, t-s)\, ds \\ &= f(x|x_0, t) \cdot \exp\left\{ -\frac{2(C-x_0)(C-x)}{\beta^2 t} \right\}, \end{aligned} \tag{1.10}$$

where $-\infty < x < \infty$. Substituting (1.10) into (1.8) results in

$$\begin{aligned} f_A(x|x_0, t) &= f(x|x_0, t)\left[1 - \exp\left\{ -\frac{2(C-x_0)(C-x)}{\beta^2 t} \right\} \right] \\ &= \frac{1}{\beta\sqrt{t}} \phi\left(\frac{x - x_0 - \alpha t}{\beta\sqrt{t}} \right) \cdot \left[1 - \exp\left\{ -\frac{2(C-x_0)(C-x)}{\beta^2 t} \right\} \right], \end{aligned} \tag{1.11}$$

where $-\infty < x < C$.

Since $\int_{-\infty}^{C} f(x|A, x_0, t)dx = 1$, it is immediate from (1.9) that

$$P(A) = \int_{-\infty}^{C} f_A(x|x_0, t)dx.$$

This probability is obtained by Park and Padgett [3] and is given by

$$P(A) = \Phi\Big(\frac{C - x_0 - \alpha t}{\beta\sqrt{t}}\Big) - \Phi\Big(\frac{-C + x_0 - \alpha t}{\beta\sqrt{t}}\Big) \cdot \exp\Big\{\frac{2\alpha(C - x_0)}{\beta^2}\Big\}. \quad (1.12)$$

Note that if C is very large or t is very small, then $P(A) \approx 1$. Thus, in that case, we can ignore $P(A)$. Otherwise, $P(A)$ should be included in the likelihood function.

1.3.2 Degradation Model Based on Geometric Brownian Motion Process

Next, we consider the damage accumulation function $c(X_t) = \ln X_t$ and the damage model function $h(X_t) = 1$ with $\{D_t, t \geq 0\}$ being a Brownian motion process with mean αt and variance $\beta^2 t$. We assume that the process started at $X_0 = x_0$. It is immediate from (1.4) that we have the following cumulative damage function of the system at time t,

$$\ln X_t - \ln X_0 = D_t.$$

Thus, the system degradation X_t becomes a geometric Brownian motion process with pdf

$$f(x|x_0, t) = \frac{1}{\beta\sqrt{t}\,x}\,\phi\Big(\frac{\ln x - \ln x_0 - \alpha t}{\beta\sqrt{t}}\Big), \quad (1.13)$$

where $x > 0$. This degradation model was originally proposed by Park and Padgett [2]. The damage threshold level for failure is assumed to be a known positive constant C and the initial value of the process X_t is given by $X_0 = x_0$. The derivations for the geometric Brownian motion process are similar to those used in the case of the Brownian motion process. Next, we derive the first passage time to the threshold C which is given by $S = \inf\{t > 0 : X_t \geq C\}$. Since

$$\inf\{t > 0 : X_t \geq C\} = \inf\{t > 0 : \ln X_t \geq \ln C\}$$

and $\ln X_t$ is normally distributed with mean $\ln x_0 + \alpha t$ and variance $\beta^2 t$, it is immediate upon using the result in (1.7) that the pdf of the first passage time S is obtained as

$$g(s|x_0, C) = \frac{\sqrt{\lambda}}{\sqrt{2\pi s^3}} \exp\Big(-\frac{\lambda(s-\mu)^2}{2\mu^2 s}\Big), \tag{1.14}$$

where $s > 0$, $\mu = (\ln C - \ln x_0)/\alpha$ and $\lambda = (\ln C - \ln x_0)^2/\beta^2$.

Then, using the law of total probability, the pdf $f(x|x_0, t)$ in (1.13) can be written in terms of the events A and A^c as

$$f(x|x_0, t) = f_A(x|x_0, t) + f_{A^c}(x|x_0, t).$$

It should be noted that for this case, the support of the sub-density $f_A(\cdot)$ is $(0, C)$ and that of the sub-density $f_{A^c}(\cdot)$ is $(0, \infty)$. Using (1.11) along with the fact that $\ln X_t$ is normally distributed, we have

$$f_A(x|x_0, t) = \frac{1}{\beta\sqrt{t}\,x}\,\phi\Big(\frac{\ln x - \ln x_0 - \alpha t}{\beta\sqrt{t}}\Big)\Big[1 - \exp\Big\{-\frac{2(\ln C - \ln x_0)(\ln C - \ln x)}{\beta^2 t}\Big\}\Big],$$

where $0 < x < C$. Finally, we can obtain $P(A)$ using (1.12)

$$P(A) = \Phi\Big(\frac{\ln C - \ln x_0 - \alpha t}{\beta\sqrt{t}}\Big) - \Phi\Big(\frac{-\ln C + \ln x_0 - \alpha t}{\beta\sqrt{t}}\Big) \cdot \Big(\frac{C}{x_0}\Big)^{2\alpha/\beta^2}.$$

1.3.3 Degradation Model Based on Shifted Gamma Motion Process

As mentioned previously, the geometric Brownian motion process is always positive but not strictly increasing. Thus, at some point in time before the degradation value is measured, the value can become larger than the measurement itself. Although this can be interpreted as a healing process, it is often an unrealistic assumption in most applications. Generally speaking, except in a few special cases, it is desirable that a degradation process always be positive and strictly increasing.

Park and Padgett [2] considered a shifted gamma process as a degradation process X_t which is always positive and strictly increasing. Assuming that the damage model function $h(u) = 1$ and the damage accumulation function $c(u) = u$, we have the following cumulative damage of the system at the time t,

$$X_t - X_0 = D_t.$$

We assume that the stochastic process $\{D_t, t \geq 0\}$ is a gamma process with positive shape coefficient αt and scale β and that the process started at $X_0 = x_0$. This implies that the system degradation X_t is a shifted gamma process with pdf

$$f(x|x_0,t) = \frac{1}{\Gamma(\alpha t)\beta^{\alpha t}}(x-x_0)^{\alpha t-1}\exp\Big(-\frac{x-x_0}{\beta}\Big), \tag{1.15}$$

where $x_0 < x < \infty$.

Let S be the first passage time of the gamma process X_t to the threshold C. If the initial value of the process X_t is x_0 and the threshold value is a known positive constant C, then $X_t - x_0$ is distributed as gamma with shape parameter αt and scale β. Since X_t is strictly increasing in t, the probability of $S > t$ is the same as that of $x_0 < X_t < C$, that is,

$$P[S > t] = P[x_0 < X_t < C].$$

Using (1.15), we have

$$\begin{aligned} P[S > t] &= \int_{x_0}^{C} \frac{1}{\Gamma(\alpha t)\beta^{\alpha t}}(x-x_0)^{\alpha t-1}\exp\Big(-\frac{x-x_0}{\beta}\Big)\,dx \\ &= \int_0^{C-x_0} \frac{1}{\Gamma(\alpha t)\beta^{\alpha t}}x^{\alpha t-1}\exp\Big(-\frac{x}{\beta}\Big)\,dx \\ &= \frac{1}{\Gamma(\alpha t)}\int_0^{C_\beta}\xi^{\alpha t-1}e^{-\xi}\,d\xi, \end{aligned}$$

where $C_\beta = (C-x_0)/\beta$. Thus, we obtain the distribution of S

$$\begin{aligned} G(s|x_0,C) &= P[S \le s] \\ &= 1 - \frac{1}{\Gamma(\alpha s)}\int_0^{C_\beta}\xi^{\alpha s-1}e^{-\xi}\,d\xi \\ &= \frac{1}{\Gamma(\alpha s)}\int_{C_\beta}^{\infty}\xi^{\alpha s-1}e^{-\xi}\,d\xi \\ &= \frac{\Gamma(\alpha s, C_\beta)}{\Gamma(\alpha s)}, \end{aligned}$$

where $\Gamma(a,z)$ is the incomplete gamma function defined as

$$\Gamma(a,z) = \int_z^{\infty}\xi^{a-1}e^{-\xi}d\xi.$$

Given the distribution of S, the derivation of the pdf of S requires the use of some complex and tedious calculus and algebra which is not shown here. The resultant pdf of S is shown below

$$g(s|x_0, C) = \frac{d}{ds}\left[\frac{\Gamma(\alpha s, C_\beta)}{\Gamma(\alpha s)}\right]$$
$$= \alpha\Big\{\Psi(\alpha s) - \ln C_\beta\Big\}\Big\{1 - \frac{\Gamma(\alpha s, C_\beta)}{\Gamma(\alpha s)}\Big\} + \frac{\alpha}{\Gamma(\alpha s)}\frac{C_\beta^{\alpha s}}{(\alpha s)^2}\cdot {}_2F_2, \quad (1.16)$$

where

$${}_2F_2 = 1 + \sum_{k=1}^{\infty}\Big(\frac{\alpha s}{\alpha s + k}\Big)^2\frac{(-C_\beta)^k}{k!}.$$

For more details on these derivations, the reader is referred to Park and Padgett [2]. Although the pdf provided in (1.16) is the exact pdf of S, it is quite complex and extremely difficult to compute in practice. In order to avoid this computation, the approximate distribution of S is derived in [2, 38] and is found to be a type of Birnbaum-Saunders distribution [41]

$$G(s|x_0, C) \approx \Phi\left[\frac{1}{\alpha^*}\Big(\sqrt{\frac{s}{\beta^*}} - \sqrt{\frac{\beta^*}{s}}\Big)\right], \quad (1.17)$$

where $\alpha^* = 1/\sqrt{C_\beta}$, $\beta^* = C_\beta/\alpha$, and $C_\beta = (C - x_0)/\beta$. If $C_\beta/\sqrt{\alpha} \gg C_\beta/\alpha$ (i.e., $\sqrt{\alpha} \gg 1$), then the above distribution is known to be quite similar to the inverse Gaussian distribution with parameters $\mu = C_\beta/\alpha$ and $\lambda = C_\beta^2/\alpha$. For more details, see Chhikara and Folks [40].

Next, we find the sub-density of X_t with the event A. The event A is defined as the event that any sample path within the interval $(0, t)$ with the initial value $X_0 = x_0$ has a terminal value $X_t = x$ that does not exceed the threshold C. The probability of the event A becomes

$$P(A) = P[S > t] = P\Big[\max_{0<u\le t} X_u < C\Big].$$

Since the shifted gamma process X_u is strictly increasing in u, we have

$$\max_{0<u\le t} X_u = X_t.$$

Therefore, we have $P(A) = P[x_0 < X_t < C]$. Using the law of total probability, we have

$$f(x|x_0, t) = f_A(x|x_0, t) + f_{A^c}(x|x_0, t)$$
$$= f(x|A, x_0, t)P(A) + f(x|A^c, x_0, t)P(A^c)$$
$$= f(x|x_0, t)\mathbb{I}(x \le C) + f(x|x_0, t)\mathbb{I}(x > C),$$

where $\mathbb{I}(\cdot)$ is the indicator function. Notice that the support of $f_A(x|x_0,t)$ is (x_0, C) and that of $f_{A^c}(x|x_0,t)$ is (C,∞). This results in the sub-density of X_t

$$f_A(x|x_0,t) = \frac{1}{\Gamma(\alpha t)\beta^{\alpha t}}(x-x_0)^{\alpha t-1}\exp\Big(-\frac{x-x_0}{\beta}\Big), \tag{1.18}$$

where $x_0 < x < C$. Notice that the support of the sub-density in (1.15) is (x_0,∞) while the support of the pdf in (1.18) is (x_0, C). Hence, the pdf conditional on the event A is given by

$$f(x|A,x_0,t) = \frac{f_A(x|x_0,t)}{P(A)},$$

where $x_0 < x < C$ and $P(A) = \int_{x_0}^{C} f_A(x|x_0,t)dx$. It should also be noted that, similar to the previous cases, $P(A)$ can be ignored in the likelihood function when C is very large or t is very small.

1.3.4 General Likelihood for Hard and Soft Failures

We assume that we observe n (hard or soft) failures from independent units. Suppose that there are p hard failures denoted by s_k with $k = 1,2,\ldots,p$ and q soft failures, where $n = p + q$. The q soft failures are obtained with the q sets of degradation measurements

$$\{x_{l,0},\ x_{l,1},\ \ldots,\ x_{l,M_l}\},$$

where $l = 1,2,\ldots,q$ and the degradation measurements $x_{l,m}$ with $m = 1,2,\ldots,M_l$ are obtained at the corresponding times $t_{l,m}$ with $t_{0,m} = 0$ in general. In the reliability literature, the degradation measurements above are referred to as the degradation path data.

The likelihood function for p units which result in hard failures at s_k is given by

$$L_{\text{hard}}(\boldsymbol{\theta}) \propto \prod_{k=1}^{p} g(s_k|x_{k,0},C_k),$$

where $\boldsymbol{\theta}$ denotes the vector of parameters. In addition, the likelihood function for q units which result in q sets of degradation path measurements is given by

$$L_{\text{soft}}(\boldsymbol{\theta}) \propto \prod_{l=1}^{q}\prod_{m=1}^{M_l} f(x_{l,m}|A,x_{l,m-1},\Delta t_{l,m}),$$

where $\Delta t_{l,m} = t_{l,m} - t_{l,m-1}$. Since we assume that failures are obtained from n independent units, the general likelihood function for both hard and soft failures on n units is obtained as

$$L(\boldsymbol{\theta}) \propto L_{\text{hard}}(\boldsymbol{\theta}) \times L_{\text{soft}}(\boldsymbol{\theta}).$$

1.4 Degradation Models with Several Accelerating Variables

As explained previously, accelerated tests have been developed in order to obtain hard or soft failures in a more timely fashion. They decrease the strength or lifetime to failure by exposing the specimens or products to harsh conditions. Modelling these accelerated tests requires the use of stochastic degradation models along with accelerating explanatory variables. Traditional acceleration rules such as the power rule or Arrhenius reaction rate rule are widely used in engineering in order to incorporate an accelerating variable. However, traditional rules such as these are usually based on a single accelerating variable.

Let us assume that the parameter α in the Brownian motion, geometric Brownian motion and shifted gamma processes depends on several accelerating variables (say, $L_1, L_2, \ldots, L_p$) through an acceleration (link) function. For the case of $p = 1$ (i.e. a single accelerating variable), there are a variety of well-known acceleration functions relating the model coefficients and the single accelerating variable. These functions yield a large family of inverse-Gaussian-type models (see [42]) with an accelerating variable that can be incorporated with hard and soft failure observations. For more details regarding accelerated tests using link functions, the reader is referred to Mann et al. [43], Nelson [4], Meeker and Escobar [8], and other references therein. Well-known traditional acceleration functions with a single accelerating variable are summarized in Table 1.2.

In order to improve upon the models that handle one accelerating variable, other specific link functions have been developed for two accelerating variables. One of them is the generalized Eyring relationship rule [1]. However this is also of somewhat limited use in practice because of their two-variable restriction. To overcome the two-variable limitation, Park and Padgett [37] suggest a *hyper-cuboidal volume* approach as an overall accelerating measure which can incorporate several accelerating variables in a straightforward fashion. This idea is essentially an

Table 1.2 Acceleration functions with a single accelerating variable

Model	Acceleration function
Power rule	$\alpha_L = \xi L^{\eta}$
Arrhenius reaction rate rule	$\alpha_L = \xi e^{\eta/L}$
Inverse-log rule	$\alpha_L = \xi (\ln L)^{\eta}$
Exponential rule	$\alpha_L = \xi e^{\eta L}$
Inverse-linear rule	$\alpha_L = \xi + \eta L$

extension of the *weakest-link* concept in one dimension which is originally attributed to Peirce [44]. For example, suppose that we have a situation where failure depends on the size of the specimen under test and that there are two accelerating variables denoted as L and A, respectively. Then one can view the two-dimensional size of the specimen as $L \times A$. Thus, it is reasonable to use $V = L \times A$ as an accelerating *volume* or *cuboidal* measure.

Similarly, if there are p accelerating variables $L_1, L_2, \ldots, L_p$, the two-dimensional rule can be further generalized by using the hyper-cuboidal volume obtained as the product of the respective accelerating variables. Note that, in order to add more flexibility, these variables can be power-transformed by generalizing the one-dimensional power-law rule. Thus, with p accelerating variables, the overall accelerating measure which generalizes two-dimensional "specimen size" can be expressed as

$$V = \prod_{\nu=1}^{p} \{T(L_\nu)\}^{\eta_\nu},$$

where $T(\cdot)$ is any monotone function. Through the use of this overall accelerating measure, one can use the following generalized acceleration function

$$\alpha_V = \eta_0 \prod_{\nu=1}^{p} \{T(L_\nu)\}^{\eta_\nu}.$$

This acceleration function includes the first four acceleration functions in Table 1.2 as special cases. The inverse-linear rule is the one exception of an acceleration function in Table 1.2 that cannot be handled as a special case of the generalized acceleration function. In order to deal with this exception, rather than using traditional inverse-linear rule, one can use the generalized version below

$$\alpha_V = \eta_0 \prod_{\nu=1}^{p} (1 + \eta_\nu L_\nu).$$

It should also be noted that, aside from the inverse linear rule, the other four rules in Table 1.2 can be rewritten in the power rule form by simply transforming the accelerating variable. For example, the Arrhenius reaction rate model is obtained from the power rule $\alpha_V = \xi V^\eta$ with $V = e^{1/L}$. The exponential rule is obtained from the power rule with $V = e^L$. Similarly, the shifted inverse-log rule is obtained from the power rule with $V = 1 + \ln L$ while the original inverse-log rule is obtained with $V = \ln L$. However, when the original inverse-log rule is used with $L = 1$ (typical value for a normal-use condition), this results in infinite values of $\mu = (C - x_0)/\alpha$ in (1.7), $\mu = (\ln C - \ln x_0)/\alpha$ in (1.14) and $\beta^* = C_\beta/\alpha$ in (1.17). In order to avoid this pitfall, one can use a shifted version.

Note that the hyper-cuboidal volume approach can be incorporated into other existing models that only use one accelerating variable so that several accelerating variables can be conveniently handled. For example, Park and Padgett [3] apply the hyper-cuboidal volume approach to models based on the weakest-link theory. For recent work on weakest-link theory, the reader is referred to Wolstenholme [45]. The hyper-cuboidal volume approach of Park and Padgett [3] extends the weakest-link model so that it can handle several accelerating variables. The extension includes the power-law Weibull model [46, 47] as a special case. It is also noteworthy that Park and Padgett [3] show that a proportional hazard model [48] can be viewed as an extended weakest-link model with hyper-cuboidal volume.

1.5 Likelihood Construction with Accelerating Variables and Model Selection

In this section, we describe how to construct a general likelihood function. Park and Padgett [3] provide likelihood function constructions for the case of two accelerating variables with hard and soft failures. This can be easily extended to more complex cases with more accelerating variables. We briefly review their likelihood construction in what follows.

Consider two accelerating variables denoted as (L_1, L_2). Suppose that accelerated lifetime tests are performed under all accelerating levels of $(L_1^{(i)}, L_2^{(j)})$ for $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$. At each of the levels of $(L_1^{(i)}, L_2^{(j)})$, there are $N^{(i,j)}$ units under testing. The $P^{(i,j)}$ hard failures, denoted as $s_k^{(i,j)}$, $k = 1, 2, \ldots, P^{(i,j)}$, are observed with the threshold denoted as $C_k^{(i,j)}$. In addition to the $P^{(i,j)}$ hard failures, for each unit under testing at the accelerating levels of $(L_1^{(i)}, L_2^{(j)})$, the $Q^{(i,j)}$ soft failures are also observed up to the termination time, which result in degradation measurements $x_{\ell,m}^{(i,j)}$ at corresponding times $t_{\ell,m}^{(i,j)}$ for $m = 0, 1, 2, \ldots, M_\ell^{(i,j)}$ and $\ell = 1, 2, \ldots, Q^{(i,j)}$. Note that $P^{(i,j)} + Q^{(i,j)} = N^{(i,j)}$. The soft failures on the ℓth unit at the accelerating levels of $(L_1^{(i)}, L_2^{(j)})$ are obtained as

$$\left\{ x_{\ell,0}^{(i,j)},\ x_{\ell,1}^{(i,j)},\ \ldots,\ x_{\ell,M^*}^{(i,j)} \right\}$$

with $M^* = M_\ell^{(i,j)}$ for brevity. Note that if $x_{\ell,m}^{(i,j)}$ is being measured at $t_{\ell,m}^{(i,j)}$, then $t_{\ell,M^*}^{(i,j)}$ is the *termination time* of the unit being observed.

Given the framework described, it is easily shown that the log-likelihood function due to hard failures is given by

$$l_{\text{hard}}(\boldsymbol{\theta}) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{P^{(i,j)}} \ln g(s_k^{(i,j)} | x_{k,0}^{(i,j)}, C_k^{(i,j)}), \tag{1.19}$$

where $\boldsymbol{\theta}$ denotes the vector of parameters. Using a similar argument, the log-likelihood function for the degradation path measurements is given by

$$l_{\text{soft}}(\boldsymbol{\theta}) \propto \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\ell=1}^{Q^{(i,j)}}\sum_{m=1}^{M_{\ell}^{(i,j)}} \ln f(x_{\ell,m}^{(i,j)}|A, x_{\ell,m-1}^{(i,j)}, \Delta t_{\ell,m}^{(i,j)}), \tag{1.20}$$

where $\Delta t_{\ell,m}^{(i,j)} = t_{\ell,m}^{(i,j)} - t_{\ell,m-1}^{(i,j)}$. Clearly, given (1.19) and (1.20), the general log-likelihood function for hard and soft failures is given by

$$l(\boldsymbol{\theta}) \propto l_{\text{hard}}(\boldsymbol{\theta}) + l_{\text{soft}}(\boldsymbol{\theta}). \tag{1.21}$$

Given the log-likelihood function, the issue of model selection becomes important. A well-known model selection criteria, quite popular in the statistics and engineering literature is the Akaike information criterion (AIC) [49, 50]. The AIC is defined as:

$$\text{AIC} = -2l(\hat{\boldsymbol{\theta}}) + 2\nu,$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate under consideration and ν is the number of independent model parameters. Note that the second term in the AIC is viewed as the likelihood penalty contribution which arises due to the fact that an increase in the number of parameters, ν, will always result in an increase the likelihood.

As a measure of a good model fit, the model with a smaller AIC is preferred. For more details, the reader is referred to Burnham and Anderson [51]. Thus, if there are several competing models, the model with the smallest AIC among them is the selected model. Alternatively, it is shown in Park and Padgett [38] that one can use the mean square error (MSE) as the criterion for comparing several potential models. Although the MSE is an attractive criterion in that it is straightforward to compute, it is unfortunately limited to the case where there are only hard failure observations.

1.6 Concluding Remarks

In this article, we have provided a general methodology for developing stochastic accelerated degradation models with hard and soft failures. By selecting the appropriate damage accumulation function $c(\cdot)$ and the appropriate damage model function $h(\cdot)$, several useful stochastic degradation models were derived. The generality of the methodology allows one to choose these respective functions accordingly so that new stochastic accelerated degradation models can be developed and existing stochastic degradation models can be viewed as special cases.

Regarding future areas of research related to the stated studies, we believe that there are several ways that would be explored as stated below. In the particular case of the stochastic accelerated degradation model based on the shifted gamma motion process, the exact distribution of the first passage time has been derived. However, due to the difficulties with calculating the exact distribution numerically, an approximate distribution is recommended. This approximation has the form of a Birnbaum-Saunders distribution. In practical applications where a more accurate model is needed, one can use a better approximation through the use of more delicate numerical methods.

Several stochastic degradation models were developed by using Brownian, geometric Brownian and gamma motion processes. We could utilize other motion processes satisfying stationary increments and independent increments along with the generalized cumulative damage model. Currently, from a theoretical perspective, the use of the AIC criterion is the most appropriate methodology for model selection. One area of research that has not yet been explored is an investigation in the empirical performance of the AIC method as a means of model selection for stochastic accelerated degradation models. It may be the case that other less-known criterion may be useful in practice. By using Monte Carlo simulation techniques, it should be possible to investigate the usefulness of the AIC criterion as well as competing criterion such as BIC and others.

To the best of our knowledge, the development of stochastic accelerated degradation models by solving the stochastic integral equations in (1.5) has not been an area of emphasis in reliability research. Interesting and challenging future work would involve derivations of the solutions to these stochastic integral equations in order to develop other useful stochastic accelerated degradation models.

In the construction of the log-likelihood function in (1.21) for both hard and soft failures, we assumed that we observe failures from independent units and we separated failures into hard and soft failures. Thus, when a hard failure occurs, we should ignore its corresponding degradation measurements although we would obtain them along with a hard failure. Because of the independence assumption, we could not use this additional information. Thus, there is an information loss. To make the most use of this information, we need to consider the correlation between a hard failure and its corresponding degradation measurements to construct the more general likelihood function.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. NRF-2017R1A2B4004169). We appreciate the valuable comments from anonymous referees which led to an improvement of the article. The author also wishes to dedicate this work to the memory and honor of Professor Byung Ho Lee in the Department of Nuclear Engineering at Seoul National University.

References

1. Glasstone S, Laidler KJ, Eyring HE (1941) The theory of rate processes. McGraw-Hill, New York
2. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Anal 11:511–527
3. Park C, Padgett WJ (2006) Stochastic degradation models with several accelerating variables. IEEE Trans Reliab 55:379–390
4. Nelson W (1990) Accelerated testing: statistical models, test plans, data analyses. John Wiley, New York
5. Nelson W (2005) A bibliography of accelerated test plans. IEEE Trans Reliab 54:194–197
6. Nelson W (2005) A bibliography of accelerated test plans part II – references. IEEE Trans Reliab 54:370–373
7. Bagdonavicius V, Nikulin M (2002) Accelerated life models, modeling and statistical analysis. Chapman & Hall/CRC, Boca Raton
8. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. John Wiley, New York
9. Doksum K, Normand S-LT (1995) Gaussian models for degradation processes – part I: methods for the analysis of biomarker data. Lifetime Data Anal 1:135–144
10. Lu J (1995) Degradation processes and related reliability models. Ph.D. thesis, McGill University
11. Whitmore GA (1995) Estimating degradation by a Wiener diffusion process subject to measurement error. Lifetime Data Anal 1:307–319
12. Whitmore GA, Schenkelberg F (1997) Modelling accelerated degradation data using Wiener diffusion with a scale transformation. Lifetime Data Anal 3:27–45
13. Whitmore GA, Crowder MJ, Lawless JF (1998) Failure inference from a marker process based on a bivariate Wiener model. Lifetime Data Anal 4:229–251
14. Pettit LI, Young KDS (1999) Bayesian analysis for inverse Gaussian lifetime data with measures of degradation. J Stat Comput Simul 63:217–234
15. Padgett WJ, Tomlinson MA (2004) Inference from accelerated degradation and failure data based on Gaussian process models. Lifetime Data Anal 10:191–206
16. Bagdonavicius V, Nikulin M (2000) Estimation in degradation models with explanatory variables. Lifetime Data Anal 7:85–103
17. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
18. Ling MH, Tsui KL, Balakrishnan N (2015) Accelerated degradation analysis for the quality of a system based on the gamma process. IEEE Trans Reliab 64(1):463–472
19. Li J, Wang Z, Zhang Y, Fu H, Liu C, Krishnaswamy S (2017) Degradation data analysis based on a generalized Wiener process subject to measurement error. Mech Syst Signal Process 94:57–72
20. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
21. Boulanger M, Escobar LA (1994) Experimental design for a class of accelerated degradation tests. Technometrics 36:260–272
22. Hamada M (1995) Analysis of experiments for reliability improvement and robust reliability. In: Balakrishnan N (ed) Recent advances in life testing and reliability. CRC Press, Boca Raton
23. Meeker WQ, Escobar LA, Lu CJ (1998) Accelerated degradation tests: modeling and analysis. Technometrics 40:89–99
24. Carey MB, Koenig RH (1991) Reliability assessment based on accelerated degradation: a case study. IEEE Trans Reliab 40:499–506
25. Yanagisawa T (1997) Estimation of the degradation of amorphous silicon cells. Microelectron Reliab 37:549–554
26. Wang L, Pan R, Li X, Jiang T (2013) A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliab Eng Syst Saf 112:38–47

27. Guan Q, Tang Y, Xu A (2016) Objective Bayesian analysis accelerated degradation test based on Wiener process models. Appl Math Model 40(4):2743–2755
28. Fan T-H, Chen C-H (2017) A Bayesian predictive analysis of step-stress accelerated tests in gamma degradation-based processes. Qual Reliab Eng Int. doi:10.1002/qre.2114
29. Ye Z-S, Chen N, Shen Y (2015) A new class of Wiener process models for degradation analysis. Reliab Eng Syst Saf 139:58–67
30. Ye Z-S, Xie M (2015) Stochastic modelling and analysis of degradation for highly reliable products. Appl Stoch Model Bus Ind 31(1):16–32
31. Bhattacharyya GK, Fries A (1982) Fatigue failure models – Birnbaum-Saunders vs. inverse Gaussian. IEEE Trans Reliab 31:439–440
32. Desmond AF (1985) Stochastic models of failure in random environments. Can J Stat 13: 171–183
33. Padgett WJ (1998) A multiplicative damage model for strength of fibrous composite materials. IEEE Trans Reliab 47:46–52
34. Durham SD, Padgett WJ (1997) A cumulative damage model for system failure with application to carbon fibers and composites. Technometrics 39:34–44
35. Park C, Padgett WJ (2005) New cumulative damage models for failure using stochastic processes as initial damage. IEEE Trans Reliab 54:530–540
36. Park C, Padgett WJ (2006) A general class of cumulative damage models for materials failure. J Stat Plan Inference 136:3783–3801
37. Park C, Padgett WJ (2007) Cumulative damage models for failure with several accelerating variables. Qual Technol Quant Manag 4:17–34
38. Park C, Padgett WJ (2008) Cumulative damage models based on gamma processes. In: Ruggeri F, Faltin F, Kenett R (eds) Encyclopedia of statistics in quality and reliability. Wiley, Chichester
39. Jacod J, Shiryaev AN (1987) Limit theorems for stochastic processes. Springer, New York
40. Chhikara RS, Folks JL (1989) The inverse Gaussian distribution: theory, methodology, and applications. Marcel Dekker, New York
41. Birnbaum ZW, Saunders SC (1969) A new family of life distributions. J Appl Probab 6: 319–327
42. Onar A, Padgett WJ (2000) Inverse Gaussian accelerated test models based on cumulative damage. J Stat Comput Simul 66:233–247
43. Mann NR, Schafer RE, Singpurwalla ND (1974) Methods for statistical analysis of reliability and life data. Wiley, New York
44. Peirce FT (1926) Tensile tests for cotton yarns: “the weakest link” theorems on the strength of long and of composite specimens. J Text Inst 17:355–368
45. Wolstenholme LC (1995) A nonparametric test of the weakest-link principle. Technometrics 37:169–175
46. Padgett WJ, Durham SD, Mason AM (1995) Weibull analysis of the strength of carbon fibers using linear and power law models for the length effect. J Compos Mater 29:1873–1884
47. Smith RL (1991) Weibull regression models for reliability data. Reliab Eng Syst Saf 34:55–77
48. Cox DR (1972) Regression models and life tables. J R Stat Soc B 34:187–220
49. Akaike H (1993) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Czáki F (eds) Second international symposium on information theory 1973, Budapest. Akademiai Kiadó, pp 267–281. Reprinted in Kotz S, Johnson NL (eds) Breakthroughs in statistics, vol 1. Springer, pp 610–624
50. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–722
51. Burnham KP, Anderson DR (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer, New York

Chapter 2
Hierarchical Bayesian Change-Point Analysis for Nonlinear Degradation Data

Suk Joo Bae and Tao Yuan

Abstract Degradation data for some products tends to present two-phase patterns during testing periods. It is caused by defects or contaminants remaining after manufacturing processes. The change-point of the two-phase degradation path represents the time when a burn-in related phase transits to an inherent degradation phase with a slower and more stable rate. This chapter discusses a hierarchical Bayesian change-point regression model to fit the two-phase degradation patterns. A Gibbs sampling algorithm is developed for the inference of the parameters in the change-point regression model. The results indicate that reliability estimation can be improved substantially by using the change-point model to account for product burn-in. Based on the hierarchical Bayesian change-point degradation model, degradation-based burn-in tests are devised according to a reliability criterion or a cost criterion.

Keywords Change-point regression • Degradation • Gibbs sampling • Hierarchical Bayesian modeling • Lifetime distribution

2.1 Introduction

Degradation analysis is an attractive alternative to the traditional failure-time analysis for highly reliable products in that it can improve reliability inference and provide additional information related to failure mechanisms [14]. In real applications of the degradation test and analysis, it has been observed that the degradation paths of some products exhibit two-phase patterns over the test periods. For example, Gebraeel et al. [11] observed that the vibration-based signals for

S.J. Bae (✉)
Department of Industrial Engineering, Hanyang University, 222 Wangsimni-ro Seongdong-gu, Seoul, South Korea
e-mail: sjbae@hanyang.ac.kr

T. Yuan
Department of Industrial and Systems Engineering, Ohio University, 279 Stocker Center, 45701, Athens, OH, USA
e-mail: yuan@ohio.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_2

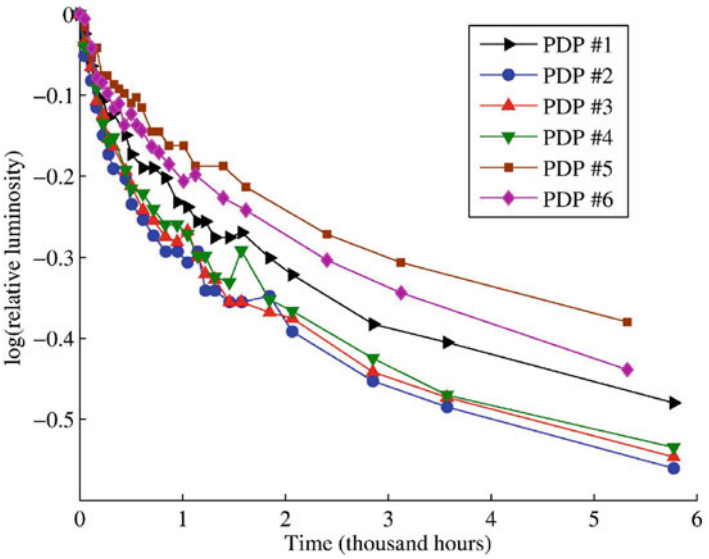

Fig. 2.1 Observed degradation paths of six PDPs: relative luminosity vs. measurement time

bearings showed a two-phase pattern, which was caused by the generation of fatigue defects. Before the fatigue defects occur, the degradation signals during the first phase are constant and stable. After the fatigue defects occur, the degradation path enters the second, unstable phase. Another example is the two-phase luminosity degradation of plasma display panels (PDPs) reported by Bae and Kvam [3] as shown in Fig. 2.1. After a rapid decrease in PDP luminosity at the first phase, the decrease in luminosity slowed and stabilized in the second phase. The two-phase degradation pattern of PDPs was due to nano-sized impurities unnecessarily produced during manufacturing processes. PDP manufacturers execute a burn-in procedure (called "aging" in the industry) to burn off the impurities. However, with an incomplete burn-in procedure, some impurities may remain in the products and cause the two-phase degradation pattern. Some other examples of products exhibiting two-phase degradation paths include direct methanol fuel cells [2], lithium-ion batteries [18], and organic light-emitting diodes (OLEDs) [17].

There have been attempts to model the two-phase degradation patterns. Bae et al. [1, 4] used, respectively, a bi-exponential model and a change-point regression model to fit the observed two-phase degradation paths of PDPs. The change-point regression model was also employed by Chen and Tsui [8] to analyze the two-phase degradation signals of bearings reported by Gebraeel et al. [11]. In this chapter, we will discuss the hierarchical Bayesian change-point degradation model developed by Bae et al. [4]. The PDP degradation data shown in Fig. 2.1 will be used as an illustrative example.

Based on the hierarchical Bayesian change-point regression model, this chapter will also discuss a degradation-based burn-in optimization problem studied by Yuan et al. [29]. Burn-in is a screening test to identify weak or defective products before shipping them to customers [13]. The burn-in test is usually conducted by stressing the products for a pre-determined duration under designed or accelerated stress conditions [9]. In traditional failure-based burn-in tests, products are subject to life tests and weak or defective products are identified by observing their failures during the tests. For highly reliable products, the failure-based burn-in tests may be ineffective because long burn-in durations may be needed in order to observe failures [28]. The degradation-based burn-in tests, which subject the products to degradation tests, are being considered as a promising alternative to the failure-based burn-in tests [24].

Tseng and Tang [23] proposed a cost-optimal burn-in policy using a Wiener process to describe the degradation pattern of the burn-in population. The burn-in decision variables were a burn-in duration and a cutoff point. At the end of the burn-in test, if a unit's degradation level exceeded the cutoff point, it was classified as a weak item. There have been extensions of this study to various directions [21, 22, 24]. Some studies considered simultaneous optimization of burn-in and preventive maintenance [10, 19, 26, 27]. In all those studies, it is assumed that there exists some proportion of weak products in the population, and the major purpose of the degradation-based burn-in tests is to weed out the weak products. Yuan et al. [29] recently considered a different type of degradation burn-in test for display products exhibiting two-phase degradation patterns. The major purpose of this burn-in ("aging") test is to eliminate the initial rapid degradation phase before shipping the display products to customers, while early failures of weak products (i.e., infant mortality) is not a major concern in terms of luminosity degradation. This chapter will discuss the Bayesian burn-in planning problem based on the hierarchical Bayesian change-point degradation model.

The reminder of this chapter is organized as follows. Section 2.2 presents the hierarchical Bayesian change-point degradation model developed to describe two-phase degradation paths. Section 2.3 discusses the degradation-based burn-in planning method based on the degradation model presented in Sect. 2.2. Section 2.4 discusses results of the proposed methodologies applied to the PDP degradation data shown in Fig. 2.1. Finally, Sect. 2.5 concludes this chapter.

2.2 Degradation Analysis Using Change-Point Regression

This section first reviews different forms of change-point regression models, then presents the hierarchical Bayesian change-point degradation model for two-phase degradation patterns, and finally derives the failure-time distribution of a unit randomly selected from the population. The PDP degradation is used as an example to illustrate the methods. The response variable y in the PDP example is the logarithm-transformed relative luminosity.

2.2.1 Change-Point Regression

This section reviews the change-point regression models that involve one change-point. A general change-point regression model with one change-point has the form of

$$y_j = \begin{cases} \alpha_1 + \beta_1 t_j + \epsilon_j, & j = 1, \ldots, \tau, \\ \alpha_2 + \beta_2 t_j + \epsilon_j, & j = \tau + 1, \ldots, n, \end{cases} \tag{2.1}$$

where y and t are the response variable and predictor, respectively, n is the number of observations, and the change occurs in the interval $[t_\tau, t_{\tau+1})$. The random error terms ϵ_j are generally assumed to be independent and identically distributed normal random variables with mean zero and variance σ^2. The mean function of the change-point regression model (2.1) includes two linear functions, where α_1 and β_1 are, respectively, the intercept and slope of the linear function before the change-point, and α_2 and β_2 are the intercept and slope of the linear function after the change-point, respectively. The mean regression function may be continuous or noncontinuous at the change-point. Let γ be the intersection point of the two linear lines, i.e., $\gamma = (\alpha_1 - \alpha_2)/(\beta_2 - \beta_1)$. A continuity constraint may be specified as $\gamma \in [t_\tau, t_{\tau+1})$. For the continuous model, γ is actually the change-point separating the two linear functions. Beem [6] developed a software package *Segcurve* for fitting the change-point regression model with or without the continuity constraint. Carlin et al. [7] considered the change-point regression model without the continuity constraint within the Bayesian framework and used the Gibbs sampling algorithm to inference the model parameters $(\alpha_1, \beta_1, \alpha_2, \beta_2, \tau, \sigma^2)$. Bae and Kvam [3] applied the change-point regression model (2.1) with the continuity constraint to fit the two-phase degradation paths of PDPs, while Chen and Tsui [8] did not impose the continuity constraint when studying the two-phase degradation signals of bearings.

An alternative form of the change-point regression model that explicitly includes the continuity constraint is given by [25]

$$y_j = \begin{cases} \alpha + \beta_1 (t_j - \gamma) + \epsilon_j, & t_j \le \gamma, \\ \alpha + \beta_2 (t_j - \gamma) + \epsilon_j, & t_j > \gamma, \end{cases} \tag{2.2}$$

where α is the expected response at the change-point γ. Muggeo [15] considered another change-point regression model given by

$$y_j = \begin{cases} \alpha\, t_j + \epsilon_j, & t_j \le \gamma, \\ \alpha\, t_j + \beta (t_j - \gamma) + \epsilon_j, & t_j > \gamma, \end{cases} \tag{2.3}$$

where α and $(\alpha + \beta)$ are, respectively, the slope of the linear function before the change-point γ and the slope of the linear function after the change-point. Model (2.3) also imposes the continuity constraint and assumes the expected response at time $t = 0$ is zero. We re-parameterize Model (2.3) as

$$y_j = g(t_j;\boldsymbol{\theta}) + \epsilon_j = \begin{cases} \alpha\, t_j - \beta t_j + \epsilon_j, & t_j \le \gamma, \\ \alpha\, t_j - \beta\gamma + \epsilon_j, & t_j > \gamma, \end{cases} \tag{2.4}$$

where $g(t_j;\boldsymbol{\theta})$ denotes the expected actual degradation path with a set of parameters $\boldsymbol{\theta} \equiv (\alpha, \beta, \gamma)$. $(\alpha - \beta)$ and α are, respectively, the slopes of the linear function before and after the change-point, respectively. Specifying $\alpha < 0$ and $\beta > 0$ yields the desired degradation pattern of the PDPs.

The change-point regression model (2.4) can be re-written as a standard regression form

$$\mathbf{y} = \mathbf{X}(\gamma)\boldsymbol{\vartheta} + \boldsymbol{\epsilon}, \tag{2.5}$$

where $\boldsymbol{\vartheta} \equiv (\alpha, \beta)$, and $\mathbf{y}$, $\mathbf{X}(\gamma)$, and $\boldsymbol{\epsilon}$ are defined, respectively, as

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X}(\gamma) = \begin{bmatrix} t_1 & -t_1 \\ \vdots & \vdots \\ t_\tau & -t_\tau \\ t_{\tau+1} & -\gamma \\ \vdots & \vdots \\ t_n & -\gamma \end{bmatrix}, \quad \text{and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}. \tag{2.6}$$

MLEs of the model parameters $(\alpha, \beta, \gamma, \sigma^2)$ can be obtained by maximizing the likelihood function given by

$$L(\boldsymbol{\theta}, \sigma^2) \propto (\sigma^2)^{-n/2} \exp\left[-\frac{(\mathbf{y} - \mathbf{X}(,)\boldsymbol{\vartheta})'(\mathbf{y} - \mathbf{X}(,)\boldsymbol{\vartheta})}{\sigma^2}\right]. \tag{2.7}$$

When the Bayesian method is used, prior distributions are needed for $\boldsymbol{\theta}$ and σ^2. This study considers the case in which there is a change-point in the interval $(0, \phi)$, and specifies a truncated trivariate normal prior for $\boldsymbol{\theta}$, i.e., $\boldsymbol{\theta} \equiv (\alpha, \beta, \gamma) \sim N_3(\boldsymbol{\mu},\ \boldsymbol{\Sigma})_{I_{\{\alpha<0,\beta>0,0<\gamma<\phi\}}}$, where $N_3(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ denotes the trivariate normal distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. The parameter σ^2 is assumed to have an inverse-gamma (IG) prior distribution with parameters a_σ and b_σ, i.e., $\sigma^2 \sim IG(a_\sigma, b_\sigma)$, which is the conjugate prior for σ^2, and this prior distribution for σ^2 is assumed to be independent of the prior for $\boldsymbol{\theta}$.

Given the observed data $\mathbf{d}$, the joint posterior distribution of the model parameters can be derived according to Bayes' formula

$$\begin{aligned} f(\boldsymbol{\theta}, \sigma^2|\mathbf{d}) &\propto f(\mathbf{d}|\boldsymbol{\theta}, \sigma^2) f(\boldsymbol{\theta}) f(\sigma^2) \\ &\propto \sigma^{-n} \exp\left[-\frac{(\mathbf{y} - \mathbf{X}(\gamma)\boldsymbol{\vartheta})'\,(\mathbf{y} - \mathbf{X}(\gamma)\boldsymbol{\vartheta})}{2\sigma^2}\right] \end{aligned}$$

$$\times \exp\left[-\frac{(\boldsymbol{\theta}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta}-\boldsymbol{\mu})}{2}\right]$$
$$\times \left(\sigma^2\right)^{-a_\sigma - 1} \exp\left(-\frac{b_\sigma}{\sigma^2}\right) I_{\{\alpha<0,\beta>0,0<\gamma<\phi\}}, \tag{2.8}$$

where $f(\cdot)$ represents a probability density function herein and hereafter, $f(\boldsymbol{\theta}, \sigma^2|\mathbf{d})$ is the joint posterior distribution, $f(\mathbf{d}|\boldsymbol{\theta}, \sigma^2)$ is the likelihood, and $f(\boldsymbol{\theta})$ and $f(\sigma^2)$ are the prior distributions for $\boldsymbol{\theta}$ and σ^2, respectively. Gibbs sampling technique is employed to draw random samples from the marginal posterior distribution of each parameter [12]. Marginal posterior distributions of functions of the model parameters, such as the expected degradation path $g(t; \boldsymbol{\theta})$ and the failure-time distribution, can also be conveniently sampled. Point or interval estimation of the model parameters and their functions can be computed using the sample statistics.

2.2.2 *Hierarchical Bayesian Change-Point Degradation Model*

This section introduces the hierarchical Bayesian change-point regression model for the degradation paths of multiple units. In the hierarchical Bayesian model, the *j*th response measured on the *i*th unit is modeled by

$$y_{i,j} = g(t_{i,j}; \boldsymbol{\theta}_j) + \epsilon_{i,j} = \begin{cases} \alpha_j\, t_{i,j} - \beta_j\, t_{i,j} + \epsilon_{i,j}, & t_{i,j} \le \gamma_j, \\ \alpha_j\, t_{i,j} - \beta_j\, \gamma_j + \epsilon_{i,j}, & t_{i,j} > \gamma_j, \end{cases} \tag{2.9}$$

for $j = 1, 2, \ldots, N$, where $y_{i,j}$ is the *j*th response on the *i*th unit, measured at time $t_{i,j}$, n_i is the number of measurements on the *i*th unit, N is the number of test units. The errors $\epsilon_{i,j}$ are assumed to be independent and identically distributed $N(0, \sigma^2)$ random variables.

A three-stage hierarchical Bayesian degradation model for multiple units is constructed as follows.The first stage models the observed two-phase degradation path of each test unit by the change-point regression model (2.9), which can be written equivalently as

$$\mathbf{y}_j = \mathbf{X}_j(\gamma_j)\boldsymbol{\vartheta}_j + \boldsymbol{\epsilon}_j, \tag{2.10}$$

for $j = 1, 2, \ldots, N$, where $\boldsymbol{\vartheta}_j \equiv (\alpha_j, \beta_j)$, and $\mathbf{y}_j$, $\mathbf{X}_j(\gamma_j)$, and $\boldsymbol{\epsilon}_j$ are defined similarly as in (2.6).

The second-stage model describes the unit-to-unit variability by assuming that all coefficient vectors $\boldsymbol{\theta}_j \equiv (\alpha_j, \beta_j, \gamma_j)$ form a random sample from a multivariate distribution whose probability density function is denoted by $f(\boldsymbol{\theta}|\boldsymbol{\phi})$ with a parameter vector $\boldsymbol{\phi}$. Specially, we assume this multivariate distribution is a truncated trivariate normal distribution, that is

$$\boldsymbol{\theta}_j \sim N_3(\boldsymbol{\mu},\ \boldsymbol{\Sigma})_{I_{\{\alpha_j<0,\beta_j>0,0<\gamma_j<\phi\}}}, \qquad j = 1, 2, \ldots, N. \tag{2.11}$$

Herein $\boldsymbol{\phi} \equiv (\boldsymbol{\mu},\ \boldsymbol{\Sigma})$.

In the third stage of the hierarchical model, prior distributions are specified for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and σ^2. For σ^2, the conjugate prior $IG(a_\sigma, b_\sigma)$ is employed. For the mean vector $\boldsymbol{\mu}$, a multivariate normal prior distribution with a mean vector $\boldsymbol{\xi}$ and a covariance matrix $\mathbf{C}$ is used, i.e., $\boldsymbol{\mu} \sim N(\boldsymbol{\xi},\ \mathbf{C})$. The multivariate normal distribution is the conditional conjugate prior for the multivariate normal mean vector. For the covariance matrix $\boldsymbol{\Sigma}$, the scaled inverse-Wishart prior discussed in [5] is adopted. The covariance matrix $\boldsymbol{\Sigma}$ is decomposed into variance and correlation components as $\boldsymbol{\Sigma} = \boldsymbol{\Delta}\mathbf{Q}\boldsymbol{\Delta}$, where the diagonal matrix $\boldsymbol{\Delta} = \text{diag}(\delta_1, \ldots, \delta_\nu)$ with $\delta_k > 0$, and $\mathbf{Q}$ is a $\nu \times \nu$ symmetric positive definite matrix. Herein $\nu = 3$ is the dimension of the vector $\boldsymbol{\theta}$. Then independent Gamma priors, $G(a_k, b_k)$, are assumed for δ_k, for $k = 1, 2, \ldots, \nu$, and an inverse-Wishart prior, $IW(\rho, \mathbf{S})$, is adopted for $\mathbf{Q}$, where ρ is the degrees of freedom, and $\mathbf{S}$ is a $\nu \times \nu$ symmetric positive-definite scale matrix.

Applying Bayes' formula, we can derive the joint posterior distribution of all model parameters, including the first-stage parameters $\theta_1, \theta_2, \ldots, \theta_N$ and the second-stage parameters $\boldsymbol{\mu}, \mathbf{Q}, \delta_1, \ldots, \delta_\nu, \sigma^2$, conditioning on the data $\mathbf{d}$ as follows

$$\begin{aligned}
&f(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N, \boldsymbol{\mu}, \mathbf{Q}, \delta_1, \ldots, \delta_\nu, \sigma^2 | \mathbf{d}) \\
&\quad \propto \prod_{i=1}^{N} \frac{1}{\sigma^{n_i}} \exp\left[-\frac{(\mathbf{y}_i - \mathbf{X}_i(\gamma_i)\boldsymbol{\vartheta}_i)'(\mathbf{y}_i - \mathbf{X}_i(\gamma_i)\boldsymbol{\vartheta}_i)}{2\sigma^2} \right] I_{\{\alpha_i<0,\beta_i>0,0<\gamma_i<\phi\}} \\
&\qquad \times [P(\boldsymbol{\mu}, \boldsymbol{\Sigma})]^{-N} |\boldsymbol{\Delta}|^{-N/2} |\mathbf{Q}|^{-N/2} |\boldsymbol{\Delta}|^{-N/2} \exp\left[-\frac{\sum_{i=1}^{N} (\boldsymbol{\theta}_i - \boldsymbol{\mu})' \boldsymbol{\Delta}^{-1} \mathbf{Q}^{-1} \boldsymbol{\Delta}^{-1} (\boldsymbol{\theta}_i - \boldsymbol{\mu})}{2} \right] \\
&\qquad \times \exp\left[-\frac{(\boldsymbol{\mu} - \boldsymbol{\xi})' \mathbf{C}^{-1} (\boldsymbol{\mu} - \boldsymbol{\xi})}{2} \right] \times |\mathbf{Q}|^{-(\nu+\rho+1)/2} \exp\left[-\frac{\text{tr}(\mathbf{S}\mathbf{Q}^{-1})}{2} \right] \\
&\qquad \times \prod_{k=1}^{\nu} \delta_k^{a_k - 1} \exp(-b_k \delta_k) \times (\sigma^2)^{-(a_\sigma + 1)} \exp\left(-\frac{b_\sigma}{\sigma^2} \right),
\end{aligned} \tag{2.12}$$

where $[P(\boldsymbol{\mu}, \boldsymbol{\Sigma})]$ is a normalizing constant that is attributable to the truncation in the second-stage model (2.11). A Gibbs sampling algorithm is developed to fit the three-stage hierarchical Bayesian change-point regression model and draw random samples from the marginal posterior distributions such as $f(\sigma^2|\mathbf{d})$, $f(\boldsymbol{\phi}|\mathbf{d})$, $f(\boldsymbol{\mu}|\mathbf{d})$, $f(\boldsymbol{\Sigma}|\mathbf{d})$, and $f(\boldsymbol{\theta}_j|\mathbf{d})$, for $j = 1, 2, \ldots, N$.

2.2.3 *Deriving the Failure-Time Distribution*

This section derives the failure-time distribution of a unit randomly selected from its population. The failure time T of a random unit is defined as the first hitting time at which the expected actual degradation path $g(t; \boldsymbol{\theta})$ reaches a pre-determined degradation level y^*. For display products, a unit is considered to have failed when its luminosity falls below 50% of its initial luminosity. Therefore, we will assume $y^* = \ln(0.50)$ for the PDP example. Conditioning on $\boldsymbol{\phi} \equiv (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the failure-time

cumulative distribution function is defined as

$$\begin{aligned} F_T(t|\boldsymbol{\phi}) &= \Pr(T \le t|\boldsymbol{\phi}) = \int_{\boldsymbol{\theta}} \Pr[g(t;\boldsymbol{\theta}) \le y^*] f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta} \\ &= \int_{\boldsymbol{\theta}} \Pr[g^{-1}(y^*;\boldsymbol{\theta}) \le y^*] f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta}. \end{aligned} \tag{2.13}$$

where $f(\boldsymbol{\theta}|\boldsymbol{\phi})$ is probability density function of the truncated trivariate normal distribution specified in the second-stage model (2.11). Because $F_T(t|\boldsymbol{\phi})$ is a function of $\boldsymbol{\phi}$, we can derive the marginal posterior distribution of the failure-time cumulative distribution function, $f(F_t(t)|\mathbf{d})$, from the joint posterior distribution of $\boldsymbol{\phi}$, i.e., $f(\boldsymbol{\phi}|\mathbf{d})$ via transformation of random variables. However, it is impossible to obtain a closed-form expression for $f(F_t(t)|\mathbf{d})$. Instead, we developed a simulation-based procedure and embedded it in the Gibbs sampling algorithm. In brief, at each iteration of the Gibbs sampling algorithm, a sample vector for $\boldsymbol{\mu}$ and a sample matrix for $\boldsymbol{\Sigma}$ are drawn from their conditional posterior distributions, respectively. Then, we randomly generate L $\boldsymbol{\theta}$ vectors according to the truncated trivariate normal distribution $f(\boldsymbol{\theta}|\boldsymbol{\phi})$, and compute the expected failure times by $g^{-1}(y^*|\boldsymbol{\theta})$. Finally, a sample value for $F_T(t)$ is estimated by the fraction of failure times that are less than or equal to t. Once the Gibbs sampling algorithm is executed, we obtain a set of $F_T(t)$ values that can be regarded as random sample drawn from the marginal posterior distribution $f(F_t(t)|\mathbf{d})$. Please refer to [4] for the detailed steps of the Gibbs sampling algorithm.

2.3 Degradation-Based Burn-in Optimization

This section presents the proposed Bayesian degradation-based burn-in method. The PDP example is again used as an illustrative example. A three-stage hierarchical Bayesian degradation model is employed to describe the burn-in population. The first-stage model assumes that the expected actual degradation path of the logarithm transformed relative luminosity of a random unit from the burn-in population is described by the mean function, $g(t;\boldsymbol{\theta})$, of the change-point regression model (2.4), that is,

$$g(t;\boldsymbol{\theta}) = \begin{cases} \alpha\, t - \beta t, & t \le \gamma, \\ \alpha\, t - \beta\gamma, & t > \gamma, \end{cases} \tag{2.14}$$

The second-stage model accounts for the unit-to-unit variability in the burn-in population by assuming the $\boldsymbol{\theta}$ to be random following a multivariate distribution $f(\boldsymbol{\theta}|\boldsymbol{\phi})$ with the parameter vector $\boldsymbol{\phi}$. Finally, the third-stage model specifies a joint prior distribution for $\boldsymbol{\phi}$, whose probability density function is denoted by $f(\boldsymbol{\phi})$.

The mission reliability, $R(t_m)$, is the probability that a random unit released to field operation will survive a pre-specified mission duration t_m. Conditioning on a given $\boldsymbol{\phi}$, the mission reliability without burn-in is defined as

$$R(t_m|\boldsymbol{\phi}) = \Pr(T > t_m|\boldsymbol{\phi}) = \int_{\boldsymbol{\theta}} \Pr[g(t_m;\boldsymbol{\theta}) > y^*] f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta}. \tag{2.15}$$

Because $\boldsymbol{\phi}$ is a random vector, $R(t_m)$, as a function of $\boldsymbol{\phi}$, is also a random variable. The prior distribution for the mission reliability $R(t_m)$, denoted by $f_{R_m}(r)$, can be derived from the prior distribution of $\boldsymbol{\phi}$, i.e., $f(\boldsymbol{\phi})$, via transformation of random variables according to Eq. (2.15). Then, the prior mean mission reliability $E[R(t_m)]$ is defined as

$$E[R(t_m)] = \int_0^1 r f_{R_m}(r) dr = \int_{\boldsymbol{\phi}} \left(\int_{\boldsymbol{\theta}} \Pr(g(t_m;\boldsymbol{\theta}) > y^*) f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta} \right) f(\boldsymbol{\phi}) d\boldsymbol{\phi}.$$

$f_{R_m}(r)$ and $E(R_m)$ have no closed-form expressions, and Monte Carlo simulation methods can be used to estimate them.

2.3.1 Reliability Criterion

The initial rapid degradation phase caused by impurities significantly reduces the mission reliability perceived by customers. Removing the rapid degradation phase through burn-in is crucial to the improvement of field reliability and customer satisfaction. This section presents a mission reliability criterion for planning the degradation-based burn-in test.

Assume all units in the burn-in population are subject to a degradation-based burn-in test at the designed operation conditions, and the burn-in duration is denoted by t_b. Upon completion of the burn-in degradation test, a unit is released to customers if its end-of-burn-in relative luminosity does not fall below a cutoff percentage η_b.

Conditioning on a given $\boldsymbol{\phi}$, the mission reliability of a unit that passed the burn-in test is defined as

$$\begin{aligned} R(t_m|t_b, \eta_b, \boldsymbol{\phi}) &= \Pr(T > t_m + t_b | T > t_b, \boldsymbol{\phi}) \\ &= \int_{\boldsymbol{\theta}} \Pr\left(g(t_m + t_b;\boldsymbol{\theta}) - g(t_b;\boldsymbol{\theta}) > y^* \,|g(t_b;\boldsymbol{\theta}) > \ln(\eta_b)\right) f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta}, \end{aligned} \tag{2.16}$$

where $T > t_b$ indicates that a unit is released to customers after the burn-in test because its end-of-burn-in relative luminosity does not fall below the cutoff value η_b, i.e., $g(t_b;\boldsymbol{\theta}) > \ln(\eta_b)$. Note that we take the logarithm for η_b because the function $g(\cdot)$ models the expected actual degradation path of the logarithm-transformed relative luminosity. The time-to-failure for a random unit released to field operation after the burn-in test is defined when the relative luminosity $g(t_m + t_b;\boldsymbol{\theta}) - g(t_b;\boldsymbol{\theta})$ is less than y^* because the initial luminosity perceived by the customers is the luminosity at the end of the burn-in test. The prior mean mission reliability after burn-in is, then, defined by

$$E[R(t_m|t_b,\eta_b)] = \int_{\boldsymbol{\phi}} R(t_m|t_b,\eta_b,\boldsymbol{\phi}) f(\boldsymbol{\phi}) d\boldsymbol{\phi} \tag{2.17}$$

$$= \int_{\boldsymbol{\phi}} \left[\int_{\boldsymbol{\theta}} \Pr\left(g(t_m + t_b;\boldsymbol{\theta}) - g(t_b;\boldsymbol{\theta}) > y^* \,|\, g(t_b;\boldsymbol{\theta}) > \ln(\eta_b) \right) f(\boldsymbol{\theta}|\boldsymbol{\phi}) d\boldsymbol{\theta} \right] f(\boldsymbol{\phi}) d\boldsymbol{\phi}.$$

The reliability criterion $E[R(t_m|t_b,\eta_b)]$ defined in Eq. (2.17) does not have a closed-form expression, but can be evaluated using the following Monte Carlo simulation algorithm:

1. Simulate many $\boldsymbol{\phi}$ vectors from $f(\boldsymbol{\phi})$;
2. For each $\boldsymbol{\phi}$ vector obtained in Step (*i*), simulate a $\boldsymbol{\theta}$ vector from $f(\boldsymbol{\theta}|\boldsymbol{\phi})$;
3. For all $\boldsymbol{\theta}$ vectors obtained in Step (*ii*), compute the ratio between the number of $\boldsymbol{\theta}$ vectors satisfying both $g(t_m + t_b;\boldsymbol{\theta}) - g(t_b;\boldsymbol{\theta}) > y^*$ and $g(t_b;\boldsymbol{\theta}) > \ln(\eta_b)$ and the number of $\boldsymbol{\theta}$ vectors satisfying only $g(t_b;\boldsymbol{\theta}) > \ln(\eta_b)$. This ratio provides an estimate of the $E[R(t_m|t_b,\eta_b)]$ criterion.

If the search space contains finite number of candidate plans, it may be possible to enumerate all candidate plans, compute their $E[R(t_m|t_b,\eta_b)]$ values, and, then, choose the desired plan. If it is infeasible to enumerate all candidate plans, the surface smoothing technique proposed by Muller and Parmigiani [16] may be used. This technique chooses a set of candidate plans that spread over the search space, compute the planning criterion values for these selected plans, fit a smooth surface using the method of kernels, and finally search for the desired plan on that smoothed surface. Other optimization techniques, such as the genetic algorithms, may also be used.

2.3.2 *Cost Criterion*

In this section, a cost criterion for planning degradation-based burn-in tests based on the change-point degradation model is presented. We adopt and extend the general cost model discussed in [20, 30] for planning failure-based burn-in tests. There are three types of costs: costs of the burn-in procedures, costs associated with rejecting weak units after burn-in, and costs for failures in field operations during the warranty period, t_w. Let C_{bs}, C_{bv}, C_{bf}, and C_{wf} denote the fixed burn-in cost per unit, the variable burn-in cost per unit per unit time, the burn-in rejection cost per unit, and the failure cost per unit during field operation. The expected cost per burn-in unit conditioning on a given $\boldsymbol{\phi}$ can be expressed by

$$\begin{aligned} E(C|t_b,\eta_b,\boldsymbol{\phi}) &= C_{bs} + C_{bv}t_b + C_{bf}\Pr(T < t_b|t_b,\eta_b,\boldsymbol{\phi}) + \\ &\quad + C_{wf}\Pr(T > t_b|t_b,\eta_b,\boldsymbol{\phi})\Pr(T < t_b + t_w|T > t_b,t_b,\eta_b,\boldsymbol{\phi}) \\ &= C_{bs} + C_{bv}t_b + C_{bf}\Pr(T < t_b|t_b,\eta_b,\boldsymbol{\phi}) \\ &\quad + C_{wf}\Pr(t_b < T < t_b + t_w|t_b,\eta_b,\boldsymbol{\phi}), \end{aligned}$$

where $T < t_b$ indicates that a unit is rejected at the end of the burn-in test because $g(t_b; \boldsymbol{\theta}) < \ln(\eta_b)$. $\Pr(T < t_b | t_b, \eta_b, \boldsymbol{\phi})$ and $\Pr(t_b < T < t_b + t_w | t_b, \eta_b, \boldsymbol{\phi})$ are defined as

$$\Pr(T < t_b | t_b, \eta_b, \boldsymbol{\phi}) = \int_{\boldsymbol{\theta}} \Pr[g(t_b; \boldsymbol{\theta}) < \ln(\eta_b)] f(\boldsymbol{\theta} | \boldsymbol{\phi}) d\boldsymbol{\theta}, \tag{2.18}$$

and

$$\begin{aligned} &\Pr(t_b < T < t_b + t_w | t_b, \eta_b, \boldsymbol{\phi}) = \\ &\int_{\boldsymbol{\theta}} \Pr[g(t_b; \boldsymbol{\theta}) > \ln(\eta_b) \cap (g(t_b + t_w; \boldsymbol{\theta}) - g(t_b; \boldsymbol{\theta})) < y^*] f(\boldsymbol{\theta} | \boldsymbol{\phi}) d\boldsymbol{\theta}, \end{aligned} \tag{2.19}$$

respectively.

The prior uncertainty in the parameter vector $\boldsymbol{\phi}$ is measured by its prior distribution $f(\boldsymbol{\phi})$, and, hence, the prior expected cost is defined as

$$\begin{aligned} E(C | t_b, \eta_b) &= \int_{\boldsymbol{\phi}} E(C | t_b, \eta_b, \boldsymbol{\phi}) f(\boldsymbol{\phi}) d\boldsymbol{\phi} \\ &= C_{bs} + C_{bv} t_b + C_{bf} \Pr(T < t_b | t_b, \eta_b) + C_{wf} \Pr(t_b < T < t_b + t_w | t_b, \eta_b), \end{aligned} \tag{2.20}$$

where $\Pr(T < t_b | t_b, \eta_b) = \int_{\boldsymbol{\phi}} \Pr(T < t_b | t_b, \eta_b, \boldsymbol{\phi}) f(\boldsymbol{\phi}) d\boldsymbol{\phi}$, and $\Pr(t_b < T < t_b + t_w | t_b, \eta_b) = \int_{\boldsymbol{\phi}} \Pr(t_b < T < t_b + t_w | t_b, \eta_b, \boldsymbol{\phi}) f(\boldsymbol{\phi}) d\boldsymbol{\phi}$. The Monte Carlo simulation algorithm presented in Sect. 2.3.1 can be modified to evaluate these two probabilities. In the step 3, $\Pr(T < t_b | t_b, \eta_b)$ is estimated by the fraction of simulated $\boldsymbol{\theta}$ vectors that satisfy $g(t_b; \boldsymbol{\theta}) < \ln(\eta_b)$, and $\Pr(t_b < T < t_b + t_w | t_b, \eta_b)$ is estimated by the fraction of simulated $\boldsymbol{\theta}$ vectors satisfying both $g(t_b; \boldsymbol{\theta}) > \ln(\eta_b)$ and $(g(t_b + t_w; \boldsymbol{\theta}) - g(t_b; \boldsymbol{\theta})) < y^*$.

2.3.3 Incorporation of Pre-burn-in Data

The Bayesian approach can update the prior distribution $f(\boldsymbol{\phi})$ with the addition of pre-burn-in data. Let $\mathbf{d}$ denote the available pre-burn-in data. The hierarchical Bayesian change-point degradation model presented Sect. 2.2.2 can be applied to analyze the data and obtain the posterior distribution of $\boldsymbol{\phi}$, that is, $f(\boldsymbol{\phi} | \mathbf{d})$. Then we can replace the prior distribution $f(\boldsymbol{\phi})$ in the prior mean mission reliability criterion (2.17) with the posterior distribution $f(\boldsymbol{\phi} | \mathbf{d})$ to obtain the posterior mean mission reliability after the burn-in test

$$\begin{aligned} &E[R(t_m | t_b, \eta_b, \mathbf{d})] \\ &= \int_{\boldsymbol{\phi}} \left[\int_{\boldsymbol{\theta}} \Pr\left(g(t_m + t_b; \boldsymbol{\theta}) - g(t_b; \boldsymbol{\theta}) > y^* \,|\, g(t_b; \boldsymbol{\theta}) > \ln(\eta_b) \right) f(\boldsymbol{\theta} | \boldsymbol{\phi}) d\boldsymbol{\theta} \right] \\ &\quad f(\boldsymbol{\phi} | \mathbf{d}) d\boldsymbol{\phi}. \end{aligned} \tag{2.21}$$

Similarly, the prior expected cost criterion (2.20) can be modified to obtain the posterior expected cost criterion

$$E(C|t_b, \eta_b, \mathbf{d}) = C_{bs} + C_{bv}t_b + C_{bf}\Pr(T < t_b|t_b, \eta_b, \mathbf{d}) + C_{wf}\Pr(t_b < T < t_b + t_w|t_b, \eta_b, \mathbf{d}), \tag{2.22}$$

where $\Pr(T < t_b|t_b, \eta_b, \mathbf{d}) = \int_{\boldsymbol{\phi}} \Pr(T < t_b|t_b, \eta_b, \boldsymbol{\phi}) f(\boldsymbol{\phi}|\mathbf{d}) d\boldsymbol{\phi}$, and $\Pr(t_b < T < t_b + t_w|t_b, \eta_b) = \int_{\boldsymbol{\phi}} \Pr(t_b < T < t_b + t_w|t_b, \eta_b, \boldsymbol{\phi}) f(\boldsymbol{\phi}|\mathbf{d}) d\boldsymbol{\phi}$. The step 1 of the Monte Carlo simulation algorithm presented in Sect. 2.3.1 needs to be modified by generating $\boldsymbol{\phi}$ vectors from the posterior distribution $f(\boldsymbol{\phi}|\mathbf{d})$, which can be done using the Gibbs sampling technique.

2.4 Results and Discussion

This section presents the results of the proposed methodologies applied to the PDP degradation data. The degradation measurement y is the logarithm of the relative luminosity, and t is the measurement time (in thousand hours).

2.4.1 Degradation Modeling and Failure-Time Distribution Estimation

For the purpose of comparison, a simple linear degradation model and a bi-exponential model are considered as alternatives. The simple linear degradation model is given by

$$y_j = h(t_j; \zeta) + \epsilon_j = \zeta t_j + \epsilon_j, \qquad j = 1, 2, \ldots, n, \tag{2.23}$$

where $\epsilon_j \sim N(0, \sigma_h^2)$. The bi-exponential model is given by [1]

$$\eta_j = q(t_j; \boldsymbol{\varphi}) = q(t_j; \pi, \lambda_1, \lambda_2) + \epsilon_j = \pi \exp(-\lambda_1 t_j) + (1 - \pi) \exp(-\lambda_2 t_j) + \epsilon_j, \tag{2.24}$$

for $j = 1, 2, \ldots, n$, where $\epsilon_j \sim N(0, \sigma_q^2)$, $0 < \pi < 1$ denotes the initial proportion of impurities, and $\lambda_1 > 0$ and $\lambda_2 > 0$ represent the degradation rates of the impurities and phosphors, respectively. The response η is the relative luminosity, i.e., $\eta = \exp(y)$.

Table 2.1 Maximum likelihood estimates for the parameters in the change-point degradation model (2.4) and the simple linear degradation model (2.23)

	Change-point model				Linear model	
PDP	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$ (10^3 h)	$\hat{\sigma}^2$ ($\times 10^{-4}$)	$\hat{\zeta}$	$\hat{\sigma}_h^2$ ($\times 10^{-2}$)
#1	−0.0584	0.3189	0.5641	3.889	−0.1259	0.941
#2	−0.0602	0.4632	0.5346	5.077	−0.1534	1.814
#3	−0.0608	0.4088	0.5774	4.644	−0.1493	1.580
#4	−0.0620	0.4033	0.5410	4.539	−0.1440	1.404
#5	−0.0523	0.1656	0.7340	2.220	−0.0981	0.363
#6	−0.0649	0.2689	0.4536	2.572	−0.1130	0.536

2.4.1.1 Individual Degradation Modeling

First, we fitted the six individual degradation paths with the change-point regression model (2.4) and the simple linear degradation model (2.23) using the maximum likelihood method. Maximum likelihood estimates (MLEs) of the model parameters of those two models listed in Table 2.1. The likelihood ratio test was employed to compare the simple linear degradation model (the null model) and the change-point degradation model (the alternative model), and yielded p-value ≈ 0 for all six PDP degradation paths, strongly supporting the two-phase degradation model.

Next, we fitted the six individual degradation paths with the change-point degradation model (2.4), the simple linear degradation model (2.23), and the bi-exponential degradation model (2.24) using the Bayesian method. Due to an absence of prior knowledge, we adopted non-informative priors with $\boldsymbol{\mu} = (0, 0, 0)'$, $\boldsymbol{\Sigma} = \text{diag}(10^6, 10^6, 10^6)$, $a_\sigma = 1$, and $b_\sigma = 0.0001$. In addition, we assumed $\phi = t_n$ to reflect our prior assumption that there is one change-point in the test period $(0, t_n)$. The Gibbs sampling algorithm was run for 100,000 iterations and the first 50,000 iterations were discarded. Convergence were monitored and verified by running multiple chains from dispersed initial values [12]. Table 2.2 summarizes the posterior medians of the parameters in the change-point degradation model (2.4), along with the 95% Bayesian confidence intervals (in parentheses). Tables 2.1 and 2.2 indicate that the Bayesian method and the maximum likelihood method provided very similar results. This is because non-informative priors were used in the Bayesian analysis and the Bayesian inference was largely based on the data information contained in the likelihood function.

Bayesian fitting algorithms using non-informative priors were also developed for the simple linear degradation model (2.23) and the bi-exponential degradation model (2.24). For the simple linear degradation model, the prior distributions are $\zeta \sim N(0, 10^6)_{I_{\{\zeta<0\}}}$ and $\sigma_h^2 \sim IG(1, 0.0001)$. For the bi-exponential degradation model, the prior distributions are $\boldsymbol{\varphi} \equiv (\pi, \lambda_1, \lambda_2) \sim N_3(\mathbf{0},\ 10^6\mathbf{I}_3)_{I_{\{0<\pi<1, \lambda_1>0, \lambda_2>0\}}}$ and $\sigma_q^2 \sim IG(1, 0.0001)$. Figure 2.2 compares the posterior medians of the expected relative luminosity predicted by the three different models, i.e., $\exp(g(t; \boldsymbol{\theta}))$, $\exp(h(t; \zeta))$, and $q(t; \boldsymbol{\varphi})$. The linear model fails to capture the luminosity degradation, whereas the bi-exponential model and the change-point model can effectively

Table 2.2 Bayesian inference for the parameters in the individual change-point degradation model (2.4)

PDP	α median (95% interval)	β median (95% interval)	γ (10^3 h) median (95% interval)	σ^2 ($\times 10^{-4}$) median (95% interval)
#1	−0.05831 (−0.06726, −0.04919)	0.3160 (0.2556, 0.4076)	0.5694 (0.2556, 0.4076)	4.334 (2.534, 8.315)
#2	−0.06185 (−0.07264, −0.05170)	0.4855 (0.4079, 0.6259)	0.5021 (0.3688, 0.6197)	5.592 (3.282, 10.670)
#3	−0.06037 (−0.06992, −0.05069)	0.4019 (0.3417, 0.4728)	0.5887 (0.4795, 0.7208)	5.108 (2.955, 9.792)
#4	−0.06241 (−0.07180, −0.05298)	0.4068 (0.3407, 0.5061)	0.5351 (0.4058, 0.6636)	5.038 (2.949, 9.629)
#5	−0.05198 (−0.06105, −0.04283)	0.1646 (0.1395, 0.2083)	0.7461 (0.4945, 0.9648)	2.517 (1.437, 5.029)
#6	−0.06401 (−0.07234, −0.05404)	0.2594 (0.1960, 0.3608)	0.4734 (0.3116, 0.7337)	2.921 (1.676, 5.766)

Table 2.3 Average of squared residuals of the change-point degradation model (2.4), the simple linear degradation model (2.23), and the bi-exponential model (2.24)

Subject	Change-point model	Linear model	Bi-exponential model
PDP #1	0.0056	0.1381	0.0020
PDP #2	0.0068	0.2479	0.0022
PDP #3	0.0061	0.2136	0.0019
PDP #4	0.0059	0.1951	0.0030
PDP #5	0.0035	0.0566	0.0016
PDP #6	0.0034	0.0819	0.0015

capture the two-phase degradation pattern. Table 2.3 compares the average of the squared residuals of the three models. The residual is defined as the deviation of an observed relative luminosity from the posterior median of the expected relative luminosity predicted by a model. Table 2.3 indicates that the bi-exponential model and the change-point degradation model fit the PDP degradation data better than the simple linear degradation model. The bi-exponential model provides a better fit to the data than the change-point degradation model. However, the change-point degradation model can provide additional information related to incomplete burn-in through the change-point estimation.

2.4.1.2 Hierarchical Bayesian Degradation Modeling

We, then, analyzed the six observed degradation paths using the three-stage hierarchical Bayesian change-point degradation model presented in Sect. 2.2.2 and derived the failure-time distribution of a random unit from the population according

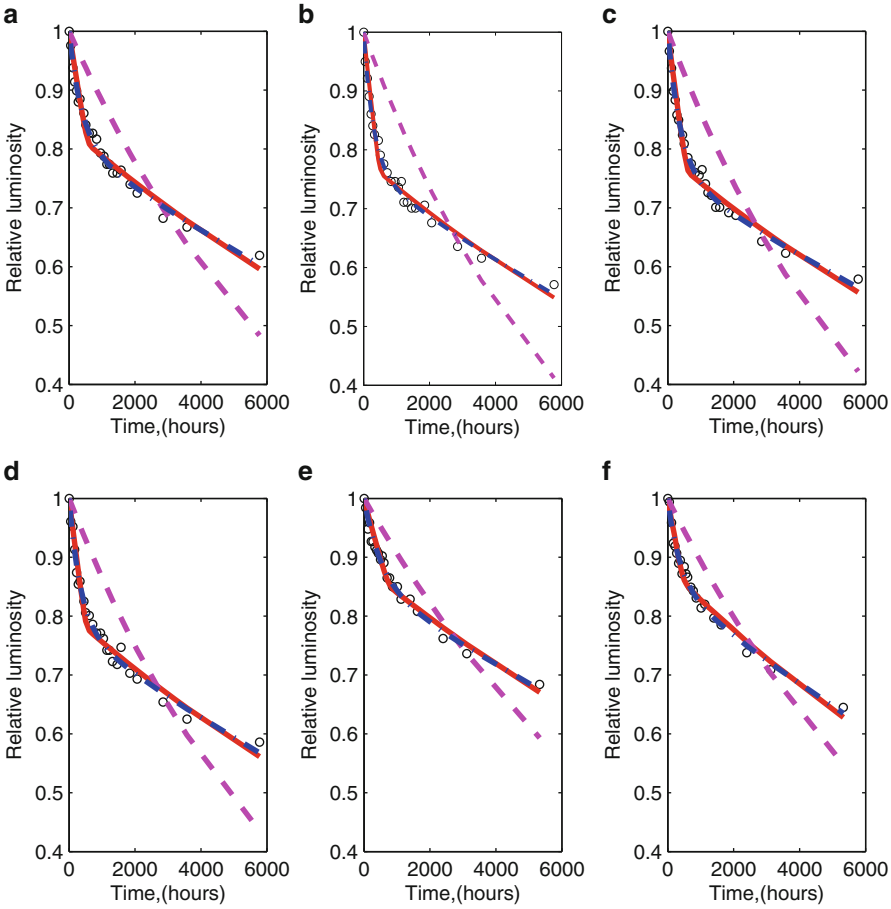

Fig. 2.2 Posterior medians of the expected degradation paths of relative luminosity estimated by different models: —: change-point degradation model (2.4); − − −: simple linear degradation model (2.23); − · −: bi-exponential degradation model (2.24); ∘: measurement. (**a**) PDP #1. (**b**) PDP #2. (**c**) PDP #3. (**d**) PDP #4. (**e**) PDP #5. (**f**) PDP #6

to the method outlined in Sect. 2.2.3. In the third stage of the hierarchical model, we assigned the following non-informative priors:

$$\boldsymbol{\mu} \sim N(\mathbf{0},\ 10^6\mathbf{I}_3),\ \mathbf{Q} \sim IW(4, \mathbf{I}_3),\ \sigma^2 \sim IG(1, 0.0001),\ \text{and}\ \delta_i \sim G(1, 0.0001), \tag{2.25}$$

for $i = 1, 2, 3$, and additionally, we assumed $\phi = 5$ (thousand hours). The failure threshold is $y^* = \ln(50\%)$. Figure 2.3 shows the posterior median of the failure-time cumulative distribution function. The number of simulated degradation coefficients, L, was selected to be 3,000. A hierarchical Bayesian linear degradation model and a hierarchical Bayesian bi-exponential degradation model were also developed in

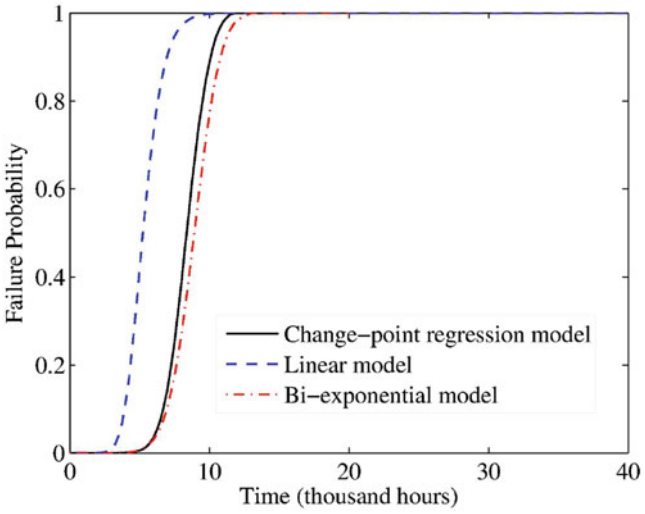

Fig. 2.3 Posterior medians of the failure-time distribution $F_T(t)$ predicted by different models

this study (not presented in this article). Their predicted failure-time distributions are also plotted in Fig. 2.3. The failure threshold for the bi-exponential model is $\eta^* = 50\%$. The hierarchical Bayesian linear degradation model underestimates the reliability of the PDPs, which is consistent with the results shown in Fig. 2.2. Because the derivation of the failure-time distribution requires extrapolation beyond the range of the observed data, different model assumptions may cause significantly different predictions for the failure-time distribution. The hierarchical Bayesian change-point and the hierarchical Bayesian bi-exponential degradation model provide reasonably consistent predictions for the failure-time distribution.

2.4.2 Burn-in Test Planning

This section presents the results of the burn-in test planning method according to the posterior mean mission reliability criterion (2.21) and the posterior expected cost criterion (2.22). We consider two scenarios of the cutoff value η_b. In the first scenario, η_b is fixed at 0%, which means that there is no end-of-burn-in inspection and all units are released to customers after the burn-in test. This scenario eliminates the rapid initial degradation phase. The burn-in duration t_b is the only decision variable in the burn-in planning problem. In the second scenario, the burn-in duration t_b and the cutoff point η_b are both decision variables. In this scenario, the purposes of the burn-in degradation test are to eliminate the rapid degradation phase and reject weak units.

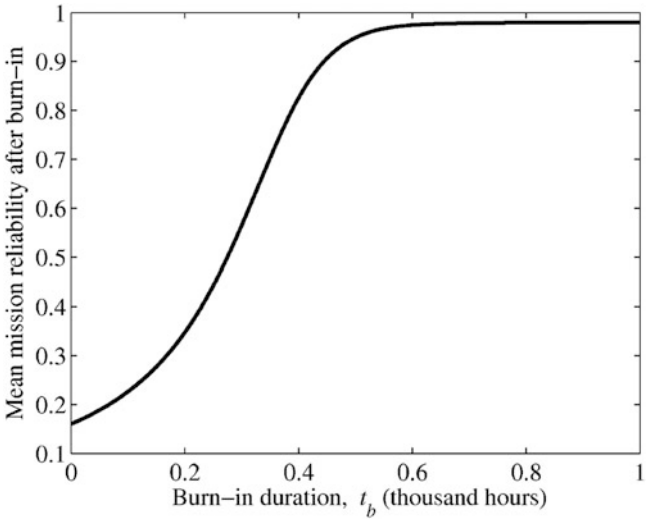

Fig. 2.4 The posterior mean mission reliability after burn-in $E[R(t_m|t_b, \eta_b, \mathbf{d})]$ vs. the burn-in duration t_b when the cutoff value $\eta_b = 0$

2.4.2.1 Planning Burn-in Without Inspection

First, we assume that there is no-end-of-burn-in inspection and all units are released to customers. In this scenario, η_b is fixed at 0%, and the burn-in duration t_b is the only decision variable. We assume the search space for t_b is from 0 to 1,000 h with an increment of 1 h and enumerate all candidate t_b values in the search space. Figure 2.4 plots the posterior mean mission reliability $E[R(t_m|t_b, \eta_b, \mathbf{d})]$ vs. the burn-in duration t_b, assuming a mission duration $t_m = 10{,}000$ h. When $t_b = 0$, the mean mission reliability is less than 20%. $E[R(t_m|t_b, \eta_b, \mathbf{d})]$ increases as t_b increases, and tends to saturate when t_b is above around 600 h. This is consistent with the results shown in Table 2.2. The posterior median of the change-point of the PDPs are around 600 h. Therefore, when the burn-in duration is around 600 h, we expect the initial rapid degradation phase to be effectively eliminated. Figure 2.4 can be used to obtain the minimum burn-in duration necessary to achieve a requirement for the mean mission reliability. For example, when the mean mission reliability is required to be 90%, the minimum burn-in duration is 445 h.

Next, we plan the burn-in test according to the posterior expected cost criterion $E(C|t_b, \eta_b, \mathbf{d})$ (2.22). For illustration purpose, we assume $C_{bs} = 10$, two different values for C_{bv} (0.10 and 0.18), two different values for C_{wf} (200 and 300), and $t_w = t_m = 10{,}000$ h. Because C_{bs} does not affect the optimal burn-in duration, only one value is assumed for C_{bs}. Since the burn-in rejection cost is not considered, $C_{bf} = 0$. Figure 2.5 plots $E(C|t_b, \eta_b, \mathbf{d})$ as a function of the burn-in duration t_b for the combination of $C_{wf} = 300$ and $C_{bv} = 0.10$. The cost-optimal burn-in plan is a

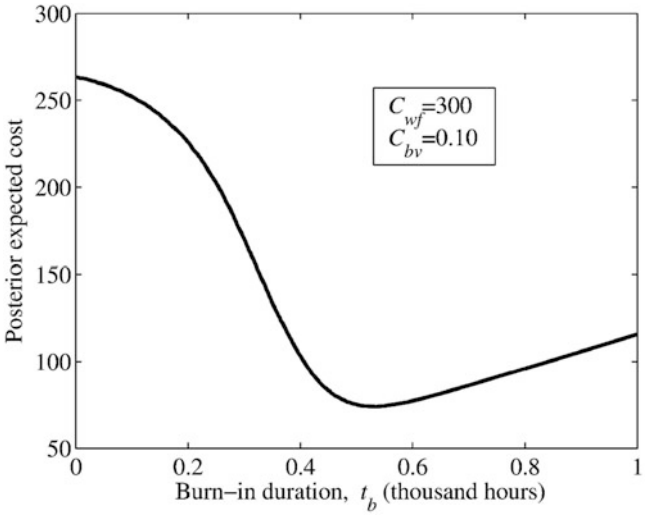

Fig. 2.5 The posterior expected cost $E(C|t_b, \eta_b, \mathbf{d})$ vs. the burn-in duration t_b when the cutoff value $\eta_b = 0\%$ and the cost coefficients are $C_{wf} = 300$ and $C_{bv} = 0.10$

Table 2.4 Cost-optimal burn-in duration t_b^* for different cost coefficients, when $\eta_b = 0$

C_{bv}	C_{wf}	C_{bv}/C_{wf}	t_b^* (h)	$\Pr(t_b^* < T < t_b^* + t_w \mid t_b^*, \eta_b, \mathbf{d})$
0.10	300	0.0003	528	0.037
0.10	200	0.0005	514	0.043
0.18	300	0.0006	497	0.052
0.18	200	0.0009	476	0.069

trade-off between the burn-in cost and the warranty cost. Table 2.4 summarizes the optimal burn-in durations for different cost coefficients.

The cost-optimal burn-in duration t_b^* depends on the relative relationship between C_{bv} and C_{wf}. As the ratio C_{bv}/C_{wf} increases, the burn-in costs increases, which results in a reduction in the optimal burn-in duration. Additionally, as the burn-in duration increases, the field failure probability after burn-in $\Pr(t_b^* < T < t_b^* + t_w|t_b^*, \eta_b, \mathbf{d})$ decreases, which is consistent with the results presented in Fig. 2.4.

2.4.2.2 Planning Burn-In with Inspection

This section considers the case with end-of-burn-in inspection. A unit is rejected if its end-of-burn-in relative luminosity is below the cutoff percentage η_b. Figure 2.6 shows the posterior mean mission reliability as a function of the burn-in duration for different values of η_b. As shown in this figure, for the same burn-in duration, increasing the cutoff value improves the mission reliability after the burn-in test. If a mean mission reliability of 90% is required, the minimum burn-in durations

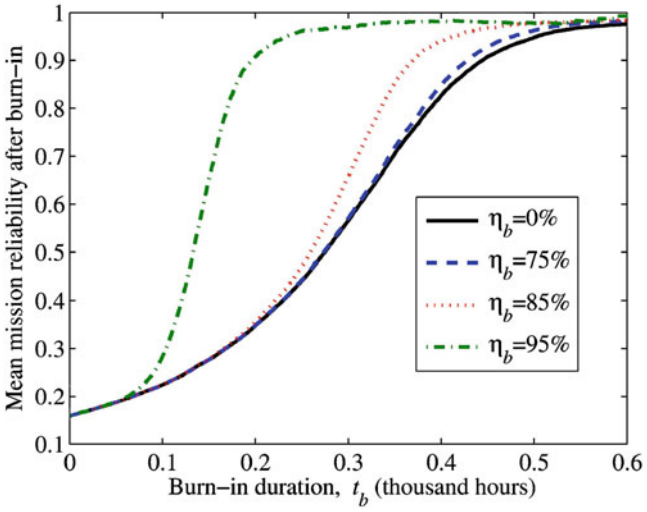

Fig. 2.6 The posterior mean mission reliability after burn-in $E[R(t_m|t_b, \eta_b, \mathbf{d})]$ vs. the burn-in duration t_b for different cutoff values η_b

Table 2.5 Minimum burn-in duration to achieve the desired mission reliability for different cutoff values

η_b	Minimum burn-in duration, t_b, (h)	$\Pr(T < t_b\|t_b, \eta_b, \mathbf{d})$
0.00	445	–
0.75	428	0.06
0.85	370	0.37
0.95	197	0.81

are listed in Table 2.5. When the cutoff value η_b increases, the minimum burn-in duration to achieve the 90% mission reliability reduces. However, this reduction of burn-in duration is achieved by rejecting more units after burn-in as shown in the $\Pr(T < t_b|t_b, \eta_b, \mathbf{d})$ column of Table 2.5. This may incur unnecessarily high burn-in rejection cost.

Next, we obtain the cost-optimal burn-in duration and cutoff values by minimizing the posterior expected cost $E(C|t_b, \eta_b, \mathbf{d})$. We assume that the candidate values for η_b are from 0% to 100% with an increment of 1%. Two different values (100 and 150) are assumed for the burn-in rejection cost C_{bf}. Table 2.6 lists the optimal burn-in test plans (t_b^*, η_b^*) for the eight combinations of C_{bv}, C_{bf}, and C_{wf} values. We have observed the following trends. Firstly, t_b^* and η_b^* are negatively correlated, that is, a longer burn-in duration tends to have a lower cutoff value. This is because the expected actual degradation path of PDPs is a monotonic decreasing function. Secondly, for the same C_{bf} value, the optimal burn-in duration t_b^* again depends on the ratio C_{bv}/C_{wf}, and a higher ratio results in a lower burn-in duration. Finally, for the same set of C_{bv} and C_{wf} values, a higher C_{bf} value causes the optimal burn-in

Table 2.6 Cost-optimal burn-in plans $(t_b^*,\ \eta_b^*)$ for different cost coefficients

C_{bf}	C_{bv}	C_{wf}	t_b^* (h)	η_b^*	$\Pr(T < t_b^* \mid t_b^*, \eta_b^*, \mathbf{d})$	$\Pr(t_b^* < T < t_b^* + t_w \mid t_b^*, \eta_b^*, \mathbf{d})$
100	0.10	200	509	0.42	0.0007	0.0456
100	0.18	200	457	0.62	0.0052	0.0793
100	0.10	300	533	0.54	0.0037	0.0360
100	0.18	300	495	0.49	0.0012	0.0512
150	0.10	200	501	0.36	0.0003	0.0478
150	0.18	200	465	0.45	0.0004	0.0735
150	0.10	300	524	0.42	0.0006	0.0386
150	0.18	300	495	0.43	0.0007	0.0524

plan to lower the burn-in rejection probability $\Pr(T < t_b^* \mid t_b^*, \eta_b^*, \mathbf{d})$. To lower this burn-in rejection probability, the optimal cutoff value is reduced.

2.5 Conclusion

Some products were found to present two-phase degradation patterns where an initial rapid degradation phase followed by a slower and more stable degradation phase. In this chapter, we demonstrated that the change-point regression is a reasonable model to describe the observed two-phase degradation paths. For the PDP example, although the bi-exponential model provides a better fit to the degradation paths, a major advantage of the change-point regression model is the estimation of the change-point between the two degradation phases, and it may provide valuable information to the manufacturers to determine the necessary duration of burn-in process in order to eliminate the initial rapid degradation phase before shipping the products to the customers. Based on the change-point regression model, this chapter also discussed the degradation-based burn-in planning according to a cost criterion or a reliability criterion. The Bayesian approach was employed in order to incorporate the available prior knowledge and/or pre-burn-in data.

References

1. Bae SJ, Kim SJ, Kim MS et al (2008) Degradation analysis of nano-contamination in plasma display panels. IEEE Trans Reliab 57:222–229
2. Bae SJ, Kim SJ, Um SK et al (2009) A predition model of degradation rate for membrance electrode assemblies in direct methanol fuel cells. Int J Hydrog Energy 34:5749–5758
3. Bae SJ, Kvam PH (2006) A change-point analysis for modeling incomplete burn-in for light displays. IIE Trans 38:489–498
4. Bae SJ, Yuan T, Ning S et al (2015) A Bayesian approach to modeling two-phase degradation using change-point regression. Reliab Eng Syst Saf 134:66–74

5. Barnard J, McCulloch R, Meng XL (2000) Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. Stat Sin 10:1281–1311
6. Beem AL (1995) A program for fitting two-phase segmented-curve models with an unknown change point, with an application to the analysis of strategy shifts in congnitive task. Behav Res Methods Instrum Comput 27:392–399
7. Carlin BP, Gelfand AE, Smith AFM (1992) Hierarchical Bayesian analysis of change point problems. Appl Stat 41:389–405
8. Chen N, Tsui KL (2013) Condition monitoring and remaining useful life prediction using degradation signals: revisited. IIE Trans 45:939–952
9. Elsayed EA (2012) Overview of reliability testing. IEEE Trans Reliab 61:282–291
10. Feng Q, Peng H, Coit DW (2010) A degradation-based model for joint optimization of burn-in, quality inspection, and maintenance: a light display device application. Int J Adv Manuf Technol 50:801–808
11. Gebraeel NZ, Lawley MA, Li R et al (2005) Residual-life distributions from component degradation signals: a Bayesian approach. IIE Trans 37:543–557
12. Gelman A, Carlin JB, Stern HS et al (2004) Bayesian data analysis, 2nd edn. Chapman & Hall, Boca Raton
13. Kuo W, Chien WTK, Kim T (1998) Reliability, yield, and stress burn-in. Wiley, Boston
14. Lu CJ, Meeker WQ, Escobar LA (1996) A comparison of degradation and failure-time analysis methods for estimating a time-to-failure distribution. Stat Sin 6:531–546
15. Muggeo VMR (2003) Estimating regression models with unknown break-points. Stat Med 22:3055–3071
16. Muller P, Parmigiani G (1995) Optimal design via curve fitting of Monte Carlo experiments. J Am Stat Assoc 90:1322–1330
17. Park JI, Bae SJ (2010) Direct prediction method on lifetime distribution of organic light-emitting diodes from accelerated degradation tests. IEEE Trans Reliab 59:74–90
18. Park JI, Baek SH, Jeong MK et al (2009) Dual features functional support vector machines for fault detection of rechargeable batteries. IEEE Trans Syst Man Cybern C: Appl Rev 39:480–485
19. Peng H, Feng Q, Coit DW (2009) Simultaneous quality and reliability optimization for microengines subject to degradation. IEEE Trans Reliab 58:98–105
20. Perlstein D, Jarvis WH, Mazzuchi TA (2001) Bayesian calculation of cost optimal burn-in test durations for mixed exponential populations. Reliab Eng Syst Saf 72:265–273
21. Tsai CC, Tseng ST, Balakrishnan N (2011) Optimal burn-in policy for highly reliable products using gamma degradation process. IEEE Trans Reliab 60:234–245
22. Tseng ST, Peng CY (2004) Optimal burn-in policy by using an integrated Wiener process. IIE Trans 36:1161–1170
23. Tseng ST, Tang J (2001) Optimal burn-in time for highly reliable products. Int J Ind Eng 8:329–338
24. Tseng ST, Tang J, Ku IH (2003) Determination of burn-in parameters and residual life for highly reliable products. Nav Res Logist 50:1–14
25. Winbugs example volume 2 (2012) http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/Vol2.pdf, pp 38–41
26. Xiang Y, Coit DW, Feng Q (2013) n Subpopulations experiencing stochastic degradation: reliability modeling, burn-in, and preventive replacement. IIE Trans 45:391–408
27. Ye ZS, Shen Y, Xie M (2012) Degradation-based burn-in with preventive maintenance. Eur J Oper Res 221:360–367
28. Ye ZS, Xie M, Tang LC et al (2012) Degradation-based burn-in planning under competing risks. Technometrics 54:159–168
29. Yuan T, Bae SJ, Zhu X (2016) A Bayesian approach to degradation-based burn-in optimzation for display products exhibiting two-phase degradation patterns. Reliab Eng Syst Saf 155:55–63
30. Yuan T, Kuo Y (2010) Bayesian analysis of hazard rate, change point, and cost-optimal burn-in time for electronic devices. IEEE Trans Reliab 59:132–138

Chapter 3
Degradation Modeling, Analysis, and Applications on Lifetime Prediction

Lirong Hu, Lingjiang Li, and Qingpei Hu

Abstract Degradation signals provide more information for product life status than failure data, when specific degradation mechanism can be identified. Modeling and analysis with the degradation signal is helpful to extrapolate for product lifetime prediction. In this chapter, comprehensive review has been conducted for different kinds of modeling and analysis approaches, together with the corresponding lifetime prediction results. Furthermore, discussions over related issues like product initial performance are presented.

Keywords Degradation models • Acceleration models • Parameter estimation • Lifetime prediction • Initial performance

3.1 Introduction

With the rapid development of technology, products of high-quality gradually appear in a wide range of areas including electronic communications, weapons manufacturers, aerospace industries, and so on. As a result, reliability evaluation on those products has been brought to the forefront in recent decades.

We may assess whether a product is of high-quality from various views, such as its appearance and its multiformity of function. However, nobody will regard an unreliable product as a high-quality one. What's reliability? Reliability is the probability of a product or a system performing its intended function without failure for a specified period of time under specific condition [1].

L. Hu • Q. Hu (✉)
Center of Quality and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
e-mail: qingpeihu@amss.ac.cn

L. Li
Search and Big Data Division, Bejing Jingdong Shangke Information Technology Co., LTD, Beijing, China

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_3

3.1.1 Traditional Reliability Analysis

Traditionally, reliability analysis is based on failure-time data. A product's failure time is the time it failed. Generally, there're two modes of failure: soft failure and hard failure. Soft failure happens when the performance of a product which do not fail catastrophically, reach a specified level of degradation. For example, Wang and Chu [2] defined the failure time of an LED-based light bar as the time required for its light output to achieve 50% lumen maintenance. Hard failure is caused by sudden accidents, and products always lose all their functionality when hard failure occurs.

In reliability analysis, we're supposed to get product's lifetime distribution from failure-time data. Assume that $T_i \sim F(\bullet;\theta)(i = 1, 2, \ldots, n)$ where $T_1, T_2, \ldots, T_n$ is a set of lifetimes of the product. The distribution type may not be known at the very beginning, however, we can use the probability plot to assess which distribution type the failure-time data fit best. And we may find in many applications, product's lifetimes usually fit a Weibull or lognormal distribution. After the form of $F(\bullet;\theta)$ is known, maximum likelihood method can be used to estimate the parameter θ.

3.1.2 Degradation Data

In many cases, we may find it difficult to observe the failure time of a highly reliable product. Therefore, looking for an alternative way to do reliability analysis is urgent. As a result, researchers started to use degradation data.

It is easy to find out that degradation data can be collected only soft failures occur. Generally, degradation can be defined as the reduction in performance. And the degradation phenomenon is very common in everyday life, for example, battery level of telephone will decrease under using condition. When the degradation of a product can be measured, we can regard it as a function of time, and we can take measurements on it at different points of time to obtain degradation data.

Comparing with failure-time data, degradation data can provide more information. On one hand, degradation data record the failure process of the product while failure-time data only record the failure time. On the other hand, it's difficult to obtain failure time of a highly reliable product even under severe conditions, as a result, reliability analysis based on failure-time data can't go on smoothly.

Nevertheless, there're also some limitations of degradation data. For one thing, hard failures can't be traced to degradation, only failure-time data is available. For another, in some applications, measurements on degradation data are difficult to obtain or the measurement is destructive. For example, Nelson [1] discussed a special situation in which the degradation measurement is destructive. And Escobar et al. [3] introduced an important class of models for accelerated destructive degradation data. In such case, failure-time data still play an important role in reliability analysis.

3.1.3 Accelerated Degradation Testing

In normal use conditions, a degradation process may be slow, and it will bring difficulties in degradation data gathering. For that reason, accelerated degradation tests (ADTs) appear in reliability analysis, and this kind of tests is applied by increasing the level of acceleration variables, such as vibration amplitude, temperature, corrosive media, load, voltage, pressure [4, 5].

3.1.3.1 Three Types of ADTs

Degradation data can provide much more information than failure-time data, as a result, accelerated degradation tests are designed to gather reliability information of product. The accelerated degradation test is widely used in reliability analysis. A well-planned accelerated degradation test can lead to the most accurate estimation. The plan variables include test stress, sample size, measurement times and termination time. There're three types of accelerated degradation tests: constant-stress ADTs, step-stress ADTs and progressive stress ADTs.

In a constant-stress ADT, several stress levels are divided, and each stress level requires a number of samples. For new developing products, it's not an easy task to determine suitable stress levels. Thus, conducting a constant-stress ADT can be very costly. Zhang et al. [6] used luminance decaying models under three groups of constant stress accelerated degradation tests to acquire the life information of OLED. Guan et al. [7] provided a constant-stress ADT in assessing product reliability based on a Wiener process. Pan et al. [8] proposed a bivariate constant-stress accelerated degradation model for product's lifetime inference.

Comparing with constant-stress ADT, the step-stress ADT is an economical and flexible test plan. The stress is increased in steps under a step-stress ADT. As a result, the step-stress ADT do not require a large sample size, and it can avoid sudden stress shocks leading to a sudden failure of the product. Wan et al. [9] found it is more efficient by carrying out ADT with step-down stress than constant stress. Pan and Sun [10] designed an efficient step-stress ADT plan for the products with multiple performance characteristics based on Gamma process. Ge et al. [11] presented a method to optimal design for step-stress ADT based on D-optimality.

Process-stress ADT is similar to the step-stress ADT, except that the stress of the test is increased with time at some constant rate. However, increasing stress at a specific rate is difficult. Therefore, there are lots of problems on the application of process-stress ADT. Guo [12] proposed an online degradation model based on a process-stress ADT. Peng and Tseng [13] proposed a progressive-stress accelerated degradation test for highly reliable products with very few test units.

3.1.3.2 Accelerated Degradation Models

In order to get product's lifetime distribution using degradation data, the core step is to set up a model describing the degradation process. And under accelerated degradation tests, this kind of model is called an accelerated degradation model.

An accelerated degradation model is the combination of an accelerated model and a degradation model. And accelerated degradation models include the physic based models and the statistics based models. Boulanger and Escobar [14] reviewed accelerated degradation models in details. Park and Padgett [15] provided some accelerated degradation models for failure based on the geometric Brownian motion or Gamma process. Tang et al. [16] discussed accelerated degradation tests modelling based on the nonlinear Wiener process. Sang and Kim [17] used an accelerated degradation model to estimate the life of elevator wire rope. Liu [18] proposed a Gauss-power law accelerated degradation model for electronic products. Yang et al. [19] used several accelerated degradation models for prediction of the time-to-failure of smart electricity meter (SEM). Hayes et al. [20] discussed accelerated degradation models as ways to characterize the interaction between disability and the functional decline of aging.

3.1.4 Overview

The remainder of this paper is organized as follows: Section 3.2 reviews acceleration models for reliability analysis, and it focuses on three types of widely used acceleration models. In Sect. 3.3, two types of degradation model are introduced, and they are general path models and stochastic models. Furthermore, three common methods of parameter estimation are also discussed in Sect. 3.3.3. Section 3.3.4 describes methods of lifetime distribution evaluation, including approximate accelerated degradation analysis. Section 3.4 shows the initial performance is often nonzero and it is random. And mixed-effect general path models are set up to describe the relationship between initial performance and the degradation rate. A short conclusion is given in Sect. 3.5, as well as discussions on future study.

3.2 Acceleration Models

This paper will focus on three types of widely used acceleration models: the usage rate acceleration models, the temperature acceleration models and the voltage acceleration models.

3.2.1 Usage Rate Acceleration Models

Running the product at a higher usage rate is a way to accelerate degradation. Yang and Zaghati [21] increased the usage rate of the product in an accelerated life testing. Yellamati et al. [22] used vacuum fluorescent displays (VFD) heat sink temperatures under worst-case by accelerating the usage rate. Hwang et al. [23] developed the usage rate acceleration model using the degradation data. Assuming the degradation is usage rate independent, there are usually two ways of doing such accelerated test.

One way is to run the product faster. For example, in many applications, it's common to increase the rotate speed of a rolling bearing. The acceleration factor can be written as $AF(UR) = UR/UR_0$, where UR_0 is the operating usage rate of the product. Another way is to increase the operating frequency or reduce the off time of the product. Some products, such as the washing machine, run about an hour a day, however, they can run 24 hours a day under accelerated testing. The acceleration factor can also be written as $AF(UR) = UR/UR_0$, where UR_0 is the operating usage rate of the product.

3.2.2 Temperature Acceleration Models

As we mentioned before, temperature is one common cause of degradation. For example, increasing temperature can speed up many chemical reactions. Arrhenius and Eyring relationship are two widely used models when the accelerated variable is thermal.

3.2.2.1 Arrhenius Relationship

The Arrhenius relationship is probably the most common model used in accelerated testing. It has been widely used when the accelerated variable is temperature. Shen et al. [24] used the Arrhenius accelerated aging model to estimate the product lifespan. Cooper [25] used the Arrhenius equation for analysis of product's life test data. Zhou et al. [26] converted the failure data to normal stress employing acceleration factors of Arrhenius model. The Arrhenius relationship is derived from the Arrhenius reaction rate equation, and it can be expressed as:

$$R_{Ar}(temp) = \gamma_0 \cdot \exp\left(\frac{-E_a}{k_b \cdot temp}\right) = \gamma_0 \cdot \exp\left(\frac{-E_a \cdot 11605}{temp}\right). \tag{3.1}$$

Where R is the reaction rate; γ_0 is a constant depending on product or material characteristics; E_a is the activation energy; $k_b = 8.6171 \times 10^{-5} = 1/11605$ is the Boltzmann's constant; *temp* is the ambient temperature in degrees Kelvin.

The Arrhenius acceleration factor is:

$$AF_{Ar}(temp) = \frac{R_{Ar}(temp)}{R_{Ar}\,(temp_0)} = \exp\left[\frac{E_a}{k_b}\left(\frac{1}{temp_0} - \frac{1}{temp}\right)\right]. \tag{3.2}$$

Where $temp_0$ is the operating temperature of the product under test. From the Eq. (3.2), when $temp > temp_0$, we have $AF(temp) > 1$; and when $temp < temp_0$, we have $AF(temp) < 1$.

3.2.2.2 Eyring Relationship

Just like the Arrhenius reaction rate equation, the Eyring equation describes the temperature dependence of reaction rate. Dai et al. [27] built the degradation model with the theoretical description of side reaction based on Eyring equation. Endicott et al. [28] discussed the relation between the Eyring model and the Weibull distribution for both progressive and constant stress tests. The Eying relationship is given by:

$$R_{Ey}(temp) = \gamma_0 \cdot A(temp) \cdot \exp\left(\frac{-E_a}{k_b \cdot temp}\right) = A(temp) \cdot R_{Ar}(temp). \tag{3.3}$$

Where $A(temp)$ is a function of Kelvin temperature $temp$. In a lot of applications, the function $A(temp)$ is usually denoted as $A(temp) = temp^m$ with a constant m.

The Eyring acceleration factor is:

$$AF_{Ey}(temp) = \frac{A(temp)}{A\,(temp_0)} \cdot AF_{Ar}(temp). \tag{3.4}$$

When $A(temp) = temp^m$, the Eyring acceleration factor can be written as:

$$AF_{Ey}(temp) = \left(\frac{temp}{temp_0}\right)^m \cdot AF_{Ar}(temp). \tag{3.5}$$

3.2.3 Voltage Acceleration Models

For products like light bars and heaters, increasing voltage is a simple way to accelerate degradation. Voltage is defined as the difference in electrical potential between two points, and physically it can be thought of as the amount of pressure behind an electrical current [29]. Rohner et al. [5] proposed the intrinsic voltage acceleration model to describe the evolution of the post breakdown conductivity. Wu and Sune [30] proposed three independent experimental methodologies applicable to the investigation of voltage acceleration models of time to breakdown for dielectric breakdown. Hu et al. [31] presented a voltage-acceleration lifetime model

to predict the product's lifetime distribution. And the most frequently used model for voltage is the inverse power relationship. Padgett et al. [32] applied the inverse power law model to carbon composite materials.

The inverse power relationship can be expressed as:

$$T(V) = \frac{T(V_0)}{AF(V)} = \left(\frac{V}{V_0}\right)^{\beta_1} \cdot T(V_0), \tag{3.6}$$

where $T(V)$ and $T(V_0)$ are failure time of the product under the general voltage V_0 and the operating voltage V; generally, we have $\beta_1 < 0$.

The acceleration factor is:

$$AF(V) = \frac{T(V_0)}{T(V)} = \left(\frac{V}{V_0}\right)^{-\beta_1}. \tag{3.7}$$

When $V > V_0$ and $\beta_1 < 0$, we have $AF(V) > 1$.

3.2.4 *Other Acceleration Models*

Generally, an acceleration model is suitable for a single degradation process. If the failure of a product involves simultaneous degradation processes, each process will require its own acceleration model. However, there're some accelerated tests which choose acceleration models with more than one accelerated variables. Escobar and Meeker [33] reviewed this kind of acceleration models well, including generalized Eyring relationship, temperature-voltage acceleration models, temperature-current density acceleration models and temperature-humidity acceleration models.

3.3 Degradation Modeling and Analysis

Acronym	D_f Failure Threshold
PDF Probability Density Function	ε Measurement Error
CDF Cumulative Distribution Function	n Sample Size
Symbol	m_i Measuring Times of Sample i
	Δ Sampling Frequency
B(•) Standard Brown Motion	D(•) Real Value of Performance
T Failure Time	Y(•) Measured Value of Performance
M$_i$ Degradation Model$i$	$\Gamma(\alpha)$ Gamma Function
f_T PDF of Failure Time$T$	$\Gamma(a, z)$ Incomplete Gamma Function
F_T CDF of Failure Time T	Φ(•) Standard Normal Distribution Function

One objection of reliability analysis is to make predictions about the product lifetime. When the product follows a soft failure, its lifetime can be defined as the first time that degradation reaches a specified level. In order to make connections between degradation data and the product lifetime, it's essential to establish a proper probability model to describe the behavior of the collected degradation data.

For degradation test, if all product units along with their test conditions and environment were identical, the lifetimes would be the same. However, in reality, there are some variables in degradation test. First, it is almost impossible that the product units are exactly the same, for instance, its initial degradation level can be different from each other. Second, we shouldn't ignore the difference between each unit in test conditions and environment. For example, environmental temperature often changes slightly over time, and this change may affect the degradation process of the units. Next, measurement errors of degradation can be reduced, but they cannot be totally avoided.

Two types of degradation model are commonly used in real applications, and they are general path models and stochastic models. When referring stochastic process models, this paper will discuss three frequently-used types: Wiener process, Gamma process and the inverse Gaussian process.

3.3.1 General Path Models

The fundamental notion under the general degradation path models is to limit the sample space of the degradation process and assume all sample functions admit the same functional form but with different parameters. In the degradation model, a unit's actual degradation path over time is denoted by $D(t)$, which is a monotone function. A soft failure occurs when $D(t)$ reaches a threshold value D_f. The degradation level of unit i at time t_j, obtained from measurements, can be expressed as:

$$y_{ij} = D_{ij} + \varepsilon_{ij} = D\left(t_j; \varphi, \theta_i\right) + \varepsilon_{ij}, i = 1, \ldots, n, j = 1, \ldots, m_i, \tag{3.8}$$

where ε_{ij} follows a normal distribution $N(0, \varepsilon^2)$, and it is measurement error for unit i at time t_j; φ is the vector of fixed-parameter which is identical for all units; θ_i is the vector of random parameter which vary from unit to unit; ε_{ij} and θ_i $(i = 1, \ldots, n, j = 1, \ldots, m_i)$ are independent of one another. If we know that the cumulative distribution function (CDF) of θ_i is $\Phi_\theta(\bullet)$, and $D(t)$ is a monotone increasing function, then the CDF of lifetime T can be expressed as:

$$F_T(t) = P\{T \le t\} = P\left\{D(t) \ge D_f\right\} = F_T\left(t; \varphi, \Phi_\theta(\cdot), D_f, D\right). \tag{3.9}$$

3.3.1.1 Two Basic Methods of Model Application

There are two basic methods to process the degradation data, which the general path model is used. One of them is degradation path fitting method, where a path model is built for the performance of product under test, such as the luminous flux of a LED bar. Zuo et al. [34] presented the general path model to fit parameters of the model as functions of time. The other is degradation amount distribution method. Sometimes, the degradation can be taken as a random process, that is to say the degradation level follows a random distribution class with time-dependent parameters. For example, assume that the degradation level $y(t)$ obtained from measurements obeys the normal distribution with mean value $\mu_y(t)$ and variance $\sigma_y^2(t)$, and the distribution pattern of $y(t)$ do not change with time. Then, if the actual degradation path $D(t)$ is a monotone increasing function, the reliability function can be described as:

$$R(t) = 1 - F(t) = 1 - P\left\{y(t) \geq D_f\right\} = \Phi\left(\frac{D_f - \mu_y(t)}{\sigma_y(t)}\right). \tag{3.10}$$

Thus, the core problem of this situation falls into setting up path models with mean value $\mu_y(t)$ and variance $\sigma_y^2(t)$. Wang et al. [35] proposed a lifetime prediction approach to predict degradation amount distribution of products using a composite time series modeling procedure based on degradation data.

3.3.1.2 Incorporation of Accelerated Models

This section will use a simple example to introduce the way of incorporation accelerated models into general path models. Assume that a product's degradation path under operating stress s_0 can be written as:

$$D\left(t; s_0\right) = \theta_0 + \theta_1 t + \varepsilon, \tag{3.11}$$

where θ_0 is the vector of fixed-parameter which is identical for all units; θ_1 is a vector of random parameter; and ε is measurement error which is independent of time t. The failure time of the product can be defined as:

$$T\left(s_0\right) = \frac{D_f - \theta_0}{\theta_1}. \tag{3.12}$$

Then according to accelerated model, the failure time of the product under stress level s can be defined as:

$$T\left(s\right) = \frac{T\left(s_0\right)}{AF\left(s\right)} \quad \frac{D_f - \theta_0}{\theta_1 \times AF\left(s\right)}, \tag{3.13}$$

where $AF(s)$ is the acceleration factor. Then, correspondingly, the product's degradation path under stress level s can be written as:

$$\begin{aligned} D(t;s) &= \theta_0 + \theta_1 \times AF(s) \times t + \varepsilon \\ &= \theta_0 + h(s) \times t + \varepsilon, \end{aligned} \tag{3.14}$$

where $h(s) = \theta_1 \times AF(s)$ is a function of stress s.

From (3.14), when incorporating accelerated models into general path models, we can regard some of the model parameters as a function of covariate s which is called link function. And the link function is related to the type of accelerated model. For instance, in the example above, if the covariate is temperature and the type of accelerated model is Arrhenius relationship, then the link function can be express as:

$$\begin{aligned} h(s) &= \theta_1 \times AF_{Ar}(s) = \theta_1 \times \frac{R_{Ar}(s)}{R_{Ar}(s_0)} \\ &= \frac{\theta_1}{R_{Ar}(s_0)} \times \gamma_0 \times \exp\left(\frac{-E_a}{k_b \times s}\right) \\ &= \beta_0 \times \exp\left(-\frac{\beta_1}{s}\right), \end{aligned} \tag{3.15}$$

where $\beta_0 = \frac{\theta_1}{R_{Ar}(s_0)} \times \gamma_0$ is a vector of random parameter; $\beta_1 = \frac{E_a}{k_b}$ is the vector of fixed-parameter.

3.3.2 Stochastic Processes Models

The general path model is a type of simplified model for degradation data, and this kind of model can be used when all units are tested under a particular homogenous environment. Ye and Xie [36] pointed out that the randomness in the unobserved environmental factors will convert to unexplained randomness in the observed degradation, and stochastic processes are a natural choice for modeling the randomness in degradation processes caused by inherent randomness and unexplained randomness due to environmental factors.

The Wiener process, the gamma process and the inverse Gaussian process are three common stochastic processes which have received lots of applications in degradation modeling. Tseng et al. in paper [37] used a Wiener process to model the light intensity of LED lamps. Joseph and Yu [38] used a Wiener process model for reliability improvement. Si et al. in [39] developed a Wiener-process-based degradation model with a recursive filter algorithm to estimate the RUL from the observed degradation data. Wang et al. [40] considered the gamma process with a time transformation and random effects for lifetime prediction. Iervolino et al. [41]

used a gamma distribution to model damages which is produced by earthquakes. Noortwijk [42] surveyed the application of gamma processes in maintenance. Wang and Xu [43] proposed an inverse Gaussian process model to fit laser data. Peng [44] proposed a degradation model based on an inverse normal-gamma mixture of an inverse Gaussian process. Ye et al. [45] discussed the accelerated degradation test planning when the underlying degradation follows the inverse Gaussian process.

3.3.2.1 The Wiener Process

In many practical cases, degradation path of product performance is the result of joint action by a large number of external environmental factors. From the central limit theorem, the increments $\Delta D(t) = D(t + \Delta t) - D(t)$ should approximately obey the normal distribution. And the increments of degradation path in Wiener process exactly obey normal distribution, so the Wiener process is commonly used in the field of reliability analysis.

1. Wiener Process with Drift Parameter

Before introduction of Wiener process with drift parameter, a standard Wiener process is given as follow first. $\{D(t), t \geq 0\}$ is known as Wiener process or Brownian motion, if

① $D(0) = 0$;
② $\{D(t), t \geq 0\}$ has stable independent increments, it is to say, the increments $\Delta D(t) = D(t + \Delta t) - D(t)$ are independent;
③ $\forall t > 0$, $D(t)$ follow a normal distribution $N(0, \sigma^2 t)$.

Let σ be 1, then $\{D(t), t \geq 0\}$ is called a standard Wiener process or standard Brownian motion process.

Here comes the one variable Wiener process with drift parameter. $\{D(t), t \geq 0\}$ is called a Wiener process with drift parameter μ and volatility parameter $\sigma > 0$ [36, 46, 47], if

① $D(0) = 0$;
② the increments $\Delta D(t) = D(t + \Delta t) - D(t)$ are independent;
③ $\forall t > 0$, $D(t)$ follow a normal distribution $N(\mu \Lambda(t), \sigma^2 \Lambda(t))$.

Where $\Lambda(t)$ is the scales time, and it is a continuous monotone increasing function. Let $B(\bullet)$ be a standard Brownian motion, according to the definition above, the degradation path in a Wiener process is often expressed as

$$D(t) = \mu \Lambda(t) + \sigma B(\Lambda(t)). \tag{3.16}$$

Because $D(t)$ obeys a normal distribution $N(\mu \Lambda(t), \sigma^2 \Lambda(t))$, the coefficient of variation is given as

$$CV\left(D(t)\right) = \frac{\sqrt{Var\left(D(t)\right)}}{E\left(D(t)\right)} = \frac{\sigma}{\mu\sqrt{\Lambda(t)}}. \tag{3.17}$$

And it is a reduction function of time, which shows that with the increase of time, sample degradation path of $D(t)$ gradually move closer to the mean curve.

One important property of Wiener process is that the increments $\Delta D(t) = D(t+\Delta t) - D(t)$ are independent, and they have a normal distribution with mean $\mu\Delta\Lambda(t) = \mu\Lambda(t+\Delta t) - \mu\Lambda(t)$ and variance $\sigma^2\Delta\Lambda(t) = \sigma^2\Lambda(t+\Delta t) - \sigma^2\Lambda(t)$. The PDF of $\Delta D(t)$ is

$$f(\Delta D(t); \Delta\Lambda(t), \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi\cdot\Delta\Lambda(t)}} \cdot exp\left\{-\frac{(\Delta D(t) - \mu\Delta\Lambda(t))^2}{2\sigma^2\Delta\Lambda(t)}\right\}. \tag{3.18}$$

The failure time T is defined as the first time that $D(t)$ reaches the threshold value D_f. In Wiener process, the failure time T follows a transformed inverse Gaussian distribution, that is $\Lambda(T) \sim IG\left(\frac{D_f}{\mu}, \frac{D_f^2}{\sigma^2}\right)$. Thus, we can induce the CDF of $\Lambda(T)$ is given by

$$\begin{aligned} F_{\Lambda(T)}(x) = \Phi_{nor}&\left(\sqrt{\frac{D_f^2}{\sigma^2 x}}\left(\frac{\mu x}{D_f} - 1\right)\right) \\ &+ \exp\left(\frac{2\mu D_f}{\sigma^2}\right)\cdot\Phi_{nor}\left(-\sqrt{\frac{D_f^2}{\sigma^2 x}}\left(\frac{\mu x}{D_f} + 1\right)\right). \end{aligned} \tag{3.19}$$

And $E\left(\Lambda(T)\right) = \frac{D_f}{\mu}$, $Var\left(\Lambda(T)\right) = \frac{D_f\sigma^2}{\mu^3}$.

2. Wiener Process with Measurement Errors

Although the Wiener process (3.16) can express inner randomness of the performance degradation process, it is difficult to show measurement error in degradation test. The existence of measurement error may due to imperfect measuring tool, and the randomness of external environmental factors may also affect the measured values of product performance. In fact, measurement error is very common, especially when the data are obtained in an indirect way. For example, in many cases, current probe and sensor are used to measure the product performance. As a result, when using Wiener process, measurement error is often taken into account.

Wiener process with measurement error can be expressed as

$$Y(t) = D(t) + \varepsilon, \tag{3.20}$$

where $Y(t)$ is measured value of product performance at time t; $D(t)$ represent the model (3.16); and ε is measurement error which is independent of time t. Li et al. [48] provided a Wiener process model for accelerated degradation analysis considering measurement errors.

3. Wiener Process with Random Effects

It is common to find differences among degradation path for product units from the same population. This kind of differences is very likely to be the result of unobservable random effects, such as the distribution of bacteria in the test. Based on this, it is necessary to discuss Wiener process with random effects.

Considering unobservable random effects, it is common to allow one or more model parameters being unit-specific when Wiener process applied. And a Wiener process will be set if a certain parametric distribution is given. For example, we can let drift parameter μ in model (3.16) be unit-specific and follow a normal distribution, while the volatility parameter $\sigma > 0$ be a constant.

Here are two common Wiener process with random effects.

Random Effects Model 1

$$D(t) = \beta t + \sigma B(t). \tag{3.21}$$

Where drift parameter β is a random variable, which describes the difference of individual degradation rate; volatility parameter σ is a constant for all product units. In particularly, the drift parameter β can be considered to follow a normal distribution $N\left(\mu_\beta, \sigma_\beta^2\right)$.

Random Effects Model 2

$$D(t) = \beta t + \sigma B(t). \tag{3.22}$$

Let $V = \sigma^2$, and β, V be both random variable. Because $V > 0$, $V^{-1} > 0$. Due to the flexibility of gamma distribution, V^{-1} can be assumed to obey a gamma distribution. Then V obeys an inverse gamma distribution, and its PDF can be expressed as

$$f(V; \alpha_1, \alpha_2) = \frac{{\alpha_2}^{\alpha_1}}{\Gamma(\alpha_1)} V^{-\alpha_1 - 1} e^{-\frac{\alpha_2}{V}}, \quad V > 0, \alpha_1 > 0, \alpha_2 > 0, \tag{3.23}$$

where α_1 is shape parameter and α_2 is rate parameter.

Furthermore, the drift parameter β can be considered to follow a normal distribution $N(v, \eta V)$ under the condition of a given V.

4. Incorporation of Accelerated Models

Environmental stress factors may affect performance degradation process, for instance, the rise of junction temperature can accelerate the degradation of LED bars. And common environmental factors include voltage, temperature, humidity, vibration, sample size and so on. These stress factors are known as covariates. When covariate is measurable, then an acceleration model can be used to integrate covariate into Wiener process.

The most common way of incorporating acceleration model into Wiener process is to regard some of the model parameters as a function of covariate, which is called link function. And the choice of function form should depend on the way acceleration factor influent model parameters. Tang et al. [49] regarded volatility parameter $\sigma > 0$ as a constant, and drift parameter μ as a linear function of covariate s, which can be expressed as $\mu(s) = \mu_0 + \mu_1 s$. Padgett and Tomlinson [50] regarded the drift parameter μ as a power function of covariate s, which can be expressed as $\mu(s) = \mu_0 s^{\mu_1}$.

3.3.2.2 The Gamma Process

In the process of accelerated degradation tests, the degradation path of product performance is not necessarily monotonous. For many products, their performance degradation is often irreversible, at this time the monotonicity of degradation path becomes a necessary condition. Compared with the Wiener process, the gamma process can be used when the degradation path is monotone. From another aspect, a lot of product failure is caused by the impact of the external random factors. These kinds of random factors tend to be very small, and their impact time follows a compound Poisson distribution. In this case, Gamma process is a good choice because it precisely obeys a compound Poisson distribution.

A process $\{D(t); t \geq 0\}$ is called a gamma process, if:

① the increments $\Delta D(t) = D(t + \Delta t) - D(t)$ are independent;
② the increments $\Delta D(t)$ follow a gamma distribution $Ga(\mu, \Delta\eta(t))$, where μ is the scale parameter, and the shape function $\eta(\bullet)$ is a given, monotone increasing function.

If $\eta(0) = 0$ and $D(0) = 0$, then $D(t)$ follows the distribution $Ga(\mu, \eta(t))$, and its PDF and CDF can be expressed as:

$$\text{PDF} : g(x; \mu, \eta) = \frac{\mu^{-\eta}}{\Gamma(\eta)} \cdot x^{\eta-1} \cdot \exp\left\{-\frac{x}{\eta}\right\}; \tag{3.24}$$

$$\text{CDF} : G(x; \mu, \eta) = \int_0^x g(x; \mu, \eta)\, dx. \tag{3.25}$$

The CDF of the failure time can be expressed as:

$$F_T(t) = P\{T \le t\} = P\left\{D(t) \ge D_f\right\} = 1 - G\left(D_f; \mu, \eta\right). \tag{3.26}$$

Similar to Wiener process, gamma process with covariates and gamma process with random effects can also be set up. Lawless and Crowder [51] regarded the scale parameter μ as a function of covariates when setting up gamma process with covariates, and assumed μ to be unit-specific which follows a gamma distribution when setting up gamma process with random effects.

In addition, the measurement error is also hard to avoid in the product degradation process. Generally, the measurement error ε obeys the white noise. If $D(t)$ represents the gamma process, ε represents the measurement error, then gamma process with measurement errors can be expressed as $Y(t) = D(\mathrm{t}) + \varepsilon$. Kallen and Noortwijk [52] took the measurement errors of degradation into account in the gamma process.

3.3.2.3 The Inverse Gaussian Process

There are many degradation paths that the Wiener process and the gamma process fail to model, Wang and Xu [43] introduced the inverse Gaussian process as an alternative. A process $\{D(t); t \ge 0\}$ is called an inverse Gaussian process, if

① the increments $\Delta D(t) = D(t + \Delta t) - D(t)$ are independent;
② the increments $\Delta D(t)$ follow a IG distribution $IG(\Delta\eta(t), \xi(\Delta\eta(t))^2)$, where ξ is the scale parameter, and the mean function $\eta(\bullet)$ is a nonnegative, monotone increasing function.

Let $\eta(0) = 0$ and $D(0) = 0$, then $D(t)$ denotes the distribution $IG(\eta(t), \xi\eta^2(t))$, and its PDF and CDF can be expressed as:

$$PDF: f_{IG}\left(x; \eta, \xi\eta^2\right) = \sqrt{\frac{\xi\eta^2}{2\pi x^3}} \cdot \exp\left(-\frac{\xi(x-\eta)^2}{2x}\right), x > 0; \tag{3.27}$$

$$\begin{aligned} \mathrm{CDF}: F_{IG}\left(x; \eta, \xi\eta^2\right) &= \int_0^{\mathrm{x}} f_{IG}\left(x; \eta, \xi\eta^2\right) dx \\ &= \Phi_{nor}\left(\sqrt{\frac{\xi}{x}}(x-\eta)\right) + \exp\left(2\xi\eta\right) \cdot \Phi_{nor}\left(-\sqrt{\frac{\xi}{x}}(x+\eta)\right). \end{aligned} \tag{3.28}$$

The CDF of the failure time can be expressed as:

$$\begin{aligned} F_T(t) &= P\{T \le t\} = P\left\{D(t) \ge D_f\right\} = 1 - F_{IG}\left(x; \eta, \xi\eta^2\right) \\ &= \Phi_{nor}\left(\sqrt{\tfrac{\xi}{D_f}}\left(\eta - D_f\right)\right) - \exp\left(2\xi\eta\right) \cdot \Phi_{nor}\left(-\sqrt{\tfrac{\xi}{D_f}}\left(\eta + D_f\right)\right). \end{aligned} \tag{3.29}$$

And when $\xi \bullet \eta(t)$ is large enough, especially when t is large enough, $D(t)$ approximately obeys a normal distribution with mean $\eta(t)$ and variance $\eta(t)/\xi$. At this time, the distribution function of failure time T can be expressed as

$$F_T(t) \approx 1 - \Phi\left(\frac{l - \eta(t)}{\sqrt{\frac{\eta(t)}{\xi}}}\right) = \Phi\left(\sqrt{\xi \cdot \eta(t)} - \frac{l\sqrt{\xi}}{\sqrt{\eta(t)}}\right). \tag{3.30}$$

In Wiener process, the failure time T follows a transformed inverse Gaussian distribution, that is $\Lambda(T) \sim IG\left(\frac{D_f}{\mu}, \frac{D_f^2}{\sigma^2}\right)$. If there are a series of threshold l indexed by t, we can gain corresponding transformed failure times $\Lambda T(t)$. Then $\Lambda T(t)$ follows an IG process $\Lambda T(t) \sim IG(\gamma l(t), \lambda l^2(t))$, where $\gamma = 1/\mu$ and $\lambda = 1/\sigma^2$. As a result, the IG process is very flexible in setting up IG process with covariates and IG process with random effects. That is to say, when it exists a way to integrate covariates and random effects into Wiener process, there is a corresponding way to integrate covariates and random effects into IG process.

3.3.3 Estimation of Model Parameters

In the field of reliability analysis, there are three most commonly used methods of parameter estimation: maximum likelihood method, the least squares approach and Bayesian methods.

Maximum likelihood method is one of the most widely-used methods in reliability analysis. There're two primary reasons for using maximum likelihood method [53]. First, there are hardly any statistical problems where the maximum likelihood method can't be used. Second, the maximum likelihood method is known to be optimum in large samples under mild conditions. The principle of maximum likelihood method is actually choosing a parameter which can make test result to have the greatest probability. Meeker and Escobar [29] proposed the parametric maximum likelihood in Chap.11. Hu et al. [54] gave a log-likelihood function of parameters from the N test units. Escobar et al. [3] used likelihood-based methods to infer both the degradation and the lifetime distributions.

The least squares approach is a standard approach in regression analysis to the approximate solution of overdetermined systems, and the least squares problems include linear least squares and non-linear least squares. It also has wide application in reliability analysis. Lu and Meeker [55] provided a least-squares-based two-stage method to estimate a time-to-failure distribution using degradation data. Wu and Shao [56] used the linear least squares methods for degradation analysis.

Bayesian methods are closely related to the likelihood methods. Ignoring the philosophical differences in the theory of statistical inference, Bayesian methods can be viewed as an extension of likelihood-based methods, where one com-

bines prior information, in the form of a prior probability distribution, with the likelihood [53]. In the Bayesian estimation, the parameter to be estimated is often regarded as random variable which obeys a prior distribution. Robinson and Crowder [57] explored a Bayesian approach for a growth curve degradation model with repeated measures. Wakefield et al. [58] proposed a Bayesian EM algorithm to estimate parameters of pharmacokinetics. Peng et al. [59] conducted a Bayesian analysis of inverse Gaussian process models for degradation modeling and inference.

3.3.4 Lifetime Prediction

In degradation analysis, a path model $D(t; \varphi, \theta)$ and a threshold value D_f is usually assumed, where φ is the vector of fixed-parameter and θ is the vector of random parameter. Usually, the inference on lifetime distribution can be induced from the degradation model parameters, see equation:

$$F_T(t) = P(T \leq t) = P\left(D(t; \varphi, \theta) \geq D_f\right). \tag{3.31}$$

The Eq. (3.31) shows that the lifetime distribution depends on the random parameter θ. In some simple situations, if the functional form of the distribution of random parameter θ is assumed, we can induce the lifetime distribution $F_T(t)$ with a closed-form expression.

However, for some complex situation, it is not an easy task to obtain the lifetime cumulative distribution function (CDF) $F_T(t)$ with a closed-form expression. And when there is more than one random parameter in the model, the situation will become more complicated. Thus, other methods of evaluating $F_T(t)$ are needed.

Approximate accelerated degradation analysis is an alternative method of analyzing accelerated degradation data. Generally, there are two steps in conducting the approximate method. In the first step, we should predict the failure time for each sample path, the obtained n failure times $T_1, \ldots, T_n$ are called pseudo failure times. In the second step, the pseudo failure times $T_1, \ldots, T_n$ can be analyzed using the methods of traditional reliability analysis to estimate $F_T(t)$.

Chen and Zheng [60] used the degradation data to derive predictive intervals of individual lifetimes first, then an imputation algorithm is invoked to obtain the estimate of the lifetime distribution. Lu and Meeker [55] used a non-linear mixed-effects model and developed methods based on Monte Carlo simulation to obtain point estimates and confidence intervals for reliability assessment. Bae et al. [61] explored the link between degradation model and the resulting lifetime model, and found out that seemingly innocuous assumptions of the degradation path create surprising restrictions on the lifetime distribution.

3.4 Initial Degradation Levels

Initial degradation level is product's degradation performance at time zero. In degradation analysis, most studies regarded initial degradation level as zero, and focused on the degradation increments when stochastic models are used. For example, Hu et al. [54] assumed that the degradation path under standardized stress level x_i follows a Wiener process denoted as $W_i(t) = \eta_i t + \sigma B(t), t \geq 0$.

However, the initial degradation is often nonzero and it is random. Bae and Kvam [62] proposed a degradation model which reflects the variation in initial luminosity of VFDs, and further analysis revealed that the initial performance of VFDs is a random coefficient. Furthermore, plenty of studies showed that the initial degradation always correlates with the degradation rate.

3.4.1 Motivating Examples

Here we use the example of sliding metal wear application in Chap. 21 of [29] to support the view intuitionally. In the experiment, the wear resistance is tested among metal pieces of different weight. Measurements of the scar width are taken at different points in time, which are shown in Fig. 3.1. And this example shows that the low initial degradation level of scar width indicates a lower degradation rate.

There also exist some examples that the initial degradation is negatively correlated with the degradation rate. Weaver et al. [63] introduced an example involving inkjet printer heads. Figure 3.2 shows the scatterplot of print head migration data, and from the figure we can see that when the initial degradation level is low, the corresponding degradation rate can be high. That implies the degradation rate can be roughly estimated using the initial performance data.

3.4.2 Mixed-Effect General Path Model

In order to make use of the initial degradation as well as its correlation with a degradation rate when doing acceleration degradation test, a mixed-effect general path model is usually used. For unit $i (i = 1, 2, \ldots, n)$ the mixed-effect degradation path can be described as following form:

$$\boldsymbol{y_i} = \boldsymbol{X_i}\boldsymbol{\varphi} + \boldsymbol{Z_i}\boldsymbol{\Theta_i} + \boldsymbol{\varepsilon_i}, i = 1, \ldots, n. \tag{3.32}$$

Where $\boldsymbol{y_i} = (y_{i1}, \ldots, y_{im_i})'$ denotes the vector of the degradation measurement data; $\varphi = (\varphi_0, \varphi_1, \ldots, \varphi_{p-1})'$ is the coefficient corresponding to the fixed-effect covariates $\boldsymbol{X_i}$; $\boldsymbol{X_i}$ is a $m_i \times p$ matrix with each row denoted by $\boldsymbol{x_{ij}} = (1, t_{ij}, x_{ij2}, \ldots, x_{ij,p-1})$, $j = 1, \ldots, m_i$; $\boldsymbol{\Theta_i} = (\theta_{i1}, \ldots, \theta_{iq})'$ is the coefficients corresponding to the

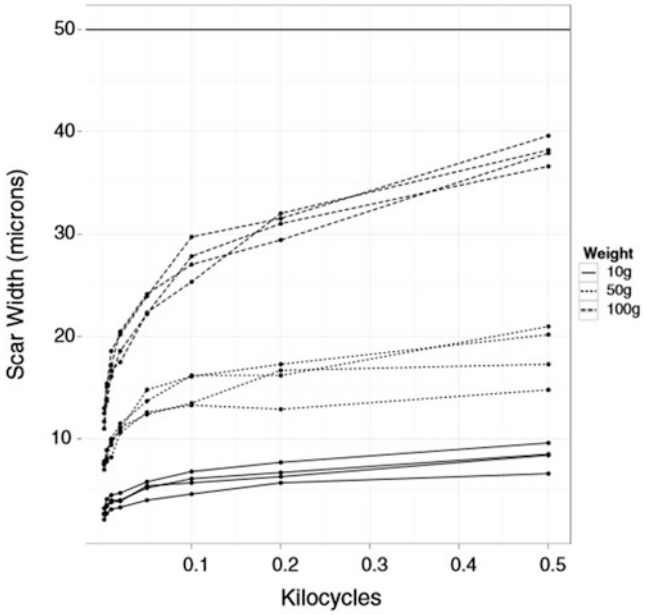

Fig. 3.1 Scar width data of metal alloy

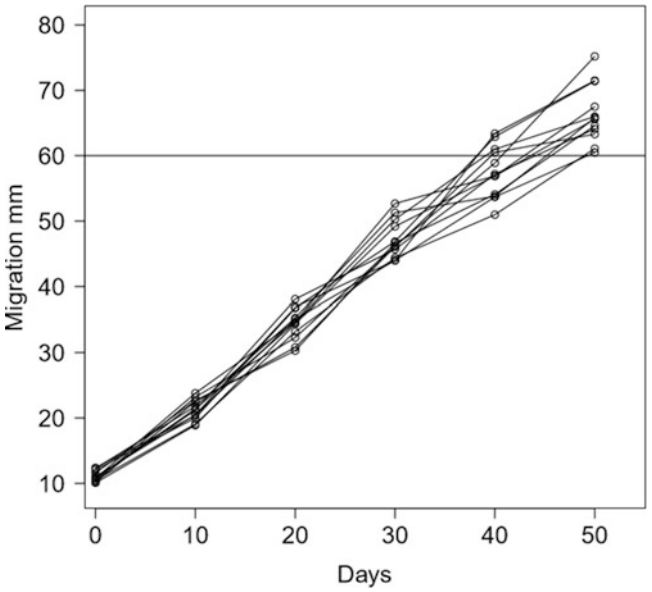

Fig. 3.2 Scatterplot of print head migration data

random-effect covariates $\boldsymbol{Z_i}$; $\boldsymbol{Z_i}(m_i \times q)$ is a sub-matrix of $\boldsymbol{X_i}$, and each row can be denoted as $\boldsymbol{z_{ij}} = (z_{ij1}, \ldots, z_{ijq})$; and $\boldsymbol{\varepsilon_i} = (\varepsilon_{i1}, \ldots, \varepsilon_{im_i})'$ denotes the measurement errors.

Without loss of generality, let

$$\boldsymbol{x_{ij}} = \left(1, t_{ij}, x_{ij1}, x_{ij2}t_{ij}\right), j = 1, \ldots, m_i; \tag{3.33}$$

$$\boldsymbol{\varphi} = (0, 0, \varphi_1, \varphi_2)'; \tag{3.34}$$

$$\boldsymbol{z_{ij}} = \left(1, t_{ij}\right), j = 1, \ldots, m_i; \tag{3.35}$$

$$\boldsymbol{\Theta_i} = (\theta_0, \theta_1)'. \tag{3.36}$$

Then the model (3.32) can be simplified as:

$$\mathrm{y} = x_1\phi_1 + x_2\phi_2 t + \theta_0 + \theta_1 t + \varepsilon. \tag{3.37}$$

Where the term $x_1\varphi_1$ denotes how the initial degradation levels change with accelerating variable x_1; $x_2\varphi_2$ denotes how the degradation rate changes with accelerating variable x_2; θ_0 and θ_1 describe the unit-to-unit variability in the initial performance and degradation rate; $\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$ denotes measurement errors.

In many cases, the random in the initial degradation levels is often modelled by a normal distribution, and it often has a strong correlation with the degradation data. As a result, we usually assume that parameter θ_0 and θ_1 follow a bivariate normal distribution, which can be described as $(\theta_0, \theta_1)' \sim BVN(\boldsymbol{\beta}, \boldsymbol{V})$. Where $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ is the mean vector, and $\boldsymbol{V}$ is the covariance matrix expressed as

$$\boldsymbol{V} = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \tag{3.38}$$

Yuan and Pandey [64] used the model (3.37) to describe the flow accelerated corrosion in pipes, and found the correlation coefficient in the wall thinning data of carbon steel pipes is as high as −0.91. Lu et al. [65] used a transformed model $y = \theta_0 + \theta_1 \log(t) + \varepsilon$ to describe the HCI degradation data, and found the correlation coefficient between θ_0 and θ_1 is $\rho = -0.62$. Weaver and Meeker [66] introduced an example of carbon-file resistors using model (3.37), and in the example $\rho = 0.628$.

3.5 Discussions on Future Study

There are two types of general degradation model which have been widely used, general path models and stochastic models. Above have already introduced them in details, including the general path model, the Wiener process, the Gamma process and the inverse Gaussian process. However, the degradation of a product is often very slow, so it is common to use harsh test conditions such as high temperature and usage rate. For that reason, three types of acceleration models are presented, and they are the usage rate acceleration models, the temperature acceleration models and the voltage acceleration models. This paper also discussed the accelerated degradation test (ADT) analysis, including the estimation methods for the ADT model parameters, and the methods of evaluation of failure time distribution.

As we come across more accelerated degradation problems, we will find that above knowledge is far from enough. For example, the initial status of a product at time zero is usually not quite consistent, and it always correlates with the degradation rate. As a result, we cannot ignore the initial degradation when we perform an acceleration degradation test, and this refers to the strategic allocation of test units which can be further discussed.

References

1. Nelson W (1981) Analysis of performance-degradation data from accelerated tests. IEEE Trans Reliab 30(2):149–155
2. Wang FK, Chu TP (2012) Lifetime predictions of led-based light bars by accelerated degradation test. Microelectron Reliab 52(7):1332–1336
3. Escobar LA, Meeker WQ, Kugler DL, Kramer LL (2015) Accelerated destructive degradation tests: data, models and analysis. Math Stat Methods Reliab:319–337
4. Shiau JJH, Lin HH (1999) Analyzing accelerated degradation data by nonparametric regression. IEEE Trans Reliab 48(2):149–158
5. Rohner M, Kerber A, Kerber M (2006) Voltage acceleration of TBD and its correlation to post breakdown conductivity of N- and P-Channel MOSFETs. IEEE Int Reliab Phys Symp Proc:76–81
6. Zhang J, Li W, Cheng G, Chen X, Wu H, Shen MHH (2014) Life prediction of OLED for constant-stress accelerated degradation tests using luminance decaying model. J Lumin 154:491–495
7. Guan Q, Tang Y, Xu A (2015) Objective Bayesian analysis accelerated degradation test based on Wiener process models. Appl Math Model 40(4):2743–2755
8. Pan Z, Balakrishnan N, Sun Q (2011) Bivariate constant-stress accelerated degradation model and inference. Commun Stat Simul Comput 40(2):259–269
9. Wan F, Zhang C, Tan Y, Chen X (2014) Data analysis and reliability estimation of step-down stress accelerated degradation test based on Wiener process. Prog Sys Health Manage Conf:41–45
10. Pan Z, Sun Q (2014) optimal design for step-stress accelerated degradation test with multiple performance characteristics based on Gamma processes. Commun Stat Simul Comput 43(2):298–314
11. Ge Z, Li XY, Jiang TM, Huang TT (2011) optimal design for step-stress accelerated degradation testing based on D-optimality. Reliab Maint Symp:1–6

12. Guo CS (2011) Online degradation model based on process-stress accelerated test. Acta Phys Sin 60:128501
13. Peng CY, Tseng ST (2010) Progressive-stress accelerated degradation test for highly-reliable products. IEEE Trans Reliab 59(1):30–37
14. Boulanger M, Escobar LA (1994) Experimental design for a class of accelerated degradation tests. Technometrics 36(3):260–272
15. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes. Lifetime Data Anal 11(4):511–527
16. Tang S, Guo X, Yu C, Xue H, Zhou Z (2014) Accelerated degradation tests modeling based on the nonlinear Wiener process with random effects. Math Probl Eng 2014(2):1–11
17. Sang BK, Kim SH (2015) Estimation of elevator wire life using accelerated degradation model. J Korean Soc Qual Manag 43(3):409–420
18. Liu HC (2014) Gauss-power law accelerated degradation model about reliability assessment of electronic products. Journal of Guiyang University
19. Yang Z, Chen YX, Li YF, Zio E, Kang R (2013) Smart electricity meter reliability prediction based on accelerated degradation testing and modeling. Int J Elect Power Energy Syst 56(3):209–219
20. Hayes KC, Wolfe DL, Trujillo SA, Burkell JA (2010) On the interaction of disability and aging: accelerated degradation models and their influence on projections of future care needs and costs for personal injury litigation. Disabil Rehabil 32(5):424–428
21. Yang G, Zaghati Z (2006) Accelerated life tests at higher usage rates: a case study. Reliability and maintainability symposium, 313–317
22. Yellamati D, Arthur E, James S, Morris G, Heydt T, Graf E (2013) Predictive reliability models for variable frequency drives based on application profiles. 2013 Proceedings Reliability and Maintainability Symposium (RAMS), 1–6
23. Hwang DH, Park JW, Jung JH (2011) A study on the lifetime comparison for electric double layer capacitors using accelerated degradation test. International Conference on Quality, Reliability, Risk, Maintenance and Safety Engineering, 302–307
24. Shen CN, Xu JJ, Chao MC (2014) A study for the relationship between drive level and the activation energy in arrhenius accelerated aging model for small size quartz resonators. 2014 IEEE International Frequency Control Symposium (FCS), 1–3
25. Cooper MS (2005) Investigation of arrhenius acceleration factor for integrated circuit early life failure region with several failure mechanisms. IEEE Trans Compon Packag Technol 28(3):561–563
26. Zhou J, Yao J, Song Y, Hu HH (2014) The step-down-stress accelerated storage testing evaluation methods of small sample electronic products based on arrhenius model. 2014 10th international conference on reliability, maintainability and safety (ICRMS), 908–912
27. Dai HF, Zhang XL, Gu WJ, Wen XZ, Sun ZC (2013) A semi-empirical capacity degradation model of ev li-ion batteries based on eyring equation. IEEE Vehicle Power and Propulsion Conference (VPPC), 1–5
28. Endicott H, Hatch B, Sohmer R (1965) Application of the eyring model to capacitor aging data. IEEE Trans Compon Parts 12(1):34–41
29. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data
30. Wu EY, Sune J (2009) On voltage acceleration models of time to breakdown-part I: experimental and analysis methodologies. IEEE Trans Electron Devices 56(7):1433–1441
31. Hu LC, Kang AC, Wu TY, Shih JR, Lin YF, Wu K, King YC (2006) Efficient low-temperature data retention lifetime prediction for split-gate flash memories using a voltage acceleration methodology. IEEE Trans Device Mater Reliab 6(4):528–533
32. Padgett WJ, Durham SD, Mason AM (1995) Weibull analysis of the strength of carbon fibers using linear and power law models for the length effect. J Compos Mater 29(14):1873–1884
33. Escobar LA, Meeker WQ (2007) A review of accelerated test models. Stat Sci 21(4):552–577
34. Zuo MJ, Renyan J, Yam R (1999) Approaches for reliability modeling of continuous-state devices. IEEE Trans Reliab 48(1):9–18

35. Wang L, Zhang H, Xue H (2012) Life prediction based on degradation amount distribution using composite time series analysis. Ieri Procedia 1(9):217–224
36. Ye ZS, Xie M (2014) Stochastic modelling and analysis of degradation for highly reliable products. Appl Stoch Model Bus Ind 31(1):16–32
37. Tseng ST, Tang J, Ku IH (2003) Determination of burn-in parameters and residual life for highly reliable products. Nav Res Logist 50(1):1–14
38. Joseph VR, Yu IT (2006) Reliability improvement experiments with degradation data. IEEE Trans Reliab 55(1):149–157
39. Si XS, Wang W, Hu CH, Chen MY, Zhou DH (2013) A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. Mech Syst Signal Process 35(1–2):219–237
40. Wang H, Xu T, Mi Q (2015) Lifetime prediction based on gamma processes from accelerated degradation data. Chin J Aeronaut 28(1):172–179
41. Iervolino I, Giorgio M, Chioccarelli E (2013) Gamma degradation models for earthquake-resistant structures. Struct Saf 45(45):48–58
42. Noortwijk JMV (2009) A survey of the application of gamma processes in maintenance. Reliab Eng Syst Saf 94(1):2–21
43. Wang X, Xu D (2010) An inverse Gaussian process model for degradation data. Technometrics 52(2):188–197
44. Peng CY (2015) Inverse Gaussian processes with random effects and explanatory variables for degradation data. Technometrics 57(1):100–111
45. Ye ZS, Chen LP, Tang LC, Xie M (2014) Accelerated degradation test planning using the inverse Gaussian process. IEEE Trans Reliab 63(63):750–763
46. Wang X (2010) Wiener processes with random effects for degradation data. J Multivar Anal 101(2):340–351
47. Ye ZS, Wang Y, Tsui KL, Pecht M (2013) Degradation data analysis using wiener processes with measurement errors. IEEE Trans Reliab 62(4):772–780
48. Li J, Wang Z, Liu X, Zhang Y, Fu H, Liu C (2016) A Wiener process model for accelerated degradation analysis considering measurement errors. Microelectron Reliab 65:8–15
49. Tang LC, Yang GY, Xie M (2004) Planning of step-stress accelerated degradation test, Reliability and Maintainability Annual Symposium, 287–292
50. Padgett WJ, Tomlinson MA (2004) Inference from accelerated degradation and failure data based on Gaussian process models. Lifetime Data Anal 10(2):191–206
51. Lawless J, Crowder M (2004) Covariates and random effects in a Gamma process model with application to degradation and failure. Lifetime Data Anal 10(3):213–227
52. Kallen MJ, Noortwijk JMV (2005) Optimal maintenance decisions under imperfect inspection. Reliab Eng Syst Saf 90(2–3):177–185
53. Meeker WQ (2009) Trends in the statistical assessment of reliability. Adv Degrad Model:3–16
54. Hu CH, Lee MY, Tang J (2014) Optimum step-stress accelerated degradation test for Wiener degradation process under constraints. Eur J Oper Res 241(2):412–421
55. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35(2):161–174
56. Wu SJ, Shao J (1999) Reliability analysis using the least squares method in nonlinear mixed-effect degradation model. Stat Sin 9(3):855–877
57. Robinson ME, Crowder MJ (2001) Bayesian methods for a growth-curve degradation model with repeated measures. Lifetime Data Anal 6(4):357–374
58. Wakefield JC, Smith AFM, Racine-Poon A, Gelfand AE (1994) Bayesian analysis of linear and nonlinear population models using the Gibbs sampler. J R Stat Soc 43(1):201–221
59. Peng W, Li YF, Yang YJ, Huang HZ, Zuo MJ (2014) Inverse Gaussian process models for degradation analysis: a Bayesian perspective. Reliab Eng Syst Saf 130(1):175–189
60. Chen Z, Zheng S (2005) Lifetime distribution based degradation analysis. IEEE Trans Reliab 54(1):3–10
61. Bae SJ, Kuo W, Kvam PH (2007) Degradation models and implied lifetime distributions. Reliab Eng Syst Saf 92(5):601–608

62. Bae SJ, Kvam PH (2012) A nonlinear random-coefficients model for degradation testing. Technometrics 46(4):460–469
63. Weaver BP et al (2011) Methods for planning repeated measures degradation studies. Technometrics 55(2):122–134
64. Yuan XX, Pandey MD (2009) A nonlinear mixed-effects model for degradation data obtained from in-service inspections. Reliab Eng Syst Saf 94(2):509–519
65. Lu J, Park J, Yang Q (1997) Statistical inference of a time-to-failure distribution derived from linear degradation data. Technometrics 39(4):391–400
66. Weaver BP, Meeker WQ (2014) Methods for planning repeated measures accelerated degradation tests. Appl Stoch Model Bus Ind 30(6):658–671

Chapter 4
On Some Shock Models with Poisson and Generalized Poisson Shock Processes

Ji Hwan Cha and Maxim Finkelstein

Abstract We consider systems subject to shocks that follow the generalized Polya process (GPP), which has been recently introduced and characterized in the literature. Distinct from the nonhomogeneous Poisson process that has been widely used in applications, the important feature of this process is the dependence of its future behaviour on the number of previous events (shocks). We consider the delayed events model and the corresponding shot noise process governed by the GPP. We also present some results on the preventive maintenance for systems with failure/repair times following the GPP.

Keywords Generalized Polya process • Extreme shock model • Imperfect repair • Preventive maintenance

4.1 Introduction

The goal of this paper is to present a new useful tool for reliability analyses of technical systems. The most popular point processes that are used in reliability analysis are the Poisson process (homogeneous and non-homogeneous) and the renewal process. In reliability, renewal processes are mostly applied in repair and maintenance modelling, whereas Poisson processes are effective when, e.g., modelling an effect of environment in the form of shocks and in other applications. It is well known that the Poisson process possesses the property of independent increments, which leads to the corresponding Markov property. Thus the past, or the history that describes operation of technical systems cannot influence the future reliability characteristics, which is not the case in numerous reliability application. The history of the renewal process (time since the last renewal) is useful in models of

J.H. Cha
Department of Statistics, Ewha Womans University, Seoul, 120-750, Republic of Korea
e-mail: jhcha@ewha.ac.kr

M. Finkelstein (✉)
Department of Mathematical Statistics, University of the Free State, 339, Bloemfontein, 9300, South Africa
e-mail: FinkelM@ufs.ac.za

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_4

repair and maintenance, however, considering even this simplest history can create certain difficulties in the corresponding mathematical description.

The generalized Polya process that was recently introduced and fully characterized in [6], on one hand allows for the 'richer' history and on the other hand, presents tractable solutions for probabilities of interest that can be applied in practice. We will present a short overview of some of the developed results on the GPP modelling in reliability and discuss some of the recent findings in this direction. Hopefully, it will bring this new modelling tool into attention of the wider audience of specialists in stochastic modeling.

In various engineering applications, by a shock we usually mean a 'point' potentially harmful event (see, e.g., [14] and references therein). Each shock from a stochastic process of shocks can be fatal (failure) or just result in the decrease in operational characteristics of a system. As the third option it can be also harmless thus showing a resistance of an object to this event. An example of a shock process can be voltage peaks over a threshold in an electrical system, earthquakes effecting the infra-structure of a large territory, insurance claims, etc.

Various shock models have been intensively studied in the literature (see, e.g., [7, 12, 14, 20, 22, 23] and references therein). In reliability studies, the most popular model is, probably, the so-called, extreme shock model, where each shock can result in a system failure with the specified probability and a system survives it with the complementary probability. It should be noted that survival probabilities of systems subject to shock processes for extreme shock model can be obtained explicitly only for the Poisson process of shocks (see, e.g., [9]). Even for the renewal processes of shocks, everything becomes more cumbersome and asymptotic or approximate methods should be used for the corresponding calculations.

Assume for simplicity that shocks constitute the only cause of failure of a system. In applications, we are usually interested in the probability of survival in $[0, t)$. Denote this probability by $P(t)$. The simplest model is when an item is subject to the homogeneous Poisson process (HPP) of shocks with the constant rate r; it survives each shock with probability q and it fails with the complementary probability $p = 1 - q$ (extreme shock model). It is well known that the probability of survival in $[0, t)$ in this case (an item has survived all shocks) is

$$P(t) = \exp\{-rt\} \sum_{i=0}^{\infty} \frac{(rt)^i}{i!} q^i = \exp\{-prt\}, \tag{4.1}$$

whereas for the nonhomogeneous Poisson process (NHPP) with rate $r(t)$ and time-dependent $p(t), q(t)$, this expression turns to [4, 13].

$$P(t) = \exp\left\{-\int_0^t p(u)r(u)du\right\}. \tag{4.2}$$

Obviously, when $r(t) = r$ and $p(t) = p$, (4.2) reduces to (4.1). The GPP process, as will be shown in the next sections, presents a useful and practically important generalization to (4.2) and to other characteristics that describe performance of technical systems subject to shocks.

As was mentioned, we assume for simplicity that shocks constitute the only failure mode for the affected systems. All results can be trivially generalized to the case when there are other independent failure modes. However, the dependence between the process of shocks and other failure mechanisms can create numerous interesting problems to be addressed in the future research.

4.2 Definition of the GPP

A new counting process, called the 'Generalized Polya Process' (GPP) has been recently described and studied in detail in [6]. Its definition to follow is based on the notion of stochastic intensity, therefore, let us first briefly discuss this notion in a way suitable for further presentation.

Let $\{N(t),\ t \geq 0\}$ be a point process, where $N(t)$ denotes the number of events (points) in $[0, t]$. Let $H_{t-} \equiv \{N(u),\ 0 \leq u < t\}$be the history of this process in $[0, t)$, i.e., the set of all point events in $[0, t)$. The history H_{t-}can be defined by the number of events in $[0, t)$ denoted by$N(t-)$ and the sequential arrival times of events $T_0 \equiv 0 \leq T_1 \leq T_2 \leq \ldots \leq T_{N(t-)} < t$. It is well known that the rate of the orderly point process is defined by the following expectation

$$r(t) = \lim_{\Delta t \to 0} \frac{E[N(t, t + \Delta t)]}{\Delta t},$$

where $N(t_1, t_2), t_1 < t_2$, represents the number of events in $[t_1, t_2)$. However, the rate does not fully characterize the point processes and the appropriate characterization should be employed. Point processes can be conveniently and mathematically described by using the concept of stochastic intensity λ_t, $t \geq 0$ [2]. This stochastic process is often also called the *intensity process*. For a process $\{N(t),\ t \geq 0\}$, the stochastic intensity is defined as the following limit (see, e.g., [14]):

$$\lambda_t = \lim_{\Delta t \to 0} \frac{\Pr[N(t, t + \Delta t) = 1 | H_{t-}]}{\Delta t} = \lim_{\Delta t \to 0} \frac{E[N(t, t + \Delta t) | H_{t-}]}{\Delta t}. \tag{4.3}$$

Obviously, for the NHPP with rate $r(t)$, the stochastic intensity is deterministic and equal to $r(t)$. Specifically, for HPP $\lambda_t = r,\ t \geq 0$. Note that the stochastic intensity for the renewal process is

$$\lambda_t = \lambda_F(t - T_{N(t)}),$$

where $\lambda_F(t)$ is the failure rate of the governing distribution for the renewal process and $T_{N(t)}$ is the notation for the time of the last before t renewal.

The formal definition of the GPP, via its stochastic intensity is given in [6] as

Definition 1 (Generalized Polya Process (GPP)) A counting process $\{N(t),\ t \geq 0\}$ is called the Generalized Polya Process (GPP) with the set of parameters $(\lambda(t), \alpha), \alpha \geq 0, \beta > 0$, if

(i) $N(0) = 0$;
(ii) $\lambda_t = (\alpha N(t-) + \beta)\lambda(t)$.

Thus, when $\alpha = 0$ the GPP reduces to the NHPP with rate $r(t) = \beta\lambda(t)$ and, accordingly, the GPP can be understood as a generalized version of the NHPP. A similar model has been recently studied in [1] (see also [3, 17]). However, the focus of these papers was different, mostly considering the corresponding issues of statistical inference and related frailty modeling, whereas here we will mostly concentrate on different results with respect to shock modeling and related aspects. It follows from the definition that the increments of this process are dependent, as probability of an event occurrence in the next infinitesimal interval of time depends on the number of the previous events.

The NHPP, due to its simplicity and possibility to derive the probabilities of interest in the explicit form, is the most popular and convenient point process in reliability applications, specifically, in shocks modeling. However, it possesses the property of independent increments (and therefore, it is a Markov process), whereas in real life the probabilities of events to follow most often depend on history. The GPP as the simplest point process with history, can effectively deal with this problem, which makes it the useful tool in many practical reliability applications. Most of the applications to be considered in this paper will be related to shock modeling, but the following introductory meaningful example deals with the model of imperfect repair.

Example 1 Imperfect repair. Consider the specific GPP process $(\lambda(t) = \lambda)$ of *imperfect repairs* defined by the following stochastic intensity:

$$\lambda_t = (\alpha N(t-) + 1)\lambda. \tag{4.4}$$

The time to the first failure (and instantaneous repair) is described by the Cdf $1 - \exp\{-\lambda t\}$. The first imperfect repair increases the initial failure rate from λ to $\lambda(1 + \alpha)$ and that is why it is called 'imperfect'. Therefore, the Cdf of the duration of the second cycle is $1 - \exp\{-\lambda(1 + \alpha)t\}$. On the third cycle, the failure rate is $\lambda(1 + 2\alpha)$ and the Cdf is $1 - \exp\{-\lambda(1 + 2\alpha)t\}$, etc. Thus it increases the failure rate as compared with that on the previous cycle and can be considered as a useful model for imperfect repair. Note that, we define imperfect repair as the repair when each consecutive cycle is stochastically smaller (in a suitable stochastic sense) then the previous one.

It is shown in [6] and [1] that probabilities of occurrence of n events in $[0, t)$ for the GPP can be described by the corresponding negative binomial distribution:

$$P(N(t) = n) = \frac{\Gamma(\beta/\alpha + n)}{\Gamma(\beta/\alpha)n!} (1 - \exp\{-\alpha\Lambda(t)\})^n (\exp\{-\alpha\Lambda(t)\})^{\frac{\beta}{\alpha}},$$
$$n = 0, 1, 2, \ldots \tag{4.5}$$

where $\Lambda(t) \equiv \int_0^t \lambda(u)du$.

See also [6] for relationships for other probabilities of interest (e.g., $P(N(t, t + h) = n)$. It immediately follows from (4.5) that

$$E[N(t)] = \sum_1^n nP(N(t) = n) = \frac{\beta}{\alpha} (\exp\{\alpha\Lambda(t)\} - 1). \tag{4.6}$$

Therefore, the rate of the GPP can be obtained as the corresponding derivative

$$r(t) = \frac{d}{dt}E[N(t)] = \beta\lambda(t) \exp\{\alpha\Lambda(t)\}. \tag{4.7}$$

Thus, for example, for the constant baseline function, $\lambda(t) = \lambda$, the rate of the process of shocks is *exponentially* increasing which reflects the cumulative effect of the previous events on the probability of occurrence of an event at the current instant of time.

4.3 Extreme Shock Model

Let our system be subject to the GPP process of external shocks and assume, for simplicity, that shocks constitute the only cause of its failure. Consider the corresponding extreme shocks model when a system survives each shock with probability $q(t)$ and fails with the complementary probability $p(t) = 1 - q(t)$. Denote by T the time to failure of a system. The extreme shock model with the NHPP of shocks (4.2) was generalized in [10] to the following result ($\beta = 1$, for convenience).

The survival function for an object exposed to the GPP of shocks in the described extreme shock model is

$$P(T > t) = \frac{1}{\left(1 + \int_0^t \alpha p(v)\lambda(v) \exp\{\alpha\Lambda(v)\}dv\right)^{\frac{1}{\alpha}}}, \tag{4.8}$$

whereas the corresponding failure rate is

$$\lambda_S(t) = -\frac{dP(T>t)}{P(T>t)} = \frac{p(t)\lambda(t)\exp\{\alpha\Lambda(t)\}}{\left(1+\int\limits_0^t \alpha p(v)\lambda(v)\exp\{\alpha\Lambda(v)\}dv\right)}. \tag{4.9}$$

Example 2 Consider the specific case

$$p(t) \equiv p;\ \lambda(t) \equiv \lambda.$$

Then (4.9) becomes:

$$\lambda_S(t) = \frac{p\lambda e^{\alpha\lambda t}}{1+p(1-e^{\alpha\lambda t})}. \tag{4.10}$$

There are two parameters in this model, i.e., parameter λ describes some general properties. On the other hand, α can be interpreted as an aging parameter (extent of aging), i.e., the increase in α results in a more pronounced influence of the history of the corresponding GPP.

4.4 Delayed Failures and Shot-Noise Processes

Let, as previously, $N(t), t \geq 0$ be an orderly point process (without multiple occurrences) of some 'initiating' events (IEs) with arrival times $T_1 < T_2 < T_3 < \ldots$. Assume now that each event from this process triggers the 'effective event' (EE), which occurs after a random time (delay) $D_i, i = 1, 2, \ldots$ since the occurrence of the corresponding IE at T_i [8]. The sequence of EEs $\{T_i + D_i\}, i = 1, 2, \ldots$ form now a new point process. This setting can be encountered in many practical situations, when, e.g., initiating events start the process of developing the non-fatal faults in a system and we are interested in the number of these faults in $[0, t)$. For instance, the initiation of cracks in a material was modelled by the nonhomogeneous Poisson process (NHPP), where $D_i, i = 1, 2, \ldots$ were assumed to be i.i.d random variables (see [16] and [5]) for the NHPP case. Alternatively, each EE can result in a fatal, terminating failure and then one can be interested in the survival probability of a system. Therefore, the latter setting means that the first EE results in the failure of our system. Obviously, a failure (the first EE) should not necessarily correspond now to the first IE. This setting in a slightly more generality was considered in [8], however, also only for the case of the NHPP. Note that the IEs can often be interpreted as some external shocks affecting a system, and, for convenience, we will use this term (interchangeably with the "IE").

It was proved in [16] that when the process of imitating events is the NHPP, the process of effective events is also NHPP. This is a rather unique property. Indeed, denote the rate of the IEs of a general ordinary process by $r(t)$. Then the rate of the point process with delays (EEs) is

$$r_D(t) = \int_0^t r(x)g(t-x)dx, \tag{4.11}$$

where 'D' stands for "delay" and $g(t)$ $(G(t))$ is the pdf (Cdf) of the i.i.d. $D_i, i = 1, 2, \ldots$. For instance, for the HPP with $r(t) = \lambda$, the rate of the NHPP of EEs, $r_D(t) = \lambda G(t)$ just follows the shape of the distribution of the delay and is increasing asymptotically to λ.

We will focus now on the corresponding survival model and relevant properties of the delayed model when the process of imitating events is GPP. Consider a system subject to the GPP (with the set of parameters $(\lambda(t), \alpha)$) of IEs $N(t), t \geq 0$, to be called, for convenience, shocks. Let the corresponding arrival times be denoted as $T_1 < T_2 < T_3 \ldots$. The sequence of EEs $\{T_i + D_i\}, i = 1, 2, \ldots$ form now a new point process, $\{N_E(t),\ t \geq 0\}$, where $D_i, i = 1, 2, ..$ are i.i.d., non-negative random variables with the pdf (Cdf) $g(t)$ $(G(t))$. Denote also the time to the first event in this process, which is the survival time if the EEs are fatal, by T_S. Thus, in this case, T_S is considered as the time to failure of our system. The following result for the distribution of the $N_E(t)$ (for each fixed t) and the survival function of the time to the first EE can be obtained in the spirit of [6] generalizing [8]. The latter paper considered the corresponding delay model for the NHPP of shocks, whereas the formulated result is already for the GPP process of shocks.

Under the given assumptions, the survival function that corresponds to the time to the first effective event and the corresponding failure rate are given respectively by

$$P(T_S > t) = \left(1 + \int_0^t G(t-x)\alpha\lambda(x)\exp\{\alpha\Lambda(x)\}dx\right)^{-\frac{1}{\alpha}}. \tag{4.12}$$

$$\lambda_S(t) = \frac{\int_0^t g(t-x)\lambda(x)\exp\{\alpha\Lambda(x)\}dx}{\left(1 + \int_0^t G(t-x)\alpha\lambda(x)\exp\{\alpha\Lambda(x)\}dx\right)}. \tag{4.13}$$

These formulas can be effectively used for obtaining reliability characteristics for the described models.

We will discuss now the corresponding shot noise process governed by the GPP. Denote by T the lifetime of a system subject to a shot noise process $X(t)$ and assume, for simplicity, that it is the only cause of the system's failure. Recall that the 'standard' shot noise process $X(t)$ is defined in the literature as (e.g., [19, 21]):

$$X(t) = \sum_{j=1}^{N(t)} D_j h(t - T_j),$$

where T_j is the j-th arrival time in the point (shock) process $\{N(t), t \geq 0\}$, $D_j, j = 1, 2, \ldots$ are the i.i.d. magnitudes of shocks and $h(t)$ is a nonnegative and non-increasing deterministic function for $t \geq 0$ and $h(t) = 0$ for $t < 0$. Throughout this paper, we assume that $\{N(t), t \geq 0\}$ and $\{D_j, j = 1, 2, \ldots\}$ are independent. Obviously, if $D_j \equiv 1$, $h(t) \equiv 1$, then $X(t) = N(t)$. Note that the simplified version

$$X(t) = \sum_{j=1}^{N(t)} h(t - T_j)$$

can be loosely considered as the generalization of the *counting* process $\{N(t), t \geq 0\}$ (although, strictly speaking, $X(t)$ is not a counting process) in the sense that when $h(t)$ is decreasing, it gives different, time-dependent weights to the previous events (counts) in the process, i.e., the larger the time elapsed since the event, the smaller is its input.

Obviously, $X(t)$ does not possess the independent increments property. It should be noted that, in most of the applications of the shot noise processes, it is assumed that the underlying shock process $\{N(t), t \geq 0\}$ is Poisson (homogeneous or nonhomogeneous), whereas, in real life, shock processes usually do not possess the independent increments property. For instance, the incidence of the subsequent heart attack depends on how many heart attacks the patient had experienced previously. Similar considerations can be true, e.g., for earthquakes. Therefore, a more adequate model will be the one that takes into account the history of the shock process in the form of the number of the previously occurred events. The GPP perfectly conforms to this intention. Moreover, it can be shown that this process possesses a *positive dependent increments property,* which means that the susceptibility of the event occurrence in an infinitesimal interval of time increases as the number of events in the previous interval increases. Thus, in what follows we will assume that our system is subject to the GPP of shocks $\{N(t), t \geq 0\}$ with the set of parameters $(\lambda(t), \alpha)$ and will describe some properties of the shot noise process $X(t)$ before addressing the corresponding survival model.

The cumulative impact of external shocks modeled by the shot noise process $X(t)$ can be probabilistically described in different ways. In this paper, we follow the meaningful approach of [18] by assuming that the corresponding failure (hazard) rate process [15] (on condition that $\{N(t), T_1, T_2, \ldots, T_{N(t)}\}$ and $\{D_1, D_2, \ldots, D_{N(t)}\}$ are given) is proportional to $X(t)$. This is a reasonable assumption that describes the proportional dependence of the probability of failure of a system in the infinitesimal interval of time on the level of stress, i.e.,

$$r_t \equiv kX(t) = k \sum_{j=1}^{N(t)} D_j h(t - T_j), \tag{4.14}$$

where r_t stands for the corresponding failure (hazard) rate process and $k > 0$ is the constant of proportionality. In general, the survival probability of an item described by the hazard rate process r_t is given by the following expectation with respect to

this process

$$P(T > t) = E\left[\exp\left\{-\int_0^t r_t dt\right\}\right].$$

The following result generalizing the corresponding NHPP model for the GPP of shocks [10] can be obtained using approached developed in [6].

The system survival function and the failure rate function when $\{N(t), t \geq 0\}$ is GPP *are given by*

$$P(T > t) = \left(\frac{1}{\exp\{\alpha\Lambda(t)\} - \int_0^\infty \int_0^t \exp\{-kuH(t-x)\}\alpha\lambda(x)\exp\{\alpha\Lambda(x)\}dxf_D(u)du}\right)^{\frac{1}{\alpha}}, \tag{4.15}$$

and

$$\lambda_S(t) = \frac{\int_0^\infty \int_0^t kuh(t-x)\exp\{-kuH(t-x)\}\beta\lambda(x)\exp\{\alpha\Lambda(x)\}dxf_D(u)du}{\exp\{\alpha\Lambda(t)\} - \int_0^\infty \int_0^t \exp\{-kuH(t-x)\}\alpha\lambda(x)\exp\{\alpha\Lambda(x)\}dxf_D(u)du}, \tag{4.16}$$

where $\Lambda(t) = \int_0^t \lambda(u)du, \quad H(t) = \int_0^t h(u)du.$

4.5 GPP for the Preventive Maintenance Model

As was mentioned before, the most attractive feature of this process is that it combines probabilistic tractability of the NHPP with a more realistic description of the process of failures (i.e., it takes into account the corresponding history). As another meaningful and practical example, we will show now the usefulness of the GPP for modeling the processes of failures/repairs.

Let each failure of an item which arrives in accordance with the GPP and is GPP-type repaired, meaning that the next cycle of an item's operation after the repair is in accordance with the next cycle of the GPP process. An item is replaced at $t = T$ and the process restarts. Denote the cost of the GPP repair by C_{GPP}. Thus, the long-run mean cost rate function for this case can be obtained as

$$C(T) = \frac{C_{GPP}E[N(T)] + C_r}{T}, \tag{4.17}$$

where $E[N(T)]$ is given by (4.6) and we can find the optimal T that minimizes this function. The following result was proved in [6]:

Let

$$\lambda'(x) + c^{-1}\lambda^2(x) > 0, t \geq 0. \tag{4.18}$$

Then there exists the unique, finite optimal T that minimizes C(T) in (4.17).

It is important to note that for the existence of the optimal PM time for the NHPP case the failure rate should be an increasing function, however, in (4.18) it can even decrease. Let us consider the 'marginal case' when $\lambda(x) = 1/t$ (the integral of this function in $[0, \infty)$ is still infinity, i.e., it cannot decrease as a power function faster than $1/t$ in order the corresponding distribution to be proper). Then (4.18) is satisfied for $c^{-1} = \alpha > 1$. Thus, GPP brings additional deterioration, which is creating the possibility for PM.

Let now each failure be classified as a minor failure with probability $q(t)$ (and then GPP-repaired) or as a major failure with probability $p(t) = 1 - q(t)$ (and then an item is replaced). The cost structure is as follows: the cost of the PM and the failure are C_r (repair) and C_f (failure), respectively, whereas the cost of the GPP repair is C_{GPP} and $C_f > C_r > C_{GPP}$. The PM is scheduled at $t = T$ and the length of a cycle is defined either by T, or the time of a major failure, whichever comes first. The distribution of time to a major failure/replacement (in the absence of the truncating T) in this case is given by (4.8), whereas the corresponding failure rate function is given by (4.9).

Note, that the rate of the GPP process of repairs is $\lambda(t)\exp\{\alpha\Lambda(t)\}$, where $\Lambda(t) = \int_0^t \lambda(u)du$. However, we need the rate (in fact, the integral of the rate, which is the mean number of events) of the conditional process on condition that the major failure did not occur. For the NHPP, this rate was $q(t)\lambda(t)$ and $E[N(t)] = \int_0^t q(x)\lambda(x)dx$. This quantity for the GPP was also obtained in [10] as

$$E\left[N_q(t)|N_p(t) = 0\right] = \frac{\int\limits_0^t q(u)\lambda(u)\exp\{-\alpha(\Lambda(t) - \Lambda(u))\}du}{1 - \int\limits_0^t \alpha q(u)\lambda(u)\exp\{-\alpha(\Lambda(t) - \Lambda(u))\}du}. \tag{4.19}$$

For the sake of notation, denote the right hand side of (4.19) by $G(t)$. Thus, similar to the NHPP case, we are able now to define the mean long-run cost rate and obtain the optimal time of replacement as,

$$C_p(T) = \frac{\bar{F}(T)\ (C_r + C_{GPP}G(T)) + \int_0^T \left(C_f + C_{GPP}G(u)\right)f(u)du}{\mu_{Tp}}, \tag{4.20}$$

where $\mu_{Tp} = \int_0^T \bar{F}_p(x)dx$ and $f(t) = F'(t)$. To find the optimal value of T that minimizes $C_p(T)$, consider $C_p'(T) = 0$, which can be written after simple algebra as

$$(\lambda_S(T)\mu_{Tp} - F_p(T))$$

$$+\frac{C_{GPP}(G'(T)\mu_{Tp} - (\bar{F}_p(T)G(T) + \int_0^T G(u)f_p(u)du))}{C_f - C_r} = \frac{C_r}{C_f - C_r}. \tag{4.21}$$

It is easy to verify via considering the corresponding derivative that the left hand side of (4.21) is increasing when for all $T \geq 0$,

$$(C_f - C_r)\lambda_S'(T) + C_{GPP}G(T)' > 0. \tag{4.22}$$

It can be easily shown that, e.g., when $p(t) \equiv p$, (22) holds for our assumptions (we assume that $\lambda(t)$ is increasing). In order to cross the line $y = C_r/(C_c - C_r)$ and to ensure a single, finite optimal T, the left hand side of (21) (that is equal to 0 at $T = 0$) should obey the following condition (at $T = \infty$):

$$\lambda_S(\infty) + \frac{C_{GPP}G'(\infty)}{C_f - C_r} + \frac{C_{GPP}\int_0^\infty G(u)f_p(u)du}{\mu_p(C_f - C_r)} > \frac{C_f}{(C_f - C_r)\mu_p}. \tag{4.23}$$

It is clear that when, e.g., $\lim_{t\to\infty}\lambda_S(t) = \infty$, this condition is always met and we have a single finite solution for optimal T.

In the following example, the baseline $\lambda(t)$ is set to be a constant, however, $\lambda_S(t)$ is increasing, although counter-intuitively not to infinity, which is due to heterogeneity induced by the history of the corresponding GPP [11].

Example 3 Suppose that $\lambda(t) = \lambda = 1$, $p(t) = 0.1 \forall t \geq 0$, $\alpha = 1.2$, $C_f = 12$, $C_r = 4$, $C_{GPP} = 1.5$. Then the corresponding calculations show that there exists the unique optimal solution $T \approx 1.63$.

4.6 Concluding Remarks

The most popular and well-studied point processes in reliability are the renewal and Poisson processes. Although asymptotic properties of the renewal process are quite simple and applied in many reliability problems (e.g., in optimal maintenance), the finite time solutions are usually rather cumbersome and involve numerical methods in applications. On the other hand, models involving NHPP of point events can be usually described mathematically in a close, tractable form. However, the main deficiency of a Poisson process is that it assumes the independent increments for the occurrence of the relevant events. Therefore, in this paper, we discuss a meaningful generalization of the NHPP, the generalized Polya process (GPP). This process that

was recently introduced in the literature, on one hand, allows for a more 'rich' history and on the other hand, presents tractable solutions for probabilities of interest that can be applied in practice.

In this paper, we have briefly reviewed and discussed some of the GPP models in reliability. The focus was on shock modelling, however in Sect. 8.5, the corresponding optimal preventive maintenance problem was also considered. We hope that it will attract some attention of the wider audience of specialists to this process as real applications usually imply the dependence of events of interest on the history. The GPP presents a tractable and mathematically clear tool for that.

References

1. Asfaw ZG, Linqvist B (2015) Extending minimal repair models for repairable systems: a comparison of dynamic and heterogeneous extensions of a nonhomogeneous Poisson process. Reliab Eng Syst Saf 140:153–158
2. Aven T, Jensen U (1999) Stochastic models in reliability. Springer, New York
3. Babykina G, Couallier V (2014) Modelling pipe failures in water distribution systems: accounting for harmful repairs and a time-dependent covariate. Int J Perform Eng 10:31–42
4. Block HW, Borges WS, Savits TH (1985) Age-dependent minimal repair. J Appl Probab 22:370–385
5. Caballé NC, Castro IT, Pérez CJ, Lanza-Gutiérrez JM (2015) A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. Reliab Eng Syst Saf 134:98–109
6. Cha JH (2014) Characterization of the generalized Polya process and its applications. Adv Appl Probab 46:1148–1171
7. Cha JH, Finkelstein M (2009) On a terminating shock process with independent wear increments. J Appl Probab 46:353–362
8. Cha JH, Finkelstein M (2011) On new classes of extreme shock models and some generalizations. J Appl Probab 48:258–270
9. Cha JH, Finkelstein M (2015) On some mortality rate processes and mortality deceleration with age. J Math Biol 72:331–342
10. Cha JH, Finkelstein M (2016) New shock models based on the generalized Polya process. Eur J Oper Res 251:135–141
11. Cha JH, Finkelstein M (2016) Justifying the Gompertz curve of mortality via the generalized Polya process of shocks. Theor Popul Biol 109:54–62
12. Cox DR, Isham V (1980) Point processes. University Press, Cambridge
13. Finkelstein M (2008) Failure rate modelling for reliability and risk. Springer, London
14. Finkelstein M, Cha JH (2013) Stochastic modelling for reliability. Shocks, burn-in and heterogeneous populations. Springer, London
15. Kebir Y (1991) On hazard rate processes. Nav Res Logist 38:865–877
16. Kuniewski SP, van der Weide JA, van Noortwijk JM (2009) Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection. Reliab Eng Syst Saf 94:1480–1490
17. Le Gat Y (2014) Extending the Yule process to model recurrent pipe failures in water supply networks. Urban Water J 11:617–30
18. Lemoine AJ, Wenocur ML (1986) A note on shot-noise and reliability modeling. Oper Res 34:320–323

19. Lund R, McCormic W, Xiao U (2004) Limiting properties of poisson shot noise processes. J Appl Probab 41:911–918
20. Nakagawa T (2007) Shock and damage models in reliability theory. Springer, London
21. Rice J (1977) On generalized shot noise. Adv Appl Probab 9:553–565
22. Ross SM (1996) Stochastic processes. Wiley, New York
23. Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, New York

Chapter 5
Degradation-Based Reliability Modeling of Complex Systems in Dynamic Environments

Weiwen Peng, Lanqing Hong, and Zhisheng Ye

Abstract Benefiting from the intimate link and the sufficient information conveyed by the degradation-threshold failure mechanism, degradation analysis has gradually become a hot topic in reliability engineering, which has been investigated extensively in the recent two decades. Various degradation models have been introduced to facilitate the reliability modeling and assessment of modern products, especially for highly reliable products. As the continual evolving of these models, there is a growing trend of investigation of reliability modeling and assessment based on degradation analysis. However, modern complex systems are characterized as multi-functional and subject to dynamic environments. Two aspects are indispensable for the investigation of degradation based reliability modeling and assessment of modern complex systems: (1) how to deal with complex systems with more than one degradation indicators, and (2) how to incorporate the effects of dynamic environments. To advance the research on degradation modeling and analysis of complex systems, this paper presents a summary of the state of the arts on the research of reliability modeling of complex systems by taking account of these two aspects. In this paper, the review is delivered in two progressive stages: multiple degradation processes under static environments, and multiple degradation processes under dynamic environments. Some discussion on further research topics from both theoretical and practical perspectives are presented to conclude the paper.

Keywords Stochastic process models • Inverse Gaussian process • Reliability modelling • Reliability assessment • Data-driven methods

W. Peng
School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, China

L. Hong • Z. Ye (✉)
Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, Singapore
e-mail: yez@nus.edu.sg

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_5

5.1 Introduction

Continually evolving paces of technology advances and ever-increasing requirements of system performance have made modern systems become more and more complex. Reliability engineers are encountered with paradoxical situations of reliability modelling of complex systems, for which the situation introduced by limited failure time data is an inevitable challenge [1], however this situation is contradicted by the growing reality that reliability data are stepping into big data situation [2], which is enhanced by the gradual availability of the system performance data, system operating profile, and working environment information. To facilitate the reliability analysis of complex system with limited failures under big data situation, degradation based reliability modeling and assessment has gradually become a hot topic, benefiting from the intimate link and the sufficient information conveyed by the degradation-threshold failure mechanism [3]. For the degradation based reliability analysis, the failure process of a complex system is reflected by degradation processes of some performance indicators, where the system fails when these degradation processes reach predefined thresholds. Accordingly, the failure time distribution can be determined through the degradation analysis of system performance data together with the system operating profile and working environment information, which provides a promising solution to the difficulty introduced by limited failure time data.

In recent two decades, degradation based reliability modelling has been investigated extensively. Classical examples include, but not limited to the works presented recently by Pan and Balakrishnan [4], Wang and Pham [5], Kharoufeh et al. [6], Liao and Tian [7], Wang et al. [8], Bian and Gebraeel [9], Si et al. [10], Ye and Xie [11], Hong et al. [12], Peng et al. [13], and Xu et al. [14]. Generally, two critical assumptions are used in the degradation based reliability analysis, which include the assumptions of (1) single degradation indicator, and (2) constant external factors. For complex systems under dynamic environments, these two assumptions are challenged greatly. These challenges are raised from the growing awareness that complex systems are composed of multiple components with multiple functionalities, and these systems are subject to various operating profiles under dynamic working conditions. The failure process of a complex system generally relates to degradation processes of multiple performance indicators, which are often presented with various characteristics. Moreover, the degradation processes are affected by the operating profiles and working conditions, which lead to the indicator-to-indicator dependency within the multiple degradation processes, and the unit-to-unit variability among the system population group. A classic example is the multivariate dependent degradation analysis of one type of heavy duty lathes presented in Peng et al. [15]. Within this study, the heavy duty lathes are founded subject to two types of gradually-evolving failures: losing of machining accuracy, and accumulation of lubrication debris, which are critical to the reliability of these lathes. In addition, by summarizing operating and maintenance records, these types of failures are found to vary from factories to factories, which is caused by differences of loading profiles

and environmental conditions experienced by the heavy duty lathes. Accordingly, two critical aspects arise for the degradation based reliability modelling of complex systems under dynamic environments: (1) how to deal with complex systems with more than one degradation indicators, and (2) how to incorporate the effects of dynamic operating profile and time-varying working environments.

To facilitate the research on degradation based reliability analysis, this paper devotes to the review of degradation based reliability modelling of complex systems under dynamic environments. There are excellent review papers on degradation based reliability modelling over the past few decades, such as the reviews by Singpurwalla [16], Lee and Whitmore [17], Bae et al. [18], Si et al. [19], and Ye and Xie [11]. These review papers have covered various aspects of degradation based reliability modelling. However, the challenges introduced by multiple degradation processes and dynamic environments are still underdeveloped. Therefore, this paper is to highlight the state of arts on multiple degradation analysis under dynamic environments through two progressive categories: multiple degradation processes under static environments, and multiple degradation processes under dynamic environments.

The remainder of this paper is organized as follows. Section 5.2 describes the characteristics and modelling of dynamic environments encountered by complex systems. Section 5.3 reviews the degradation based reliability modelling of complex systems with multiple degradation process models under static environments. Section 5.4 reviews the degradation based reliability modelling through multiple degradation processes under dynamic environments. A short discussion on further research topics is presented in Sect. 5.5 to conclude the paper.

5.2 Dynamic Environments

Degradation processes of a complex system are closely related to the system's operation profile and working environments. This is because failure processes of a complex system are often driven or affected by the factors and stresses originated from the dynamic environments experienced by the system. The rates and the modes by which the system degrades are closely related to what type of missions the system is fulfilling, and which kind of environments is experienced. Take a manufacturing system as an example, the system can fulfill the missions of turning, drilling, and milling with different working speeds and depths on various types of materials. These parameters consist of the basic operational profile of the manufacturing system, for which a specific mission with particular working loads on one specific working piece determine the basic degradation rate and mode of the manufacturing systems. In addition, the temperature, humidity, and vibration condition experienced by the manufacturing system consist of the basic working environments, which further modify or change the degradation rate or even degradation mode of the system.

The dynamic environments experienced by a system is generally composed by its operational profile and working environment. The operating profile is composed of the variables that can be modified or determined by the operators or users to fulfill different missions. The working environment mainly refers to the factors that can hardly be controlled and are mainly determined by natural forces. The differentiation of these two groups is aimed to highlight the effect introduced by different types factors, and to deliver the notion that some variables can be well modeled and incorporated into the degradation modelling yet some factors can hardly be well characterized or integrated. As a result, two aspects are critical for the handling of dynamic environments within the degradation based reliability analysis: (1) how to characterize the dynamic environments, and (2) how to incorporate them in degradation modelling.

5.2.1 Characterization of Dynamic Environments

The characterization of dynamic environments is implemented based on the availability of environmental information and the capability of mathematical models [16, 20]. General methods for the characterization of dynamic environments include the time-varying deterministic function [22], probability distribution model [23], stochastic processes model [24, 29], time series models [2, 12], and so on.

Time-varying deterministic function is generally used for the situation where environmental variables are well controlled and their dynamic behavior follows specific patterns, such as accelerated degradation tests [25, 26]. The probability distribution model is used to model the environmental variables which are presented as shocks with random strengths, such as vibration shocks for mechanical systems [27, 28]. The stochastic processes model is often adopted for the situation where the environmental variables are evolved stochastically with temporal variability and epistemic uncertainty, such as different operating speeds for a rotational machine [24, 29]. The time series model is used for the situation where the environmental variables have highly variable behaviors with periodicity and autocorrelations, such as the solar radiation and temperature for organic coatings [12, 30].

These methods for the characterization of dynamic environments are mainly used for the environmental variables that can be well identified and collected, where system operating/environmental data highlighted by Meeker and Hong [2] can be obtained for the corresponding model derivation and parameter estimation. There are situations that environmental variables cannot be identified or collected, such as the micro-shock for a micro-electromechanical system. This kind of environmental variables and their effect are lumped together and incorporated into the degradation modelling through a random effect or frailty term [31–34].

5.2.2 *Incorporation of Dynamic Environments*

The incorporation of dynamic environments into degradation modelling is carried out based on the availability of influence mechanism and the flexibility of mathematical models. General methods for the incorporation of dynamic environments include the covariate method [20–23] and cumulative damage method [6, 24, 35].

The covariate method has been used extensively to incorporate the effect of environmental factors into degradation modelling. The environmental factors are represented as explanatory variables, and their effects are modelled through a covariate-effect function. The covariate-effect function is further used to modify the model parameters, which are related to the degradation rate, shape, mode or diffusion of a degradation process model. The covariate-effect function is generally determined based on the physical, chemical, and engineering knowledge, such as Arrhenius relationship, Eyring relationship, power low relationship, linear relationship, inverse-logit relationship and exponential relationship [3]. Classical examples of covariate method for incorporation of dynamic environments are the works presented by Ye et al. [11] and Lawless and Crowder [23] separately for Wiener process and gamma process models.

The cumulative damage method is introduced for the situation that the degradation rate of the degradation process is dominated by random environmental factors. The environmental factors are characterized as stochastic processes, such as the Wiener process [16], continuous-time Markov chain process [6], and Semi-Markov process [24]. The degradation rate of the degradation process is modelled through a functional relationship of the environmental factors. The degradation at a specific point is the accumulation result of the degradation increments throughout the time interval from the beginning to this specific point, where the degradation increments within an infinitesimal time interval is dominated by the specific environmental factors within that interval. This method is generally used for the situation that the degradation process of a system cannot be directly observed, however the environmental factors that dominate the degradation process can be observed and their relationship is well understood. Classical example of cumulative damage method for incorporation of dynamic environments can refer to the method introduced by Bian et al. [29] and Flory et al. [36].

5.3 Multiple Degradation Processes Under Static Environments

The premise of degradation based reliability assessment of complex system is to choose a proper model to characterize the degradation process of the complex system, based on the its performance indicators and dynamic environments. Various types of multiple degradation processes models have been introduced in the recent decade, which can be categorized into two types, that is, the models for multiple

degradation process models under static environments and the models for multiple degradation process models under dynamic environments. This section devotes to the first type under static environments, including the multivariate Gaussian distribution based model, the bivariate Birnbaum-Saunders distribution based model, the degradation rate interaction model, and the copula based multivariate degradation process model.

5.3.1 Multivariate Gaussian Distribution Based Model

The multivariate Gaussian distribution based model is introduced out of the consideration that a system may have multiple degradation paths and the distribution of the degradation observations of these paths at a specific time point can be described by a joint multivariate Gaussian distribution [37]. Within this model, each degradation path is described by a marginal Gaussian distribution from the joint multivariate Gaussian distribution. The dependency among the degradation paths is characterized by the variance-covariance of the joint multivariate Gaussian distribution.

Suppose a product has L performance indicators. Let $Y_l(t)$ with $l = 1, \ldots, L$ denote the degradation process of the lth performance indicator. The joint distribution of the L performance indicators, $\mathbf{Y}(t) = (Y_1(t), \ldots, Y_L(t))^{\mathrm{T}}$, at a specific time point t is given as follow.

$$f(y_1(t), \ldots, y_L(t)) = \frac{1}{\sqrt{(2\pi)^L |\mathbf{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y}(t) - \boldsymbol{\mu}(t))^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{y}(t) - \boldsymbol{\mu}(t))\right) \tag{5.1}$$

where $\mathbf{y}(t) = (y_1(t), \ldots, y_L(t))^{\mathrm{T}}$, $\boldsymbol{\mu}(t) = (\mu_1(t), \ldots, \mu_L(t))^{\mathrm{T}}$ with $\mu_l(t)$ denoting degradation mean function of the lth performance indicator, and $\mathbf{\Sigma}$ is the covariance matrix and $|\mathbf{\Sigma}|$ is the determinant of $\mathbf{\Sigma}$.

The degradation mean function is a description of the average degradation observation of the performance indicator, such as $\mu_l(t) = at$ indicating linear degradation path of the lth performance indicator. The covariance matrix is a description of the variance and correlation among the degradation processes. Let $\mathrm{Var}(Y_l(t))$ denote the variance of $Y_l(t)$, and $\mathrm{Cov}(Y_{l-1}(t), Y_l(t))$ denote the covariance of $Y_{l-1}(t)$ and $Y_l(t)$. The general form of the covariance matrix $\mathbf{\Sigma}$ is given as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathrm{Var}(Y_1(t)) & \mathrm{Cov}(Y_1(t), Y_2(t)) & \cdots & \mathrm{Cov}(Y_1(t), Y_L(t)) \\ \mathrm{Cov}(Y_2(t), Y_1(t)) & \mathrm{Var}(Y_2(t)) & \cdots & \mathrm{Cov}(Y_2(t), Y_L(t)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(Y_L(t), Y_1(t)) & \mathrm{Cov}(Y_L(t), Y_2(t)) & \cdots & \mathrm{Var}(Y_n(t)) \end{bmatrix} \tag{5.2}$$

Generally, it's difficult to specify the functional forms of $\mathrm{Var}(Y_l(t))$ and $\mathrm{Cov}(Y_{l-1}(t), Y_l(t))$. The assumption of time-invariant variance and covariance is adopted for the application of this model [37]. Under the time-invariant assumption, the covariance matrix $\mathbf{\Sigma}$ is constant over time, where the variance and covariance are given as $\mathrm{Var}(Y_l(t)) = \sigma_l^2$ and $\mathrm{Cov}(Y_{l-1}(t), Y_l(t)) = \rho_{l-1,l}\sigma_{l-1}\sigma_l$ with $\rho_{l-1,l}$ indicating the relevance between $Y_{l-1}(t)$ and $Y_l(t)$. The parameters within the covariance matrix $\mathbf{\Sigma}$ are greatly reduced under the time-invariant assumption. In addition, the marginal distribution of the degradation path for the lth performance indicator is given as $Y_l(t) \sim \mathrm{N}\left(\mu_l(t), \sigma_l^2\right)$, and its degradation increments is $\Delta Y_l(t) \sim \mathrm{N}\left(\Delta\mu_l(t), \sigma_l^2\right)$. The model is simplified to a great extent. The temporal variability within the degradation process is missing under the time-invariant assumption due to the time invariant variance $\mathrm{Var}(\Delta Y_l(t)) = \sigma_l^2$.

The failure time T of the product is defined as the first time point that any of the L degradation processes reaches its degradation threshold D_l with $l = 1, \ldots, L$. The reliability function of the product is then given as

$$\begin{aligned} R(t) &= \Pr\left(Y_1(t) \le D_1, \ldots, \mathrm{Y}_L(t) \le D_L\right) \\ &= \int_0^{D_1} \cdots \int_0^{D_L} f\left(y_1(t), \ldots, y_L(t)\right) dy_1(t) \cdots dy_L(t) \\ &= \Phi_L\left(D_1, \ldots, D_L; \boldsymbol{\mu}(t), \mathbf{\Sigma}\right) \end{aligned} \tag{5.3}$$

where $\Phi_L(\bullet; \boldsymbol{\mu}(t), \mathbf{\Sigma})$ is the cumulative distribution function (CDF) of L dimension multivariate Gaussian distribution with mean $\boldsymbol{\mu}(t)$ and covariance matrix $\mathbf{\Sigma}$.

The multivariate Gaussian distribution based model does not receive wide application in degradation data analysis. However, this idea has roused the investigation on degradation analysis with multiple performance indicators, such as the works presented by Pan and Balakrishnan [4], Bian and Gebraeel [9], Sari et al. [38] and Peng et al. [15], which are serving as key stones for the following developed multiple degradation process models.

5.3.2 *Multivariate Birnbaum-Saunders Distribution Based Model*

The multivariate Birnbaum-Saunders distribution based model is introduced by taking the idea that the failure time distribution derived for the one-dimensional degradation can be approximated closely by Birnbaum-Saunders distribution [39]. The multivariate Birnbaum-Saunders distribution and its marginal distributions are then adopted to construct a multivariate degradation process model by Pan and Balakrishnan [4] and Pan et al. [40]. Within this model, the marginal degradation processes are modelled by gamma processes. The dependency between these degradation processes is constructed by assuming that the degradation increments of the multivariate degradation process within the same time interval are dependent, where a time-invariant correlation coefficient is used to characterize this dependency.

For a product with L performance indicators, the degradation process of the lth performance indicator is modelled as a gamma process, $Y_l(t) \sim \mathrm{Ga}(\nu_l t, \gamma_l)$ with $l = 1, \ldots, L$. The degradation process $Y_l(t)$ has independent and gamma-distributed increments as $\Delta Y_l(t) \sim \mathrm{Ga}(\nu_l \Delta t, \gamma_l)$, where $\Delta Y_l(t) = Y_l(t + \Delta t) - Y_l(t)$. The probability density function (PDF) of the degradation increment $\Delta Y_l(t)$ is given as

$$f\left(\Delta y_l(t) | \nu_l \Delta t, \gamma_l\right) = \frac{1}{\Gamma\left(\nu_l \Delta t\right) \gamma_l^{\nu_l \Delta t}} \left(\Delta y_l(t)\right)^{\nu_l \Delta t - 1} \exp\left(-\frac{\Delta y_l(t)}{\gamma_l}\right) \tag{5.4}$$

The degradation increments of different performance indicators at the same time interval are dependent, which is further described by time-invariant correlation coefficient as $\mathrm{Corr}(Y_{l-1}(t), Y_l(t)) = \rho_{l-1,l}$. However, the degradation increments at different time intervals are assumed independent. To further describe the dependence among the performance indicators, a random variable is introduced by normalizing the degradation increments as $X_l(t) = \left(\Delta Y_l(t) - \nu_l \Delta t \gamma_l\right) / \left(\sqrt{\nu_l \Delta t} \gamma_l\right)$, and all the observation intervals are assumed having the same length . In addition, the dependency among the performance indicator is further described as $\mathrm{Corr}(X_{l-1}(t), \mathrm{X}_l(t)) = \rho_{l-1,l}$. Given the degradation threshold of all the performance indicators as D_l with $l = 1, \ldots, L$, the reliability function of the product is given as follows by utilizing the central limit theorem [40].

$$\begin{aligned}
R(t) &= \Pr\left(Y_1(t) \le D_1, \ldots, \mathrm{Y}_L(t) \le D_L\right) \\
&= \Pr\left(\sum_{j=1}^{t/\Delta t} \Delta Y_{1j}\left(t_j\right) \le D_1, \ldots, \sum_{j=1}^{t/\Delta t} \Delta Y_{Lj}\left(t_j\right) \le D_L\right) \\
&= \Pr\left(\sqrt{\frac{\Delta t}{t}} \sum_{j=1}^{t/\Delta t} \frac{\Delta Y_{1j}\left(t_j\right) - \nu_1 \Delta t \gamma_1}{\sqrt{\nu_1 \Delta t} \gamma_1} \le \frac{D_1 - \nu_1 t \gamma_1}{\sqrt{\nu_1 t} \gamma_1}, \ldots,\right. \\
&\quad \left.\sqrt{\frac{\Delta t}{t}} \sum_{j=1}^{t/\Delta t} \frac{\Delta Y_{Lj}\left(t_j\right) - \nu_L \Delta t \gamma_L}{\sqrt{\nu_L \Delta t} \gamma_L} \le \frac{D_L - \nu_L t \gamma_L}{\sqrt{\nu_L t} \gamma_L}\right) \\
&= \Pr\left(\sum_{j=1}^{t/\Delta t} X_{1j}\left(t_j\right) \le \frac{D_1 - \nu_1 t \gamma_1}{\sqrt{\nu_1 t} \gamma_1}, \ldots, \sum_{j=1}^{t/\Delta t} X_{Lj}\left(t_j\right) \le \frac{D_L - \nu_L t \gamma_L}{\sqrt{\nu_L t} \gamma_L}\right) \\
&\approx \Phi_L\left(\frac{D_1 - \nu_1 t \gamma_1}{\sqrt{\nu_1 t} \gamma_1}, \ldots, \frac{D_L - \nu_L t \gamma_L}{\sqrt{\nu_L t} \gamma_L}; 0, \boldsymbol{\Sigma}\right)
\end{aligned} \tag{5.5}$$

where $\boldsymbol{\Sigma}$ the covariance matrix with the correlation coefficient between different performance indicators includes, which is given as

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1L} \\ \rho_{21} & 1 & \cdots & \rho_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{L1} & \rho_{L2} & \cdots & 1 \end{bmatrix} \tag{5.6}$$

By extending the bivariate Birnbaum-Saunders distribution introduced by Kundu et al. [41], Pan et al. [40] further derived that the lifetime time distribution of the product, $F(t) = 1 - R(t)$ can be expressed by the multivariate Birnbaum-Saunders distribution of a L-dimensional vector and all its marginal distributions.

Compared with the multivariate Gaussian distribution based model, the multivariate Birnbaum-Saunders distribution based model is derived from the perspective of degradation increments of performance indicators, where the distribution and dependence of degradation increments are highlighted, and the degradation process of performance indicators are monotonic. The temporal variability within each performance indicator is also characterized by the gamma process. However, due to the assumptions made in the derivation, the multivariate Birnbaum-Saunders distribution model is limited to the situation that the shape function of the gamma process should be linear as $v(t) = v_l t$. Parameter estimation method for this model is investigated by Pan et al. [40]. The application of this model for bivariate degradation analysis has also been presented by Pan and Balakrishnan [4] and Pan et al. [42].

5.3.3 *Degradation Rate Interaction Model*

The degradation rate interaction model is introduced from the perspective of a multi-component product by leveraging the idea of degradation rate modelling [9]. The product is composed of multiple components, and these components are associated with multiple dependent performance indicators. The dependence is represented through the consideration that the degradation process of one component can be affected by the deterioration of other components. This effect is characterized through the modelling of the degradation rate function, which consists of two separated parts of the inherent degradation rate of the component and the inductive degradation rate by contributing components. A multivariate degradation process model is constructed by separately integrating the degradation rate function and associating a stationary Brownian motion noise for each component.

Suppose a product is composed of L components, and each component has a performance indicator described by the degradation process $Y_l(t)$ with $l = 1, \ldots, L$. Let $r_l(t)$ denote the degradation rate associated with the degradation process $Y_l(t)$. To incorporate the dependence among the degradation processes, a general form of degradation rate $r_l(t)$ is given as $r_l(t) = r_l(t; \kappa(t), h(\mathbf{Y}(t)))$ [9]. The degradation rate

is defined on t and further determined by the functional relationship of $\kappa(t)$ and $h(\mathbf{Y}(t))$, for which $\kappa(t)$ is the inherent degradation rate of the component without the interaction of other components, and $h(\mathbf{Y}(t))$ is the inductive degradation rate by taking account of the effect from other components. A simplified form of the degradation rate is given as follows [43].

$$r_l(t) = \kappa_l + \sum_{l \neq k} \delta_{lk} y_k(t) \tag{5.7}$$

where a time invariant inherent degradation rate κ_l, and a linear combination of inductive degradation rates $\delta_{lk} y_k(t)$ with constant coefficient δ_{lk} is used.

Based on the degradation rate function, the degradation process $Y_l(t)$ can be given as follows.

$$Y_l(t) = \int_0^t r_l\left(u; \kappa(u), h\left(\mathbf{Y}(u)\right)\right) du + \varepsilon_l(t) \tag{5.8}$$

where $\varepsilon_l(t)$ is an error item for capturing the measurement noise and unidentified uncertainty, which is generally given as a white noise process or a stationary Brownian motion process.

It is difficult to derive the failure time distribution of the product under the degradation process model given in Eq. 5.8. By assuming a simplified form of the degradation rate as given in Eq. 5.7 and a stationary Brownian error item, a approximated failure time distribution is obtained by Bian & Gebraeel [43], where the degradation processes of the product are reconstructed as

$$\mathrm{d}\mathbf{Y}(t) = (\boldsymbol{\kappa} + \boldsymbol{\delta} \times \mathbf{Y}(t))\, \mathrm{d}t + \mathrm{d}\boldsymbol{\varepsilon} \tag{5.9}$$

where $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_L)$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_L)$ which follows a multivariate normal distribution $\mathrm{MVN}(0, \boldsymbol{\sigma}^2 t)$ with $\boldsymbol{\sigma}^2 = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_L^2\right)$, and $\boldsymbol{\delta} \in \mathbb{R}^{L \times L}$ with non-diagonal entries δ_{lk} and diagonal entries $\delta_{ll} = 0$.

The stochastic differential equations given in Eq. 5.9 with initial condition $\mathbf{Y}(0) = \mathbf{Y}_0$ has been demonstrated to have a closed-form as $\mathbf{Y}(t)|(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2)$, which follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}_0(t)|(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2)$ and covariance matrix $\boldsymbol{\Sigma}_0(t)|(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2)$ as follow [43]

$$\boldsymbol{\mu}_0(t) \mid \left(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2\right) = \exp(t\boldsymbol{\delta}) \times \mathbf{Y}_0 + \int_0^t \exp\left((t-u)\,\boldsymbol{\delta}\right) \times \boldsymbol{\kappa}\, \mathrm{d}u \tag{5.10}$$

$$\boldsymbol{\Sigma}_0(t) \mid \left(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2\right) = \int_0^t \exp\left((t-u)\,\boldsymbol{\delta}\right) \times \boldsymbol{\sigma}^2 \times \exp\left((t-u)\,\boldsymbol{\delta}\right)^{\mathrm{T}} \mathrm{d}u \tag{5.11}$$

Given the degradation thresholds of the L performance indicators, the reliability function of the product can be approximated as follows.

$$\begin{aligned} R(t) &= \Pr\left(Y_1(t) \le D_1, \ldots, Y_L(t) \le D_L\right) \\ &\approx \Phi_L\left(D_1, \ldots, D_L; \boldsymbol{\mu}_0(t)\middle|\left(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2\right), \boldsymbol{\Sigma}_0(t)\middle|\left(\boldsymbol{\kappa}, \boldsymbol{\sigma}^2\right)\right) \end{aligned} \tag{5.12}$$

The calculation of the reliability function depends on the solution of the multidimensional integrals given in Eqs. 5.10 and 5.11, which can be obtained using mathematical software. Analytical solutions of Eq. 5.12 have been obtained by Bian and Gebraeel [43] for two special cases, that are the case with $\boldsymbol{\delta}$ being a diagonal matrix, and the case with $\boldsymbol{\delta}$ being a diagonalizable matrix.

Compared with the models introduced above, the degradation interaction model is introduced from the perspective of a multi-component system, where the dependence is originated from the mutual influences of components' deteriorations. A degradation interaction function and a linear system of stochastic differential equations are used to construct the multivariate degradation process model. If the characteristics of the systems, such as the structure, functionality, operating conditions, failure mechanism, and interaction patterns of components are well studied, the degradation rate interaction model is more suitable for degradation modelling and reliability analysis than the models presented above. However, limited by the availability of the product characteristic for determining the degradation interaction parameter $\boldsymbol{\delta}$, the degradation rate interaction has not received wide application in degradation based reliability modelling of the modern system.

5.3.4 Copula Based Multivariate Degradation Process Model

The copula based multivariate degradation process model is introduced by taking the copula based multivariate degradation process model is introduced by taking the advantage of the copula theory [44] for constructing multivariate probability distributions. By adopting a copula function, the dependence structure of random variables can be characterized separately from their marginal distribution functions. Within the copula based multivariate degradation process model, the degradation processes of performance indicators are separately modelled by marginal degradation processes with independent degradation increments, such as Wiener process, gamma process, and inverse Gaussian process. Their dependency is characterized by a copula function by assuming that their degradation increments of the multiple performance indicators at the same time interval are dependent. As a result, the characteristics of each performance indicator are described by its marginal degradation process. And their dependency is modelled by a copula function. This kind of multivariate degradation process model can facilitate the degradation modelling, parameter estimation, and reliability assessment, which has been widely

investigated, such as the works presented by Sari et al. [38], Pan et al. [45], Wang and Pham [5], Wang et al. [46], and Peng et al. [13].

For a product with L performance indicators, let $Y_l(t)$ denote the degradation process of the lth performance indicator, which is modelled by a stochastic process. Let $\Delta Y_l(t_j) = Y_l(t_j) - Y_l(t_{j-1})$ denote the degradation increment within the time interval $[t_{j-1}, t_j]$, which follows a probability distribution with CDF $F_l(\Delta y_l(t_j))$ under the specific stochastic process chosen for the marginal degradation process. The dependence among the degradation processes is characterized through their degradation increments. It is assumed that the degradation increments $\mathbf{\Delta Y}_l(t_j) = (\Delta Y_1(t_j), \ldots, \Delta Y_L(t_j))$ for the L performance indicators within the same time interval $[t_{j-1}, t_j]$ are dependent. The degradation increments in disjoint time intervals are independent, e.g., $Y_l(t_j) - Y_l(t_{j-1})$ and $Y_l(t_{j-1}) - Y_l(t_{j-2})$ separately in $[t_{j-1}, t_j]$ and $[t_{j-2}, t_{j-1}]$ are s-independent. The joint probability distribution of $\mathbf{\Delta Y}_l(t_j)$ within the same time interval is characterized by leveraging a multivariate copula function as follow.

$$F\left(\mathbf{\Delta y}\left(t_j\right)\right) = \mathrm{C}\left(F_1\left(\Delta y_1\left(t_j\right)\right), \ldots, F_L\left(\Delta y_L\left(t_j\right)\right); \boldsymbol{\theta}^{\mathrm{Cop}}\right) \tag{5.13}$$

where $\mathbf{\Delta y}(t_j) = (\Delta y_1(t_j), \ldots, \Delta y_L(t_j))$ and $\mathrm{C}(u_1, \ldots, u_L; \boldsymbol{\theta}^{\mathrm{Cop}})$ is a L-dimensional multivariate copula function with parameters $\boldsymbol{\theta}^{\mathrm{Cop}}$ and $u_l \sim \mathrm{Uniform}(0, 1)$.

Under the dependency structure given in Eq. 5.13, the CDF of each degradation increment $F_l(\Delta y_l(t_j))$ is the marginal distribution of the joint CDF $F(\mathbf{\Delta y}(t_j))$. This characteristic makes the modeling of dependence among performance indicators separated from the modeling of marginal degradation process of each performance indicator. Such separation makes the construction of multivariate degradation model with different marginal degradation models feasible. In general, under the multiple degradation processes model presented in Sects. 5.3.1, 5.3.2, and 5.3.3, the marginal degradation models for the performance indicators are from the same stochastic model family, such as the normal distribution, Wiener process and gamma process. However, the performance indicators of a product may have different deterioration characteristics, which need different types of stochastic model families [15]. This practical requirement can be fulfilled through the copula based multivariate degradation process model. For instance, the performance indicators can be separately modelled by Wiener processes, gamma processes, inverse Gaussian process and so on, and their dependence under different stochastic processes can be constructed through Eq. 5.13.

Since the performance indicator is assumed following a stochastic process with independent degradation increments, the multivariate degradation process model based on copula function is constructed as follows.

$$\begin{cases} Y_l\left(t_j\right) = \sum_{k=2}^{j} \Delta Y_l\left(t_k\right), l = 1, \ldots, L, j = 2, \ldots, +\infty, \\ \Delta Y_l\left(t_k\right) \sim \mathrm{PRO}\left(\Delta y_l\left(t_k\right); \boldsymbol{\theta}_l^{\mathrm{Mar}}\right) \\ F\left(\mathbf{\Delta y}\left(t_j\right)\right) = \mathrm{C}\left(F_1\left(\Delta y_1\left(t_j\right)\right), \ldots, F_L\left(\Delta y_L\left(t_j\right)\right); \boldsymbol{\theta}^{\mathrm{Cop}}\right) \end{cases} \tag{5.14}$$

where $\mathrm{PRO}\left(\Delta y_1\left(t_k\right);\boldsymbol{\theta}_l^{\mathrm{Mar}}\right)$ is the distribution of degradation increment of the lth performance indicator, which is determined based on the stochastic process chosen for the modelling of this performance indicator. The stochastic processes generally used for degradation modelling include the Wiener process, gamma process and inverse Gaussian process [3], which separately giving rise to the $\mathrm{PRO}\left(\Delta y_1\left(t_k\right);\boldsymbol{\theta}_l^{\mathrm{Mar}}\right)$ presenting as the normal distribution, gamma distribution, and inverse Gaussian distribution.

The multivariate copula function in Eq. 5.14 is a multivariate distribution with uniformly distributed marginal distributions on $[0,1]$. A group of CDFs of degradation increments $(F_1(\Delta y_1(t_j)),\ \ldots,F_L(\Delta y_L(t_j)))$ is a sample from the multivariate copula function. Accordingly, when the CDFs of degradation increments are available, the choice of copula function to model their dependency can be implemented through the methods for multivariate probability distribution selection. Qualitative method such as the scatter plots presented in Wu [47], and quantitative method such as Bayesian model selection introduced by Huard and Evin [48], both can be adopted to choose the right copula function for multivariate degradation modelling. In bivariate degradation processes, there are various types of copula functions can be used for degradation modelling, which are listed as follows.

1. Gaussian copula

$$\mathrm{C}\left(u_1,u_2\right)=\int_{-\infty}^{\Phi^{-1}(u_1)}\int_{-\infty}^{\Phi^{-1}(u_2)}\frac{1}{2\pi\sqrt{1-\alpha^2}}\exp\left(-\frac{u_1^2-2u_1u_2+u_2^2}{2\left(1-\alpha^2\right)}\right)\mathrm{d}u_1\mathrm{d}u_2,\alpha\in[-1,1] \tag{5.15}$$

2. Frank copula

$$\mathrm{C}\left(u_1,u_2\right)=-\frac{1}{\alpha}\ln\left(1+\frac{\left(\mathrm{e}^{-\alpha u_1}-1\right)\left(\mathrm{e}^{-\alpha u_2}-1\right)}{\mathrm{e}^{-\alpha}-1}\right),\quad \alpha\neq 0 \tag{5.16}$$

3. Gumbel copula

$$\mathrm{C}\left(u_1,u_2\right)=\exp\left(-\left(\left(-\ln u_1\right)^{\alpha}+\left(-\ln u_2\right)^{\alpha}\right)^{1/\alpha}\right),\quad \alpha\in[1,\infty) \tag{5.17}$$

4. Clayton copula

$$\mathrm{C}\left(u_1,u_2\right)=\max\left(\left(u_1^{-\alpha}+u_2^{-\alpha}-1\right)^{-1/\alpha},0\right),\quad \alpha\in[-1,\infty)\backslash 0 \tag{5.18}$$

In the multivariate degradation process, the multivariate Gaussian copula [49], the multivariate t-copula [50], and the vine copula [51] are commonly used in various literature.

Given the degradation thresholds of the L performance indicators, the reliability function of the product is given as follows.

$$R(t) = \Pr\left\{\sup_{s\le t} Y_1(s) < D_1, \ldots, \sup_{s\le t} Y_l(s) < D_l, \ldots, \sup_{s\le t} Y_L(s) < D_L\right\} \tag{5.19}$$

It's often difficult to calculate the reliability function due to the unavailability of the analytical solution. Simulation method is generally used to obtain the reliability of the product based on the model in Eq. 5.14 [13]. When all the parameters for the copula function and the marginal degradation processes are available, a group of random samples, $[\tilde{u}_1, \ldots, \tilde{u}_L]$, are firstly generated from the copula function, $\mathrm{C}(u_1, \ldots, u_L; \boldsymbol{\theta}^{\mathrm{Cop}})$. By calculating the inverse CDF of degradation increments based on the generated samples, the degradation increments for marginal degradation processes at a specific time interval are obtained as $\Delta\tilde{y}_l(t_k) = F_l^{-1}\left(\tilde{u}_l; \boldsymbol{\theta}_l^{\mathrm{Mar}}\right)$. The marginal degradation process at a specific time point is obtained based the degradation increments of the passed time intervals as $\tilde{y}_l\left(t_j\right) = \sum_{k=2}^{j} \Delta\tilde{y}_l(t_k)$. The failure time point of the product can then be obtained by comparing the simulated observations of the marginal degradation processes with their respective degradation thresholds. By repeating the simulation process above, a group of samples of failure time points is obtained, and the reliability function can be statistically summarized from these generated failure time samples.

The copula based model has been investigated extensively for bivariate degradation processes modeling, which can be summarized in the following three groups according to the stochastic processes incorporated, i.e., bivariate Wiener process model, bivariate gamma process model, and bivariate inverse Gaussian process model. Bivariate degradation processes model based on Wiener process and copula function has been studied by Pan et al. [45], Wang et al. [52], and Jin and Matthews [53]. In detail, Pan et al. [45] presented a Bayesian method based on Markov chain Monte Carlo method to facilitate parameter estimation and reliability assessment. Wang et al. [52] further introduced a Bayesian method for residual life estimation under this kind of bivariate degradation process model. Jin and Mathews [53] introduced a method for degradation test planning and measurement plan optimization for products modelled by the bivariate Wiener process and copula function.

Bivariate degradation process model based on gamma process and copula function has been investigated by Pan and Sun [54], Hong et al. [55], and Wang et al. [46]. Wang et al. [46] introduced a two-stage method to estimate the parameter of a bivariate non-stationary gamma degradation process, for which the residual life estimation can be implemented in an adaptive manner. Pan and Sun [54] presented a method for step-stress accelerated degradation test planning under the bivariate degradation model based on gamma process and copula function. Hong et al. [55] investigated the condition-based maintenance optimization for products

with dependent components deteriorations, where the dependent deteriorations were characterized by gamma process and copula function.

Bivariate degradation process models based on inverse Gaussian process and copula function have been studied by Liu et al. [56] and Peng et al. [13]. Within these studies, Liu et al. [56] incorporated time scale transformation and random drift into the bivariate inverse Gaussian process model to account for the nonlinear of degradation process and heterogeneity within a product population. Peng et al. [13] introduced a two-stage Bayesian parameter estimation and reliability assessment method to deal with the incomplete degradation observations for products modelled by the bivariate inverse Gaussian process and copula function.

According to the literature review presented above, it can be found that most of the models are introduced for bivariate degradation processes. Limited exceptions are the works presented by Peng et al. [15], Pan and Sun [54] and Wang et al. [57]. Although the copula based model given in Eq. 5.14 is flexible for constructing multivariate degradation process models, the study on copula based multivariate degradation process model is still limited and deserves more investigation on both the utilization of multivariate copula function and the applications. There is also a strong imperative to advance investigation on residual life assessment, degradation test planning, optimal maintenance decision based on the copula based multivariate degradation process models.

5.4 Multiple Degradation Processes Under Dynamic Environments

The model for the degradation process under dynamic environments is aimed to deal with the two critical aspects highlighted above for degradation modelling of complex systems under dynamic environments. Various methods and models have been summarized above to deal with the two critical aspects, which include the modelling of multiple degradation process and the incorporation of dynamic environmental effect. However, the research on multiple degradation modelling for complex system under dynamic environments has not been well studied. There are generally two types of models having been introduced, which are the multiple degradation process and random shock models [5, 58, 59], and the multiple degradation process and dynamic covariate model [15].

5.4.1 Multiple Degradation Process and Random Shock Models

The multiple degradation process and random shock models have been introduced for the situation where a system is subject to degradation processes and random

shocks [5]. The degradation processes are associated with the inherent failure mechanism of the system, and the random shocks are related to the exterior environmental effects. The dependence among degradation processes can be either from the inherent dependent failure mechanism of the system, or from the exterior effect of the random shocks, or both. The inherent dependent failure mechanism of the system is characterized through the copula function [5], which is similar to the copula based multivariate degradation process model. The random shocks introduced dependence is characterized through instantaneous degradation increments [58, 59] or degradation rate acceleration or both [5], where the effect of random shocks is incorporated.

Suppose a product has L degradation processes $Y_l(t)$ with $l = 1, \ldots, L$, and the arrival of random shocks experienced by the product follows a Poisson process $N(t)$. To facilitate the derivation, a simple model for the degradation process $Y_l(t)$ is often used, such as a multiplicative path function as $Y_l(t) = X_l \eta_l(t)$ with X_l being a random variable [5]. The effect of the random shocks on the degradation processes is characterized into two types, i.e., (1) the cumulative degradation increments as $S_l(t) = \sum_{k=1}^{N(t)} \omega_{lk}$ with ω_{lk} denoting the instantaneous degradation increment introduced by the kth random shock, and (2) the degradation rate acceleration, which is incorporated into $Y_l(t)$ through the idea of accelerated degradation modelling, such as the scaling of t into $te^{G_l(t)}$ with $G_l(t) = \lambda_1 N(t) + \lambda_2 \sum_{k=1}^{N(t)} \omega_k$ [5]. By considering both effects of the random shocks, the degradation process model of the lth degradation process is presented as

$$M_l(t) = Y_l\left(t; G_l(t)\right) + S_l(t) \tag{5.20}$$

The marginal distribution of the degradation process $M_l(t)$ can be derived based on the model for the degradation process $Y_l(t)$, the model for random shocks $N(t)$, and the model for the instantaneous degradation increment ω_{lk}. For instance, assume $Y_l(t) = X_l \eta_l(t)$ with $F_{X_l}(x_l)$ being the distribution of random parameter X_l, $N(t)$ follows a homogeneous Poisson process with occurring rate λ, and ω_{lk} follows an exponential distribution with mean μ_l, the distribution of $M_l(t)$ can be derived as

$$\begin{aligned} F_l\left(m_l(t)\right) &= \Pr\left(M_l(t) < m_l(t)\right) = \Pr\left(N(t) = 0\right) \Pr\left(Y_l(t) < m_l(t)\right) \\ &\quad + \sum_{n=1}^{+\infty} \Pr\left(N(t) = n\right) \Pr\left(Y_l\left(t; G_l(t)\right) + S_l(t) < m_l(t) | N(t) = n\right) \\ &= \mathrm{e}^{-\lambda t} F_{X_l}\left(\frac{m_l(t)}{\eta_l(t)}\right) + \sum_{n=1}^{+\infty} \frac{\mathrm{e}^{-\lambda t}(-\lambda t)^n}{n!} \int_{u=0}^{m_l(t)} F_{X_i}\left(\frac{m_l(t) - u}{\eta_l\left(t\mathrm{e}^{G_l(t)}\right)}\right) \frac{u^{n-1}\mathrm{e}^{-\frac{u}{\mu_l}}}{\Gamma(n)\mu_l^n} du \end{aligned} \tag{5.21}$$

The dependency among $\mathbf{M}(t) = (M_1(t), \ldots, M_L(t))$ mainly originates from two parts, which are the dependency introduced by the dependence of the inherent

degradation processes $\mathbf{Y}(t) = (Y_1(t), \ldots, Y_L(t))$, and the dependency originated from the series of random shocks. Assume the instantaneous degradation increments ω_{lk} with $l = 1, \ldots, L$ and $k = 1, \ldots, N(t)$ are independent, the dependence among $\mathbf{M}(t)$ can be characterized through the joint distribution of $\mathbf{M}(t)$ as

$$\begin{aligned} &F(m_1(t), \ldots, m_L(t)) = \Pr(M_1(t) < m_1(t), \ldots, M_L(t) < m_L(t)) \\ &= \sum_{n=0}^{+\infty} \mathbf{C}(F_1(m_1(t)|N(t) = n), \ldots, F_L(m_L(t)|N(t) = n)) \Pr(N(t) = n) \end{aligned} \tag{5.22}$$

where $\mathbf{C}(F(m_1(t)|N(t) = n), \ldots, F(m_L(t)|N(t) = n))$ is a copula function used to model the dependence originated from the inherent dependence of degradation processes $\mathbf{Y}(t)$, and $F_l(m_l(t)|N(t) = n)$ is the marginal distribution of degradation process $M_l(t)$ condition on $N(t) = n$, which can be derived similarly to Eq. 5.22.

Given the degradation thresholds of the degradation processes, the reliability of the product is given as

$$R(t) = \Pr\{M_1(t) < D_1, \ldots, M_L(t) < D_L\} = F(D_1(t), \ldots, D_L(t)) \tag{5.23}$$

By substituting the degradation thresholds into Eq. 5.22, the reliability function of the product can be obtained. However, it is often difficult to obtain an analytical solution for the reliability function. Wang and Pham [5] derive the reliability bounds for the reliability of product with bivariate degradation processes and random shocks. A two-stage parameter inference method and a comparison of the reliability under the model introduced above with constant copulas and time-varying copulas have also been studied by Wang and Pham [5]. Song et al. [58] investigated the reliability of multi-component systems with multiple degradation processes and random shocks, where the model presented above is simplified without considering the effect of degradation rate acceleration by the random shocks. The maintenance modeling and optimization under the multiple degradation process and random shock model has been studied by Song et al. [58]. Song et al. [59] further extended the model introduced above into a more advanced model, where the dependency of transmitted shock sizes to hard failure process, and shock damages to specific degradation processes (soft failure processes) for all components have been studied.

Compared with the degradation process model under static environments, the multiple degradation process and random shock model introduced in this section successfully incorporates the effect of dynamic environments through the modeling of the random shocks and the dependency caused by the effect of random shocks. However, more assumptions need to be assumed to derive the model, which requires a deep understanding of the failure mechanism of the degradation processes, the arrival of random shocks, and their influences on the degradation processes as well. A limit on the models used for the marginal degradation processes is necessary to facilitate the calculation of the reliability as given in Eq. 5.23, where a multiplicative path function or a degradation path model is generally applicable. In addition,

methods for parameter estimation, degradation analysis and residual life prediction with this kind of multiple degradation process and random shock model has not been sufficiently studied.

5.4.2 Multiple Degradation Process and Dynamic Covariate Models

The multiple degradation process with dynamic covariate model is introduced by utilizing the idea of dynamic covariate and copula function. The dynamic covariate is used to incorporate the dynamic environments into degradation process models. The copula function is used to model the dependency among the degradation processes. This kind model is similar to the copula based multivariate degradation process models introduced in Sect. 5.3.4. Major difference is that the marginal degradation processes used in Sect. 5.3.4 are substituted with marginal degradation processes with dynamic covariates.

Suppose a product has L degradation processes $Y_l(t)$ with $l = 1, \ldots, L$, and the dynamic environments experienced by the product are summarized into the external factors $\mathbf{X}^E$. Under the effect of the external factors, the marginal degradation processes $Y_l(t)$ are modelled using stochastic process models with dynamic covariates. Following the methods and models summarized in Sect. 5.2, the baseline stochastic process models $Y_l(t)$ are modified into $Y_l(t; \mathbf{X}^E)$, such as the Wiener process model $Y_l(t) = \eta_l(t) + \sigma_l \mathsf{B}(t)$ with $\mathsf{B}(t)$ being a standard Brownian motion and the inverse Gaussian process model $Y_l(t) \sim \text{IG}\left(\Lambda_l(t), \lambda_l \Lambda_l^2(t)\right)$ with $\Lambda_l(t)$ being a nonnegative and monotonically increasing function are separately modified into $Y_l(t; \mathbf{X}^E) = \eta_l(t; \mathbf{X}^E) + \sigma_l \mathsf{B}(t)$ and $Y_l\left(t; \mathbf{X}^E\right) \sim \text{IG}\left(\Lambda_l\left(t; \mathbf{X}^E\right), \lambda_l \Lambda_l^2\left(t; \mathbf{X}^E\right)\right)$.

Let $\Delta Y_l(t_j; \mathbf{X}^E) = Y_l(t_j; \mathbf{X}^E) - Y_l(t_{j-1}; \mathbf{X}^E)$ denote the degradation increment within the time interval $[t_{j-1}, t_j]$, which follows a probability distribution with CDF $F_l(\Delta y_l(t_j); \mathbf{X}^E)$ under the specific stochastic process chosen for the marginal degradation process. The dependence among the degradation processes is characterized through their degradation increments. It is assumed that the degradation increments $\boldsymbol{\Delta} \mathbf{Y}_l(t_j) = (\Delta Y_1(t_j), \ldots, \Delta Y_L(t_j))$ for the L performance indicators within the same time interval $[t_{j-1}, t_j]$ are dependent. Similar to the copula based multivariate process model presented in Sect. 5.3.4, the joint probability distribution of $\boldsymbol{\Delta} \mathbf{Y}_l(t_j; \mathbf{X}^E)$ within the same time interval is characterized by leveraging a multivariate copula function as

$$F\left(\boldsymbol{\Delta} \mathbf{y}\left(t_j\right); \mathbf{X}^E\right) = \text{C}\left(F_1\left(\Delta y_1\left(t_j\right); \mathbf{X}^E\right), \ldots, F_L\left(\Delta y_L\left(t_j\right); \mathbf{X}^E\right); \boldsymbol{\theta}^{\text{Cop}}\right) \tag{5.24}$$

where $\boldsymbol{\Delta} \mathbf{y}(t_j) = (\Delta y_1(t_j), \ldots, \Delta y_L(t_j))$ and $\text{C}(u_1, \ldots, u_L; \boldsymbol{\theta}^{\text{Cop}})$ is a L-dimensional multivariate copula function with parameters $\boldsymbol{\theta}^{\text{Cop}}$ and $u_l \sim \text{Uniform}(0, 1)$.

The multiple degradation process and dynamic covariate model, which is based on the modified marginal degradation process and copula function is constructed as

$$Y_l\left(t_j;\mathbf{X}^E\right)=\sum_{k=2}^{j}\Delta Y_l\left(t_k;\mathbf{X}^E\right),l=1,\ldots,L,j=2,\ldots,+\infty,$$
$$\begin{cases}\Delta Y_l\left(t_k;\mathbf{X}^E\right)\sim \mathrm{PRO}\left(\Delta y_l\left(t_k\right);\mathbf{X}^E,\boldsymbol{\theta}_l^{\mathrm{Mar}}\right)\\ F\left(\boldsymbol{\Delta}\mathbf{y}\left(t_j\right);\mathbf{X}^E\right)=\mathrm{C}\left(F_1\left(\Delta y_1\left(t_j\right);\mathbf{X}^E\right),\ldots,F_L\left(\Delta y_L\left(t_j\right);\mathbf{X}^E\right);\boldsymbol{\theta}^{\mathrm{Cop}}\right)\end{cases}\tag{5.25}$$

where $\mathrm{PRO}\left(\Delta y_1\left(t_k\right);\mathbf{X}^E,\boldsymbol{\theta}_l^{\mathrm{Mar}}\right)$ is the distribution of degradation increment of the lth performance indicator with the external factors incorporated.

Given the degradation thresholds of the L performance indicators, the reliability function of the product is given as

$$R(t)=\Pr\left\{\sup_{s\le t}Y_1\left(s;\mathbf{X}^E\right)<D_1,\ldots,\sup_{s\le t}Y_l\left(s;\mathbf{X}^E\right)<D_l,\ldots,\sup_{s\le t}Y_L\left(s;\mathbf{X}^E\right)<D_L\right\}\tag{5.26}$$

It is generally to obtain the analytical solution to the reliability function, simulation based method is needed to implement the calculation. To facilitate the degradation analysis of complex systems under a dynamic environment with the multiple degradation process and dynamic covariate model, Peng et al. [15] introduced a Bayesian parameter estimation method and a simulation based degradation inference, reliability assessment and residual life prediction method. However, their method is limited to the situation that the external factors $\mathbf{X}^E$ are assumed available during the degradation analysis, where no probabilistic model is constructed for these external factors. A more general model with external factors modelled as random variables or stochastic processes deserves further investigation.

Compared with the multiple degradation process and random shocks model, the model based on dynamic covariates and copula function is more flexible for incorporating external factors and characterizing degradation processes. There are various types of stochastic models, covariate models and copula functions available for constructing the multivariate degradation process model as given in Eq. 5.25. The Bayesian parameter estimation method and degradation analysis method presented by Peng et al. [15] can also be extended to the model with more general assumptions. More advanced methods for parameter estimation and residual life predication still deserve further investigation. In addition, the study on maintenance modeling and optimization and system health management under this kind of multivariate degradation process model has not been presented yet.

5.5 Conclusions

In this paper, a summary of the state of arts on the researches of reliability modelling of complex system under dynamic environments is presented by highlighting two critical aspects, i.e., (1) modelling multiple degradation processes, and (2) characterizing dynamic environments effects. We mainly focused on the various

types of multivariate degradation process models because these models are critical for reliability modelling of modern complex system with multiple performance indicators or composed by multiple components. In addition, the characterization of dynamic environments and the methods for incorporating dynamic environments into reliability models has been discussed due to the consideration that the effects of environments are often simplified or omitted in general degradation modelling. Through these two aspects, the paper is organized into two progressive sessions, i.e., multiple degradation processes under static dynamic environments, and multiple degradation processes under dynamic environments.

For the multiple degradation processes under static environments, the multivariate degradation process models based on multivariate Gaussian distribution and multivariate Birnbaum-Saunders distribution, the degradation rate interaction model, and the models based on copula function and stochastic processes are reviewed. Among these models, the copula based multivariate degradation process model has the capability of modelling various types of degradation processes while keeping the model simple enough for model construction and parameter estimation. However, limited by the utilization of multivariate copula function in reliability engineering, most of the models are bivariate degradation process model. More research is needed to extend the research on multiple degradation processes under dynamic environments.

For the multiple degradation processes under dynamic environments, the multiple degradation process and random shocks models and the multiple degradation processes and dynamic covariates models are summarized. There is not too much research published on this topic, partially due to the limitation of real case examples, the unavailability of physical mechanism, and unjustifiability of complex models. Within the proposed model, the multiple degradation process with random shocks model is basically an extension of the degradation-threshold-shock model, which has been studied for one-dimensional degradation process. The multiple degradation process with dynamic covariate model is a combination of the copula based multivariate degradation model and dynamic covariate models, which has been investigated extensively for one-dimensional degradation process and failure time models. There is a strong imperative for more investigation on multivariate degradation modelling with dynamic environments incorporated. In addition, methods for parameter estimation, residual life predication, degradation test planning, maintenance strategy optimization, and model comparison and selection, are major open area deserving extensive investigations under the multivariate degradation modelling with dynamic environments highlighted.

Acknowledgements This project is supported by National Natural Science Foundation of China (Grant No. 51605081).

References

1. Bennett R, Booker JM, Keller-McNulty S, Singpurwalla ND (2003) Testing the ustestable: reliability in the 21st century. IEEE T Reliab 52:118–124
2. Meeker WQ, Hong Y (2014) Reliability meets big data: opportunities and challenges. Qual Eng 26:102–116. doi:10.1080/08982112.2014.846119
3. Ye ZS, Xie M (2014) Stochastic modelling and analysis of degradation for highly reliable products. Appl Stoch Model Bus Ind 31:16–32. doi:10.1002/asmb.2063
4. Pan Z, Balakrishnan N (2011) Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. Reliab Eng Syst Saf 96:949–957. doi:10.1016/j.ress.2011.03.014
5. Wang Y, Pham H (2012) Modeling the dependent competing risks with multiple degradation processes and random shock using time-varying copulas. IEEE T Reliab 61:13–22. doi:10.1109/TR.2011.2170253
6. Kharoufeh JP, Cox SM, Oxley ME (2013) Reliability of manufacturing equipment in complex environments. Ann Oper Res 209:231–254. doi:10.1007/s10479-011-0839-x
7. Liao H, Tian Z (2013) A framework for predicting the remaining useful life of a single unit under time-varying operating conditions. IIE Trans 45:964–980. doi:10.1080/0740817X.2012.705451
8. Wang D, Miao Q, Pecht M (2013) Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. J Power Sources 239:253–264. doi:10.1016/j.jpowsour.2013.03.129
9. Bian L, Gebraeel N (2014) Stochastic modeling and real-time prognostics for multi-component systems with degradation rate interactions. IIE Trans 46:470–482. doi:10.1080/0740817X.2013.812269
10. Si XS, Hu CH, Kong X, Zhou DH (2014) A residual storage life prediction approach for systems with operation state switches. IEEE T Ind Electron 61:6304–6315. doi:10.1109/TIE.2014.2308135
11. Ye ZS, Chen N, Shen Y (2015) A new class of Wiener process models for degradation analysis. Reliab Eng Syst Saf 139:58–67. doi:10.1016/j.ress.2015.02.005
12. Hong Y, Duan Y, Meeker WQ (2015) Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. Technometrics 57:180–193. doi:10.1080/00401706.2014.915891
13. Peng W, Li YF, Yang YJ, Zhu SP, Huang HZ (2016) Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas. IEEE T Reliab 65:624–639. doi:10.1109/TR.2015.2513038
14. Xu X, Li Z, Chen N (2016) A hierarchical model for lithium-ion battery degradation prediction. IEEE T Reliab 65:310–325. doi:10.1109/TR.2015.2451074
15. Peng W, Li YF, Mi J, Yu L, Huang HZ (2016) Reliability of complex systems under dynamic conditions: a Bayesian multivariate degradation perspective. Reliab Eng Syst Saf 153:75–87. doi:10.1016/j.ress.2016.04.005
16. Singpurwalla ND (1995) Survival in dynamic environments. Stat Sci 10:86–103
17. Lee Ting ML, Whitmore GA (2006) Threshold regression for survival analysis: modelling event times by a stochastic process reaching a boundary. Stat Sci 21:501–513
18. Bae SJ, Kuo W, Kvam PH (2007) Degradation models and implied lifetime distributions. Reliab Eng Syst Saf 92:601–608
19. Si XS, Wang W, Hu CH, Zhou DH (2011) Remaining useful life estimation – a review on the statistical data driven approaches. Eur J Oper Res 213:1–14
20. Yashin AI, Manton KG (1997) Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. Stat Sci 12:20–34
21. Bagdonavičius V, Nikulin M (2000) Estimation in degradation models with explanatory variables. Lifetime Data Anal 7:85–103

22. Bagdonavičius V, Nikulin M (2009) Statistical models to analyze failure, wear, fatigue, and degradation data with explanatory variables. Commun Stat – Theory M 38:3031–3047
23. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
24. Kharoufeh JP, Solo CJ, Ulukus MY (2010) Semi-markov models for degradation-based reliability. IIE Trans 42:599–612
25. Meeker WQ, Escobar LA, Lu J (1998) Accelerated degradation tests: modeling and analysis. Technometrics 40:89–99. doi:10.1080/00401706.1998.10485191
26. Park C, Padgett WJ (2006) Stochastic degradation models with several accelerating variables. IEEE Trans Reliab 55:379–390
27. Bae SJ, Kvam PH (2004) A nonlinear random-coefficients model for degradation testing. Technometrics 46:460–469
28. Liao HT, Elsayed EA (2006) Reliability inference for field conditions from accelerated degradation testing. Nav Res Logist 53:576–587
29. Bian L, Gebraeel N, Kharoufeh JP (2015) Degradation modeling for real-time estimation of residual lifetimes in dynamic environments. IIE Trans 47:471–486. doi:10.1080/0740817X.2014.955153
30. Chan V, Meeker WQ (2001) Estimation of degradation-based reliability in outdoor environments. Stat Prepr 25:1–31
31. Ye ZS, Chen N (2014) The inverse Gaussian process as a degradation model. Technometrics 56:302–311. doi:10.1080/00401706.2013.830074
32. Peng W, Li YF, Yang YJ, Huang HZ, Zuo MJ (2014) Inverse Gaussian process models for degradation analysis: a Bayesian perspective. Reliab Eng Syst Saf 130:175–189. doi:10.1016/j.ress.2014.06.005
33. Wang L, Pan R, Li X, Jiang T (2013) A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliab Eng Syst Saf 112:38–47. doi:10.1016/j.ress.2012.09.015
34. Ye ZS, Hong Y, Xie Y (2013) How do heterogeneities in operating environments affect field failure predictions and test planning? Ann Appl Stat 7:2249–2271. doi:10.1214/13-AOAS666
35. Kharoufeh JP, Cox SM (2005) Stochastic models for degradation-based reliability. IIE Trans 37:533–542
36. Flory JA, Kharoufeh JP, Gebraeel NZ (2014) A switching diffusion model for lifetime estimation in randomly varying environments. IIE Trans 46:1227–1241
37. Wang P, Coit DW (2004) Reliability prediction based on degradation modeling for systems with multiple degradation measures. Reliab Maintainab Annu Symp – RAMS 2004:302
38. Sari JK, Newby MJ, Brombacher AC, Tang LC (2009) Bivariate constant stress degradation model: LED lighting system reliability estimation with two-stage modelling. Qual Reliab Eng Int 25:1067–1084
39. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes. Lifetime Data Anal 11:511–527
40. Pan Z, Feng J, Sun Q (2016) Lifetime distribution and associated inference of systems with multiple degradation measurements based on Gamma processes. Eksploat Niezawodn 18:307–313
41. Kundu D, Balakrishnan N, Jamalizadeh A (2010) Bivariate Birnbaum-Sanders distribution and associated inference. J Multivar Anal 101:113–125
42. Pan Z, Sun Q, Feng J (2016) Reliability modeling of systems with two dependent degrading components based on gamma processes. Commun Stat A-Theor 45:1923–1938
43. Bian L, Gebraeel N (2014) Stochastic framework for partially degradation systems with continuous component degradation-rate-interactions. Nav Res Logist 61:286–303
44. Nelsen RB (2006) An introduction to copulas. Springer, New York
45. Pan Z, Balakrishnan N, Sun Q, Zhou J (2013) Bivariate degradation analysis of products based on Wiener processes and copulas. J Stat Comput Simul 83:1316–1329
46. Wang X, Balakrishnan N, Guo B, Jiang P (2015) Residual life estimation based on bivariate non-stationary gamma degradation process. J Stat Comput Simul 85(2):405–421

47. Wu S (2014) Construction of asymmetric copulas and its application in two-dimensional reliability modelling. Eur J Oper Res 238:476–485
48. Huard D, Evin G, Favre AC (2006) Bayesian copula selection. Comput Stat Data Anal 51: 809–822
49. Jaworski PJ, Durante F, Härdle WK, Rychlik T (2010) Copula theory and its applications. Springer, Heidelberg
50. Frahm G, Junker M, Szimayer A (2003) Elliptical copulas: applicability and limitations. Stat Probab Lett 63:275–286
51. Kurowicka D, Joe H (2011) Dependence modeling: vine copula handbook. World Scientific Publishing, Singapore
52. Wang X, Guo B, Cheng ZJ (2014) Residual life estimation based on bivariate Wiener degradation process with time-scale transformations. J Stat Comput Simul 84:545–563
53. Jin G, Matthews D (2014) Measurement plan optimization for degradation test design based on the bivariate Wiener process. Qual Reliab Eng Int 30:1215–1231
54. Pan Z, Sun Q (2014) Optimal design for step-stress accelerated degradation test with multiple performance characteristics based on gamma processes. Commun Stat B-Simul 43:298–314
55. Hong HP, Zhou W, Zhang S, Ye W (2014) Optimal condition-based maintenance decisions for systems with dependent stochastic degradation of components. Reliab Eng Syst Saf 121: 276–288
56. Liu Z, Ma X, Yang J, Zhao Y (2014) Reliability modeling for systems with multiple degradation processes using inverse Gaussian process and copulas. Math Probl Eng 2014:829597-1–10
57. Wang X, Balakrishnan N, Guo B (2015) Residual life estimation based on nonlinear-multivariate Wiener processes. J Stat Comput Simul 85:1742–1764
58. Song S, Coit DW, Feng Q, Peng H (2014) Reliability analysis for multi-component systems subject to multiple dependent competing failure processes. IEEE Trans Reliab 63:331–345
59. Song S, Coit DW, Feng Q (2016) Reliability analysis of multiple-component series systems subject to hard and soft failures with dependent shock effects. IIE Trans 48:720–735

Chapter 6
A Survey of Modeling and Application of Non-destructive and Destructive Degradation Tests

Chih-Chun Tsai, Chien-Tai Lin, and N. Balakrishnan

Abstract These days, most products are highly reliable which makes it very difficult or even impossible to obtain failure data on such products within a reasonable period of time prior to product release. Degradation tests are one way to overcome this obstacle by collecting degradation data (measurement of degradation) on such products. Based on different measurement processes, degradation tests can be divided into non-destructive and destructive degradation tests. In this chapter, we discuss a number of these two types of degradation models that have been developed in the literature to describe the degradation paths of products. In addition, some applications of degradation models of these two classes are also discussed.

Keywords Highly reliable products • Quality characteristics • Degradation data • Linearized and nonlinear degradation paths

6.1 Introduction

As technology advances, the consumers pay more attention to quality of products that they purchase and so have higher requirements on the reliability of products. Due to strong pressure from markets, manufacturers are often required to provide information on the reliability of their products (such as the mean-time-to-failure or the 100pth percentile of the lifetime distribution) to their customers. However, for highly reliable products, it is often difficult to obtain the product's lifetime through traditional life tests within a reasonable period of time. Methods such as accelerated life tests (ALT) or step-stress accelerated life tests (SSALT) have been developed in reliability analysis for this reason. They are meant to expedite failures

C.-C. Tsai • C.-T. Lin (✉)
Department of Mathematics, Tamkang University, Tamsui, Taiwan
e-mail: chihchuntsai@mail.tku.edu.tw; chien@mail.tku.edu.tw

N. Balakrishnan
Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada
e-mail: bala@mcmaster.ca

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_6

during test intervals by stressing or step-stressing the product (for example, at elevated temperatures or voltages) beyond its normal use conditions. These methods indeed help to obtain in more information if the link between the accelerated test environment and the regular use environment is known.

In many situations, it is difficult to perform ALT or SSALT, especially when just few or no failures are recorded, thus providing little or no information about reliability. In such cases, if there exists a quality characteristic (QC) whose degradation over time can be related to reliability of the product, then the product's lifetime can be estimated well through the use of such degradation data. Based on measurement processes, different degradation tests have been developed in the literature and they can be broadly classified into two categories: non-destructive and destructive degradation tests.

In the case of non-destructive degradation tests, one is able to directly measure the degradation of a physical or mechanical characteristic over time as, for example, with the wear of brake pads, propagation of crack size, or degradation of a performance characteristic over time such as the voltage of a battery or the luminous flux of a light emitting diodes (LED) bulb or lamp. The degradation measurements can be taken on each test unit over time during the degradation test. Hence, regression models or stochastic processes are more commonly applied to analyze this type of degradation data.

In contrast, for some components or materials, degradation measures might not be possible without destructive measurement techniques (e.g., destructive strength tests) or disruptive measurement techniques (e.g., disassembly and reassembly of a device) that have the potential to affect the subsequent performance of the product; therefore, only one degradation measurement becomes possible in such cases. As a result, degradation measurements for multiple samples are required at different points of time in order to develop inference in this case. A test with such degradation data is called a destructive degradation test. In practice, destructive degradation tests are often conducted in order to gain insight into the structural performance or material behavior of products over time. Hence, it is quite suitable for objects which will be mass-produced and the cost of destruction of units becomes economical in this case. A typical example is dielectric breakdown test. During the test, test voltage that exceeds normal operating condition is applied to an electrical component such as circuit boards, appliances and electric motors. The test voltage is then increased until the insulating material fails or breaks down. Such a test can be used to determine if a component's insulation is adequate enough to protect the user from an electric shock.

In this chapter, we will provide a brief background about the models used in the literature for estimating reliability using either degradation data from non-destructive or destructive degradation tests. Models of the first type include fixed or random effect, Wiener process, gamma process, inverse Gaussian (IG) process, and some other degradation models. Next, we survey the destructive degradation models with linearized and nonlinear degradation paths. Finally, applications of these models to burn-in, maintenance as well as acceptance tests are described.

The rest of this chapter is organized as follows. In Sect. 6.2, we describe non-destructive degradation models. Section 6.3 presents linearized and nonlinear destructive degradation models. In Sect. 6.4, we describe some applications and modeling these two types of degradation data. Finally, some concluding remarks are made in Sect. 6.5.

6.2 Nondestructive Degradation Model

In the literature, two major aspects of modeling for non-destructive degradation data have been considered. One approach is to consider general statistical models. Degradation in these models is then modeled by a function of time and some possibly multidimensional random variables. These models are called general degradation path models with fixed and random effects. An alternative approach is to use a stochastic process to describe the degradation path. The advantage of random effect formulation is that it can account for the heterogeneity of degradation paths. Moreover, stochastic processes can adequately describe time-dependent variations in degradation data. However, when the unit-to-unit variations are not significant, and variations in degradation data can not completely described by time-dependent stochastic processes, neither formulation is suitable for use. For this reason, several authors have considered the inclusion of both random effects and stochastic process formulations in their degradation model. We now review some existing models and their applications.

6.2.1 Fixed or Random Effect Degradation Model

Motivated by a fatigue-crack-growth data, Lu and Meeker [26] used a general random effect model to describe the degradation paths, which is given by

$$L_i(t_j) = f(t_j; \boldsymbol{\Theta}_i, \boldsymbol{\phi}) + \varepsilon_{ij}, \quad i = 1, \ldots, n, j = 1, \ldots, m_{\boldsymbol{\Theta}_i} \leq m, \tag{6.1}$$

where $L_i(t_j)$ and $f(t_j; \boldsymbol{\Theta}_i, \boldsymbol{\phi})$ denote the sample and actual degradation paths of the ith unit at time t_j, respectively; $\boldsymbol{\Theta}_i$ is the vector of the ith unit random-effect parameters, representing individual unit characteristics; $\boldsymbol{\phi}$ is the vector of fixed-effect parameters, common for all tested units; ε_{ij} is the measurement error with constant variance σ_ε^2; $\boldsymbol{\Theta}_i$ and ε_{ij} are independent of each other ($i = 1, 2, \ldots, n; j = 1, 2, \ldots, m_{\boldsymbol{\Theta}_i}$); m is the total number of possible inspections in the experiment; and $m_{\boldsymbol{\Theta}_i}$ is the total number of inspections on the ith unit, a function of $\boldsymbol{\Theta}_i$. They assumed that $\boldsymbol{\Theta}_i$ follows a multivariate distribution function $G_{\boldsymbol{\Theta}}(\cdot)$, and that the product's lifetime T can be suitably defined as actual degradation paths $f(t_j; \boldsymbol{\Theta}_i, \boldsymbol{\phi})$ crossing a critical value ω. Then, the cumulative distribution function (cdf) of T, the failure time, can be written as

$$F_T(t) = P(T \leq t) = F_T(t; \boldsymbol{\phi}, \omega, G_{\boldsymbol{\Theta}}(\cdot)).$$

Since the parameters appear nonlinearly in the path model, full maximum likelihood estimation (MLE) of random-effect parameters is, in general, algebraically intractable and computationally quite difficult. So, they proposed a two-stage method to estimate the model parameters, and applied a bias-corrected percentile parametric bootstrap (or simulation) method to obtain pointwise confidence intervals for $F_T(t)$.

Tseng and Yu [66] proposed an on-line and real-time termination rule to determine an appropriate stopping time of a degradation experiment for LED data. They used a linearized degradation model to describe the degradation paths as follows:

$$\ln(\ln(L_i(t))) = \ln(\alpha_i) + \beta_i \ln(t) + \varepsilon_i(t), \quad t > 0,$$

where $L_i(t)$ denotes the degradation path of the ith device and $\varepsilon_i(t)$ is the measurement error of device i at time t.

Yu and Tseng [89] dealt with the optimal design of the inspection frequency, the sample size, and the termination time for a degradation experiment. Based on real applications to LED lamps, they considered the degradation model given by

$$\ln(L_i(t)) = -\beta t^{\alpha} + \varepsilon, \quad t > 0, \tag{6.2}$$

where α is fixed effect parameter and β is a random effect parameter following a lognormal distribution. It can be seen that the product's lifetime T follows a lognormal distribution as well. Under the constraint that the total experimental cost does not exceed a pre-fixed budget, they determined the optimal decision variables by minimizing the variance of the estimated $100p$th percentile of the lifetime distribution of the product.

There are several applications relating to the models in (6.1) and (6.2). Wu and Chang [78] used the general degradation path model in (6.1) to determine the number of units to test, inspection frequency, and termination time of a degradation test under a determined cost of experiment such that the variance of the estimator of a percentile of the failure time distribution is minimum. Yu and Chiao [88] proposed a systematic approach to the identification problem with respect to fractional factorial degradation experiments under the model in (6.2). By considering the criterion of minimizing the total cost of experiment, they obtained the optimal test plan of inspection frequency, the sample size, and the termination time at each run by solving a nonlinear integer programming model such that the correct probability of identifying these factors exceeds a pre-specified level. Yu [86] followed the work of Yu and Tseng [89] to discuss the optimal test plan for an accelerated degradation test (ADT) using degradation model in (6.2). In his work, the location parameter of the random effect is assumed to have a relationship with the stress level, and the criterion for optimal combination of the sample size, inspection frequency and the

termination time at each stress level is to minimize the mean-squared error of the estimated $100p$th percentile of the product's lifetime distribution at the use condition subject to the constraint that the total experimental cost does not exceed a pre-fixed budget. Yu [87] further extended this work to develop optimal ADT based on a linear degradation path model with a reciprocal Weibull-distributed rate.

The nonlinear random-coefficients modeling, which is more appropriate in handling complicated forms of degradation paths, is appealing and widely adopted by many authors. Bae et al. [2] used a bi-exponential model with random coefficients to describe the nonlinear degradation paths caused by nano-contamination in plasma display panels. Liu and Tang [24] developed a single-path power-law statistical degradation model with nonlinear stress-life relationships, and formulated a Bayesian optimal ADT design problem by minimizing the expected pre-posterior variance of a quantity of interest at use condition. Park and Bae [31] used the nonlinear random-coefficients model for estimating lifetime distributions at use condition from ADT data sets including a real application of commercial organic light-emitting diodes (OLED). They specifically proposed three methods, namely, delta approximation, multiple imputation of failure-times, and the lifetime distribution-based method, to develop direct inference on the lifetime distribution without invoking arbitrary assumptions on the degradation model. Pan and Crispin [28] considered both the unit-to-unit variability and measurement variability to develop a hierarchical approach to construct degradation model. Kim and Bae [17] suggested a cost-effective optimal degradation test plan for the degradation data of alloy-A fatigue cracks and plasma display panels by minimizing the total experimental cost in the context of a nonlinear random-coefficients model. Guida et al. [13] proposed an empirical random-effects regression model of polynomial type for describing the three components of variability (voltage decay over time for each single cell, voltage variability among cells, and noise variability) observed in long-term solid oxide fuel cells (SOFC) degradation tests. They also discussed inferential procedures for some performance measures, such as the future degradation growth, the reliability function and the cell-to-cell variability in SOFC stacks.

6.2.2 *Stochastic Process Degradation Models*

Stochastic processes are useful for describing products's failure mechanisms when the operation environment is dynamic over time and the products degrade to failure. Let $L(t)$ denote the degradation path of the product, where $t \geq 0$ and $L(0) = 0$. Then, the product's lifetime T can be suitably defined as the first passage time when $L(t)$ crosses a failure threshold level ω; that is,

$$T = \inf\{t | L(t) \geq \omega\}. \tag{6.3}$$

Singpurwalla [47] gave an overview of approaches to failure modeling under the dynamic operating conditions based on stochastic processes. In the following

subsections, we will review some important degradation models based on stochastic processes such as Wiener process, gamma process, and IG process.

6.2.2.1 Wiener Process

Assume that the degradation path of a product, $L(t)$, follows a Wiener process. Then, it can be expressed as

$$L(t) = \eta t + \sigma B(t), \tag{6.4}$$

where η is the drift parameter, $\sigma > 0$ is the diffusion parameter, and $B(\cdot)$ is the standard Brownian motion. It is well-known that the distribution of the first passage time in (6.3) is an IG distribution.

Wiener processes have received wide applications in degradation data analysis. Based on constant-stress ADT (CSADT) data, Liao and Elsayed [19] used model (6.4) to develop a statistical inference procedure to predict field reliability by considering the stress variations. Wang et al. [69] used Bayesian approach to integrate the product's reliability information from both the ADT and field-use conditions. Liao and Tseng [20] dealt with the optimal design in a step-stress ADT (SSADT) experiment. Under the constraint that the total experimental cost does not exceed a pre-fixed budget, they determined the optimal settings of the sample size, measurement frequency, and the termination time by minimizing the asymptotic variance of the estimated $100p$th percentile of the product's lifetime distribution. Tang and Su [51] used the intermediate Wiener degradation data (the first-passage times of the test units over certain nonfailure thresholds during the early stage of a degradation test) to obtain uniformly minimum variance unbiased estimator and uniformly most accurate confidence interval for the mean lifetime. Wang et al. [70] established an adapted Brownian motion-based approach with a drifting parameter to predict the residual life via the expectation-maximization (EM) algorithm.

In contrast to the work of Liao and Tseng [20], Lim and Yum [21] developed the optimal ADT plans under the constant-stress loading by determining the test stress levels and the proportion of test units allocated to each stress level by minimizing the asymptotic variance of the MLE of the pth quantile of the lifetime distribution at the use condition. Tseng et al. [65] discussed the optimal sample size allocation under D-optimality, A-optimality, and V-optimality criteria for an ADT model. They derived analytical results of the optimal sample size allocation formula when the number of stress levels equals 2. Moreover, when the number of stress levels is at least 3, they demonstrated that the optimal sample size allocation should assign test units at the lowest stress and the highest stress levels. Note that, under the Wiener degradation model with a drift parameter being a linear function of the (transformed) stress level, Hu et al. [15] recently proved that a multi-level SSADT plan will degenerate to a simple SSADT plan under many commonly used optimization criteria and with some practical constraints.

Recent applications have focused on the use of threshold information, multi-sensor information, and calibrations for the estimation of failure time distribution and remaining useful life prediction. Based on the cumulative damage data from a two-variable constant-stress loading ADT, Tsai et al. [58] proposed a general inferential procedure, which involves threshold information in the derivations of the MLE of distribution parameters as well as of percentiles, to analytically find an approximate lower confidence bound (LCB) of a lifetime percentile. They assumed that the drift parameter in the Wiener diffusion process can be expressed as an exponential function in terms of two standardized stress variables via a generalized Eyring model (GEM). Wei et al. [75] considered an on-line multi-sensor information based approach to get the remaining useful life prediction analytically with anticipated performance subject to a latent Wiener degradation. They used a distributed Kalman fusion filtering recursively to identify the hidden degradation process, and applied a two-stage method including the EM algorithm simultaneously to update all unknown parameters. Tamborrino [50] considered a Wiener process with exponentially decaying threshold to model the neuronal spiking activity and provided a method to approximate the density of the first passage time. Cui et al. [8] developed two degradation models based on Wiener processes under pre-specified periodical calibrations. They presented stochastic process-based method and partial differential equation method to derive the system reliability and moments of lifetimes of systems.

6.2.2.2 Gamma Process

A gamma process has the following properties:

(1) non-overlapping increments are independent, that is, the increments $\Delta L(t) = L(t + \Delta t) - L(t)$ are independent;
(2) with $\alpha(\cdot)$ being a given, monotone increasing function with $\alpha(0) = 0$, for fixed t and Δt,

$$\Delta L(t) \sim Ga(\beta, \Delta\alpha(t)), \tag{6.5}$$

where $\Delta\alpha(t) = \alpha(t+\Delta t)-\alpha(t)$ and β are shape (function) and scale parameters of the gamma distribution, respectively, and with probability density function (pdf) as

$$f_{\Delta L(t)}(y) = \frac{y^{\Delta\alpha(t)-1}e^{-y/\beta}}{\beta^{\Delta\alpha(t)}\Gamma(\Delta\alpha(t))}, \quad y > 0. \tag{6.6}$$

Because gamma process is strictly increasing, the cumulative distribution function (cdf) of T can be readily expressed as

$$F_T(t) = P(T \leq t) = P(L(t) \geq \omega) = \frac{\Gamma(\alpha(t), \omega/\beta)}{\Gamma(\alpha(t))},$$

where $\Gamma(a, z)$ is the incomplete gamma function defined by

$$\Gamma(a, z) = \int_z^\infty x^{a-1} e^{-x} dx.$$

Motivated by a problem in hydrological engineering, Singpurwalla [48] gave an overview on gamma processes and their generalizations and also described some of their features such as decomposition and representation. Guida et al. [12] suggested a generalization of the non-stationary gamma process, a time discretization of the extended gamma process, to describe age dependent degradation processes with a non-constant variance-to-mean ratio. Wang et al. [73] proposed the change-point gamma and Wiener processes to capture a two-stage degradation process of liquid coupling devices (LCD) and used the Bayesian method to evaluate the real-time reliability of the product. Ling et al. [23] studied the degradation of light intensity of LED based on a time-scale transformed gamma process with power acceleration laws, and also proposed methods for estimating the remaining useful life, given the information on current time and the degradation level. Bordes et al. [3] and Pulcini [39] considered the gamma process perturbed by a Brownian motion as a degradation model.

Much work has been done on the determination of the optimal design for a gamma degradation experiment. Under the constraint that the total experimental cost does not exceed a pre-specified budget, Tseng et al. [64] determined the optimal settings for the sample size, measurement frequency, and termination time for a gamma degradation process by minimizing the approximate variance of the estimated meantime-to-failure (MTTF) of the lifetime distribution of the product. Pan and Balakrishnan [29] discussed multiple-steps SSADT based on gamma processes, in which the stress level is elevated when the degradation value of a product crosses a pre-fixed value. Their work has been further extended by Amini et al. [1] with a different assumption that the stress level is elevated as soon as the measurement of the degradation of one of the test units, at one of the specified times, exceeds the threshold value. Tsai and Lin [52] developed a procedure for selecting the most reliable design among several competing ones when degradation paths of these designs follow gamma processes with common scale parameter. In this regard, they first proposed a selection rule and then derived the optimal test plan for each of the competing designs by minimizing the total experimental cost subject to a constraint that the selection precision is not lower than a pre-fixed level. Chiang et al. [6] and Tsai et al. [59] investigated the optimal ADT strategy with the loadings of two stress variables and GEM.

6.2.2.3 Inverse Gaussian Process

Although the Wiener and gamma processes have received intensive applications in degradation data analysis, it is obvious that the two models cannot handle all degradation data in practice. Another attractive degradation model with monotone paths is the IG process. Compared to the gamma process, the IG process has many interesting and useful properties when dealing with covariates and random effects (Ye and Chen [80]).

Let $L(t)$ be the IG process. Then, it has the properties of non-overlapping independent increments, that is, the increments $\Delta L(t)$ follows an IG distribution $\mathrm{IG}(\mu\Lambda(t), \lambda\Lambda^2(t))$ with pdf

$$f_{\Delta L(t)}(y) = \left(\frac{\lambda\Lambda^2(t)}{2\pi y^3}\right)^{1/2} \exp\left(-\frac{\lambda\Lambda^2(t)(y-\mu\Lambda(t))^2}{2y(\mu\Lambda(t))^2}\right), \quad y > 0, \tag{6.7}$$

where $\Lambda(t)$ is nonnegative and monotone increasing. Then, by the monotonicity property, the cdf of the product's lifetime can be obtained as

$$\begin{aligned} F_T(t) &= P(T \le t) = P(L(t) \ge \omega) \\ &= \Phi\left(\sqrt{\frac{\lambda}{\omega}}\left(\Lambda(t) - \frac{\omega}{\mu}\right)\right) - e^{\frac{2\lambda\Lambda(t)}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{\omega}}\left(\Lambda(t) + \frac{\omega}{\mu}\right)\right), \end{aligned}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution.

Ye et al. [79] investigated the optimal CSADT plan based on IG processes. They determined the optimal stress levels and the number of units allocated to each stress level by minimizing the asymptotic variance of the pth percentile under use conditions. Subsequently, Zhang et al. [92] proposed a mixed IG process to describe the degradation paths of the products. They presented a decision rule for classifying an item as a weak or a normal unit, and used a cost model to determine the optimal burn-in duration and the optimal cut-off level. Peng et al. [37] developed a coherent bivariate degradation analysis with incomplete degradation observations based on IG processes and copula functions. They suggested a two-stage Bayesian estimation method to cope with the parameter estimation problem for both complete and incomplete degradation observations.

6.2.3 Mixed Random Effect and Stochastic Process

6.2.3.1 Random-Effect Wiener Process

Wiener processes with mixed effects have been investigated by many authors. Peng and Tseng [34] investigated the mis-specification analysis of linear degradation models by incorporating the random effects in the drift coefficient and measurement

errors in a Wiener process. Extensions of their model have been done by Si et al. [46] and Peng and Tseng [35]. Si et al. [46] used a non-linear degradation model to estimate the remaining useful life of a system and derived an analytical approximation for the distribution of the first passage time of the transformed process. Peng and Tseng [35] suggested a similar degradation model by relaxing the normality assumption of the random effect (drift rate) with a skew-normal distribution, and derived analytical expressions for the product's lifetime distribution. Ye et al. [83] discussed Wiener processes with measurement errors. They explored the traditional Wiener process with positive drifts compounded with independent Gaussian noises, and improved its estimation efficiency compared to the existing inferential procedure. Furthermore, they developed a Wiener process with a mixed effects model subject to measurement errors, and investigated some basic properties of this degradation model. Wang et al. [68] presented a remaining life prediction method based on Wiener processes with ADT conjugate prior information. They established a data extrapolation method to extrapolate the accelerated data from the accelerated stress levels to the normal use stress level using acceleration factors, and then applied EM algorithm to estimate hyper-parameters from the prior information. Wang et al. [72] studied the problem of model mis-specification based on a nonlinear Wiener process with random effects in which the product-to-product variability and the temporal uncertainty of the degradation were both incorporated. In their work, the unknown parameters in the degradation models were obtained by using the two-step MLE method, while the effects of model mis-specification on the MTTF of the product were measured by relative bias and variability. Paroissin [32] proposed a randomly delayed Wiener process in which sample paths are assumed to be observed at the same regular instants (random time), and then discussed statistical inference and some asymptotic results of this degradation model.

6.2.3.2 Random-Effect Gamma Process

Lawless and Crowder [18] modeled the data on crack growth by constructing a tractable gamma-process model incorporating a random effect. They carried out the corresponding goodness-of-fit tests and developed prediction calculations for failure times defined in terms of degradation level passages. Based on the same model, Wang [71] proposed a pseudo-likelihood method to estimate the unknown model parameters. Tsai et al. [56] discussed the problem of optimal design for degradation tests based on a gamma degradation process with random effects. Under the constraint that the total experimental cost does not exceed a pre-fixed budget, they determined the optimal decision variables (the sample size, inspection frequency, and measurement numbers) by minimizing the asymptotic variance of the estimate of the pth percentile of the lifetime distribution of the product. Ye et al. [85] addressed the semiparametric inference on gamma process with random effects. They applied the EM algorithm to obtain the MLE of the unknown model parameters and used bootstrap method to construct confidence intervals. In addition, they developed a score test to examine the existence of random effects under the

semiparametric scenario. Guida and Penta [11] used a gamma process with non-stationary independent increments for each specimen to model the time to reach any given crack size in fatigue testing.

6.2.3.3 Random-Effect Inverse Gaussian Process

Wang and Xu [74] studied the maximum likelihood estimation of an IG process model with random effects and covariates. They used the EM algorithm to estimate the unknown model parameters and developed a simple graphical method to assess the goodness of fit of different models. Qin et al. [40] and Zhang et al. [94] used an IG process that was formulated in a hierarchical Bayesian framework to characterize the growth of the depth of metal-loss corrosion defects on oil and gas steel pipelines. Liu et al. [25] studied the dependence modeling for systems using IG process and copulas. They first used the IG process with random drift and time scale transformation to model the monotonic degradation process, and then employed the copula method to fit the joint distribution of multiple degradation processes. In their work, the EM algorithm with two-stage procedure was used to estimate the model parameters. Peng et al. [36, 38] conducted a Bayesian analysis of IG process models for degradation modeling and inference. Specifically, Peng et al. [36] investigated a simple IG process model and three IG process models with random effects. They made a comprehensive sensitivity analysis of prior distributions and sample sizes through Monte Carlo simulations. Peng et al. [38] discussed the optimal design of degradation tests based on the average pre-posterior variance of reliability and addressed the issue of uncertainty in the planning values. They also investigated a trade-off between sample size and number of degradation observations in the degradation test planning, and quantified the effects of priors on the optimal designs and on the value of prior information. Ye and Chen [80] showed that the IG processes is a limiting compound Poisson process and developed some random effects models by linking the IG process to the Wiener process. Under the assumption of the natural conjugate distribution, Peng [33] proposed a degradation model based on an inverse normal-gamma mixture of an IG process with random effects and time-independent explanatory variables. He then presented some properties of the lifetime distribution and used the EM-type algorithm to obtain the MLE of model parameters. In addition, he also provided a simple model-checking procedure to assess the validity of different stochastic processes.

6.2.4 Other Degradation Models

Shiau and Lin [45] proposed a nonparametric regression accelerated degradation model for ADT data wherein stress levels affect only the degradation rate, but

not the shape of the degradation curve. They proposed an algorithm to estimate the components of the model and conducted a simulation study to explore the effectiveness of the proposed model. Park and Padgett [30] developed a general model formulation for accelerated-test-type data on tensile strength of materials. Their approach assumed that initial damage exists in a material specimen which reduces its theoretical strength and can be modeled by a stochastic process such as Brownian motion, geometric Brownian motion, and the gamma process, and additive and multiplicative cumulative damage functions that result in the distribution of the specimen's initial strength. Hsieh and Jeng [14] used a non-homogeneous Weibull compound Poisson model with accelerated stress variables to assess the device reliability. They provided a general maximum likelihood approach for the estimates of model parameters and derived the breakdown time distributions. Ye et al. [82] used the cumulative hazard and the hazard potential of the natural lifetime to represent the degradation and shock mechanisms of the system. They assumed the lifetime of a degradation-oriented failure to belong to some distributional family, the shocks arrive according to a non-homogeneous Poisson process, and the destructive probability depends on the transformed remaining resource of the system. Under these assumptions, they developed the statistical inference for the single failure time model where the system is not repairable upon failure, and the recurrent event model where the system can be minimally repaired for shock failures and is discarded when natural failure occurs. Su and Shen [49] presented a multi-hidden semi-Markov model with multivariate feature variables for degradation identification and remaining useful life forecasting of a system. They used a new variable combined with forward and backward variables to depict similarities of the feature and applied modified Viterbi algorithm to identify the degradation state, in which the linear function was used to describe the contribution of each feature to the state recognition, and the weight coefficients were rectified by using the dynamic adjustment method when missing feature problem happens. Compare et al. [7] employed a four-state continuous-time semi-Markov process with Weibull distributed transition times to model oxidation degradation mechanism in gas turbine nozzles. They developed a numerical approach to obtain the MLE of the model parameters, which can be further utilized to compute the probabilities of occupying the four degradation states over time and the corresponding uncertainties, and characterize the related uncertainty and the probabilities of occupying the different states over time. By minimizing the asymptotic variance of the estimated pth quantile of the product's lifetime, Tseng and Lee [60] analytically derived the optimal allocation rules for two-level and three-level exponential dispersion ADT allocation problems whether the testing stress levels are prefixed or not. In particular, they showed that all test units for a three-level allocation problem should be allocated into two out of three stresses, depending on certain specific conditions. They also demonstrated that a three-level compromise plan with small proportion allocation in the middle stress, in general, is a good strategy for ADT allocation. Lin et al. [22] considered importance indices of components within multi-component systems by taking into account the influence of multiple competing degradation processes, degradation dependency and maintenance tasks. They used a piecewise-

deterministic Markov process to describe the stochastic process of degradation of the component under these factors, and developed a method for the quantification of the component importance measure based on the finite-volume approach.

6.3 Destructive Degradation Model

In some applications, the degradation measurement process would destroy the physical characteristic of units when tested at higher than usual stress levels of an accelerating variable such as temperature, so that only one measurement can be made on each tested unit during the degradation testing. An accelerated degradation test giving rise to such a degradation data is called an accelerated destructive degradation test (ADDT).

Let $L(t)$, denoting the measured QC of the product at time t during the degradation experiment, be

$$L(t) = f(t; \boldsymbol{\phi}) + \varepsilon, \qquad t > 0, \tag{6.8}$$

where $\boldsymbol{\phi}$ is fixed but unknown parameter vector, and the random error term ε follows a distribution. The product's lifetime T is suitably defined as the time when $L(t)$ crosses the critical level ω. Then, the cdf and the $100p$th percentile of the lifetime of the product can be obtained by using the result that the probability of failure at a given time is simply equal to the probability of the product's QC crossing the critical level ω at that time. For decreasing degradation, the lifetime cdf is $F(t) = P(T \leq t) = P(L(t) \leq \omega)$.

There have been a number of interesting studies on ADDT. Nelson [27] was the first to introduce methods for using accelerated destructive degradation test (ADDT) data to analyze the degradation and related failure time distributions for dielectric breakdown strength of insulation specimens. Escobar et al. [10] provided an important class of models for accelerated destructive degradation data, and used likelihood-based methods for inference on both the degradation and the induced failure-time distributions of the adhesive bond. Shi et al. [44] described methods to find the optimal ADDT test condition and proportional allocations by minimizing the asymptotic variance of the estimated $100p$th percentile of the lifetime distribution at use condition. Because the optimal plans can be sensitive to mis-specification of model assumptions, they also proposed a more robust and useful compromise plan and used the general equivalence theorem (GET) to verify the optimality of test plans. Jeng et al. [16] derived the large-sample approximate mean square error of the MLE of the product's lifetime quantile to address the effect of model mis-specification when the distribution is misspecified in a degradation model. Shi and Meeker [42] presented Bayesian test planning methods for ADDT problems under an important class of nonlinear regression models when prior information is available on the model parameters. They used a Bayesian criterion based on the estimation precision of a failure-time distribution quantile at use

conditions to find optimal test plans, and examined the effects of changing the amount of prior information and sample size on conducting the Bayesian test plan. Shi and Meeker [43] suggested methods for planning ADDT for applications with more than one type of failure. They gave unconstrained and constrained optimal ADDT plans under a given test optimization criterion, and applied a modification of the GET to verify the optimality of this test plan. Doganaksoy et al. [9] discussed the use of ADDT on the issues of model building, test plan, parameter estimation, and assessment of model adequacy for seal strength.

All the models mentioned above assume a linearized relationship for an ADDT model, i.e., the relationship between the mean transformed degradation path and the transformed time is linear at each level of the accelerating variable. In a different direction, Tsai et al. [57] determined the optimal choices of total sample size, measurement frequency, and the number of measurements at each stress level for a nonlinear ADDT model by minimizing the asymptotic variance of the estimated $100p$th percentile of the lifetime distribution under use condition subject to the total experimental cost not exceeding a pre-fixed budget. Tsai and Lin [53] further proposed a nonlinear ADDT model wherein the measurement error follows a skew-normal distribution, and addressed the effects of model mis-specification when the skewness parameter of measurement error is wrongly specified.

Noting that none of the above methods have considered degradation initiation times, Zhang and Liao [90] pointed out that ignoring degradation initiation times in DDT makes reliability estimates less accurate, especially when degradation initiation times are random and unobservable. For this reason, Zhang and Liao [90] developed two delayed-degradation models based on DDT data to evaluate the reliability of a product with an exponentially distributed degradation initiation time. For homogeneous and heterogeneous populations, they considered fixed-effects and random-effects gamma processes, respectively, in modeling the actual degradation of units after degradation initiation. They used EM algorithm to estimate unobserved degradation initiation times and MLE method to estimate the model parameters.

6.4 Applications on Degradation Model

Burn-in test is a useful tool to eliminate latent failures or weak components during the manufacturing process before the products are sent to customers. Tseng and Tang [62] first used degradation data instead of failure data to discuss an optimal burn-in policy. They proposed a decision rule for classifying an unit as a typical or a weak unit, and used an economic model to determine the optimal termination time of a burn-in test through a Wiener degradation process. Tseng et al. [63] discussed the same decision rule and presented a two-stage procedure to determine the optimal burn-in time and the median of the residual lifetime of the passed units. Tseng and Peng [61] studied an efficient burn-in policy based on an integrated Wiener process for the cumulative degradation of the QC of the product. Wu and Xie [77] suggested

the use of receiver operating characteristic analysis to select optimal approaches to distinguishing weak from strong components in a population, and determined the optimal time for burn-in to remove the weak population when precise information about the life distribution of sub-populations, and mis-classification costs in the entire population might not be available. Tsai et al. [55] presented a decision rule for classifying a tested unit as typical or weak based on a mixed gamma process. They derived analytically the optimal cutoff point at each time by minimizing the total cost function, and examined the effect of wrongly treating a mixed gamma process as a mixed Wiener process. Ye et al. [84] developed a degradation-based burn-in planning framework for products with independent competing risks through gamma process with random effect and a catastrophic mode. Zhang et al. [93] investigated the optimal burn-in strategies under two different cost models through the Wiener process with Gaussian measurement errors. The first model involved the misclassification cost during burn-in, while the second considered the burn-in implementation cost and field failure cost. Through analytical analysis, they showed the intimate relationship between the two models and proposed an accurate approximation method to approximate the second model, which greatly simplifies subsequent analysis.

Maintenance is a necessary operation applied to a system in companies so that the system can work properly. Noortwijk [67] gave an overview of the most important statistical and mathematical properties of gamma processes, and surveyed the application of gamma processes for optimizing time-based and condition-based maintenance. Ye et al. [81] used a Wiener process with linear drift to develop two joint degradation-based burn-in and maintenance models under the age- and block-based maintenances, respectively. They minimized the long run average cost per unit time to determine the optimal burn-in settings and the preventive replacement intervals. Zhang et al. [91] considered a non-stationary Wiener process to develop an imperfect maintenance model for systems. Shafiee et al. [41] modeled the length of a crack using a gamma process, and proposed an optimal opportunistic condition-based maintenance policy for a deteriorating multi-bladed wind turbine system. Caballe et al. [4] addressed a condition-based maintenance strategy for a system by assuming that degradation processes start at random times following a non-homogeneous Poisson process and their growths are modeled by using a gamma process. Wu et al. [76] suggested a random-effect exponential degradation model with error term that follows a Markov process with independently distributed increments to discuss degradation-based maintenance strategy under imperfect repair. They determined an optimal interval of condition monitoring and the degradation level after imperfect preventive repairs by minimizing the total cost of imperfect degradation-based maintenance. Chen et al. [5] discussed an optimal condition-based replacement policy with optimal inspection interval when the degradation paths follows an IG process with random-effects drift.

Acceptance test is extremely important stage for ensuring products to meet customer's requirements. Motivated by a resistor data, Tsai et al. [54] used a Wiener degradation model to discuss an accelerated-stress acceptance test to shorten the acceptance testing time. By minimizing the asymptotic variance of the

estimated optimal accelerated-stress acceptance testing time subject to the total experimental cost not exceeding a pre-fixed budget, they determined the optimal test plan, including the total sample size, measurement frequency, and the number of measurements, for an accelerated-stress acceptance test.

6.5 Concluding Remarks

Non-destructive and destructive degradation models have become important analytic tools for complex systems. Following the work of Nelson [27], a number of non-destructive and destructive degradation models have been developed to capture the degradation dynamics of a system and to assist in the subsequent decision-making. This chapter provides a brief background about the models used in the literature for estimating reliability using either degradation data from non-destructive tests (NDT) or destructive degradation tests (DDT). The discussed non-destructive degradation models are classified into four classes, namely, fixed or random effect models, stochastic process models including Wiener process, gamma process and inverse Gaussian (IG) process, mixed random effect and stochastic process models including random-effect Wiener process, random-effect gamma process and random-effect IG process, and some other models in addition to these classes. Several destructive degradation and accelerated destructive degradation models with linearized and nonlinear degradation paths have been highlighted. In addition to introducing these models, applications of these models in burn-in, maintenance and acceptance tests have also been described. This chapter is intended to help readers gain a broad understanding of various developments on the NDT and DDT models and their applications.

References

1. Amini M, Shemehsavar S, Pan Z (2016) Optimal design for step-stress accelerated test with random discrete stress elevating times based on gamma degradation process. Qual Reliab Eng Int 32:2391–2402
2. Bae SJ, Kim SJ, Kim MS, Lee BJ, Kang CW (2008) Degradation analysis of nano-contamination in plasma display panels. IEEE Trans Reliab 57:222–229
3. Bordes L, Paroissin C, Salami A (2016) Parametric inference in a perturbed gamma degradation process. Comput Stat – Theory Methods 45:2730–2747
4. Caballe NC, Castro IT, Perez CJ, Lanza-Gutierrez JM (2015) A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. Reliab Eng Syst Saf 134:98–109
5. Chen N., Ye ZS, Xiang Y, Zhang L (2015) Condition-based maintenance using the inverse Gaussian degradation model. Eur J Oper Res 243:190–199
6. Chiang JY, Sung WY, Tsai TR, Lio Y (2015) Sensitivity analysis of sample allocation and measurement frequency under a degradation test with gamma process. ICIC Express Lett Part B: Appl 6:737–742

7. Compare M, Martini F, Mattafirri S, Carlevaro F, Zio E (2016) Semi-Markov model for the oxidation degradation mechanism in gas turbine nozzles. IEEE Trans Reliab 65: 574–581
8. Cui L, Huang J, Li Y (2016) Degradation models with Wiener diffusion processes under calibrations. IEEE Trans Reliab 65:613–623
9. Doganaksoy N, Hahn GJ, Meeker WQ (2014) Timely reliability assessment: using destructive degradation tests. Qual Prog 47:52–55
10. Escobar LA, Meeker WQ, Kugler DL, Kramer LL (2003) Accelerated destructive degradation tests: data, models, and analysis. In: Lindqvist BH, Doksum KA (eds) Chapter 21 in mathematical and statistical methods in reliability. World Scientific Publishing Company, River Edge
11. Guida M, Penta F (2015) A gamma process model for the analysis of fatigue crack growth data. Eng Fract Mech 142:21–49
12. Guida M, Postiglione F, Pulcini G (2012) A time-discrete extended gamma process for time-dependent degradation phenomena. Reliab Eng Syst Saf 105:73–79
13. Guida M, Postiglione F, Pulcini G (2015) A random-effects model for long-term degradation analysis of solid oxide fuel cells. Reliab Eng Syst Saf 140:88–98
14. Hsieh MH, Jeng SL (2007) Accelerated discrete degradation models for leakage current of ultra-thin gate oxides. IEEE Trans Reliab 56:369–380
15. Hu CH, Lee MY, Tang J (2015) Optimum step-stress accelerated degradation test for Wiener degradation process under constraints. Eur J Oper Res 241:412–421
16. Jeng SL, Huang BY, Meeker WQ (2011) Accelerated destructive degradation tests robust to distribution mis-specification. IEEE Trans Reliab 60:701–711
17. Kim SJ, Bae SJ (2013) Cost-effective degradation test plan for a nonlinear random-coefficients model. Reliab Eng Syst Saf 110:68–79
18. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
19. Liao H, Elsayed EA (2006) Reliability inference for field conditions from accelerated degradation testing. Nav Res Logist 53:576–587
20. Liao CM, Tseng ST (2006) Optimal design for step-stress accelerated degradation tests. IEEE Trans Reliab 55:59–66
21. Lim H, Yum BJ (2011) Optimal design of accelerated degradation tests based on Wiener process models. J Appl Stat 38:309–325
22. Lin YH, Li YF, Zio E (2016) Component importance measures for components with multiple dependent competing degradation processes and subject to maintenance. IEEE Trans Reliab 65:547–557
23. Ling MH, Tsui KL, Balakrishnan N (2015) Accelerated degradation analysis for the quality of a system based on the gamma process. IEEE Trans Reliab 64:463–472
24. Liu X, Tang LC (2010) A Bayesian optimal design for accelerated degradation tests. Qual Reliab Eng Int 26:863–875
25. Liu Z, Ma X, Yang J, Zhao Y (2014) Reliability modeling for systems with multiple degradation processes using inverse Gaussian process and copulas. Math Probl Eng 2014:Article ID 829597, 10 pp.
26. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
27. Nelson W (1981) Analysis of performance degradation data from accelerated tests. IEEE Trans Reliab 30:149–155
28. Pan R, Crispin T (2011) A hierarchical modeling approach to accelerated degradation testing data analysis: a case study. Qual Reliab Eng Int 27:229–237
29. Pan Z, Balakrishnan N (2010) Multiple-steps step-stress accelerated degradation modeling based on Wiener and gamma processes. Commun Stat Simul Comput 39:1384–1402

30. Park C, Padgett WJ (2005) New cumulative damage models for failure using stochastic processes as initial damage. IEEE Trans Reliab 54:530–540
31. Park J, Bae SJ (2010) Direct prediction methods on lifetime distribution of organic light-emitting diodes from accelerated degradation tests. IEEE Trans Reliab 59:74–90
32. Paroissin C (2016) Inference for the Wiener process with random initiation time. IEEE Trans Reliab 65:147–157
33. Peng CY (2015) Inverse gaussian processes with random effects and explanatory variables for degradation data. Technometrics 57:100–111
34. Peng CY, Tseng ST (2009) Mis-specification analysis of linear degradation models. IEEE Trans Reliab 58:444–455
35. Peng CY, Tseng ST (2013) Statistical lifetime inference with skew-Wiener linear degradation models. IEEE Trans Reliab 62:338–350
36. Peng W, Li YF, Yang YJ, Huang HZ, Zuo MJ (2014) Inverse Gaussian process models for degradation analysis: a Bayesian perspective. Reliab Eng Syst Saf 130:175–189
37. Peng W, Li YF, Yang YJ, Zhu SP, Huang HZ (2016) Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas. IEEE Trans Reliab 65:624–639
38. Peng W, Liu Y, Li YF, Zhu SP, Huang HZ (2014) A Bayesian optimal design for degradation tests based on the inverse Gaussian process. J Mech Sci Technol 28:3937–3946
39. Pulcini G (2016) A perturbed gamma process with statistically dependent measurement errors. Reliab Eng Syst Saf 152:296–306
40. Qin H, Zhang S, Zhou W (2013) Inverse Gaussian process-based corrosion growth modeling and its application in the reliability analysis for energy pipelines. Front Struct Civ Eng 7: 276–287
41. Shafiee M, Finkelstein M, Berenguer C (2015) An opportunistic condition-based maintenance policy for offshore wind turbine blades subjected to degradation and environmental shocks. Reliab Eng Syst Saf 142:463–471
42. Shi Y, Meeker WQ (2012) Bayesian methods for accelerated destructive degradation test planning. IEEE Trans Reliab 61:245–253
43. Shi Y, Meeker WQ (2013) Planning accelerated destructive degradation tests with competing risks. In: Couallier V et al (eds) Statistical models and methods for reliability and survival analysis, chap. 22. Wiley, Hoboken, pp 335–355
44. Shi Y, Meeker WQ, Escobar LA (2009) Accelerated destructive degradation test planning. Technometrics 51:1–13
45. Shiau JJH, Lin HH (1999) Analyzing accelerated degradation data by nonparametric regression. IEEE Trans Reliab 48:149–158
46. Si XS, Wang WB, Hu CH, Zhou DH, Pecht MG (2012) Remaining useful life estimation based on a nonlinear diffusion degradation process. IEEE Trans Reliab 61:50–67
47. Singpurwalla ND (1995) Survival in dynamic environments. Stat Sci 10:86–103
48. Singpurwalla ND (1997) Gamma processes and their generalizations: an overview. In: Cooke NR, Mendel M, Vrijling H (eds) Engineering probabilistic design and maintenance for flood protection. Kluwer Academic Publishers, Dordrecht, pp 67–75
49. Su C, Shen J (2013) A novel multi-hidden semi-Markov model for degradation state identification and remaining useful life estimation. Qual Reliab Eng Int 29:1181–1192
50. Tamborrino M (2016) Approximation of the first passage time density of a Wiener process to an exponentially decaying boundary by two-piecewise linear threshold. Application to neuronal spiking activity. Math Biosci Eng 13:613–629
51. Tang J, Su TS (2008) Estimating failure time distribution and its parameters based on intermediate data from a Wiener degradation model. Nav Res Logist 55:265–276
52. Tsai CC, Lin CT (2014) Optimal selection of the most reliable design based on gamma degradation processes. Comput. Stat. Theory Methods 43:2419–2428
53. Tsai CC, Lin CT (2015) Lifetime inference for highly reliable products based on skew-normal accelerated destructive degradation test model. IEEE Trans Reliab 64:1340–1355

54. Tsai CC, Lin CT, Balakrishnan N (2015) Optimal design for accelerated-stress acceptance test based on Wiener process. IEEE Trans Reliab 64:603–612
55. Tsai CC, Tseng ST, Balakrishnan N (2011) Optimal burn-in policy for highly reliable products using gamma degradation process. IEEE Trans Reliab 60:234–245
56. Tsai CC, Tseng ST, Balakrishnan N (2012) Optimal design for degradation tests based on gamma processes with random effects. IEEE Trans Reliab 61:604–613
57. Tsai CC, Tseng ST, Balakrishnan N, Lin CT (2013) Optimal design for accelerated destructive degradation test. Qual Technol Quant Manag 10:263–276
58. Tsai TR, Lin CW, Sung YL, Chou PT, Chen CL, Lio YL (2012) Inference from lumen degradation data under Wiener diffusion process. IEEE Trans Reliab 61:710–718
59. Tsai TR, Sung WY, Lio YL, Chang SI, Lu JC (2016) Optimal two-variable accelerated degradation test plan for gamma degradation processes. IEEE Trans Reliab 65:459–468
60. Tseng ST, Lee IC (2016) Optimum allocation rule for accelerated degradation tests with a class of exponential-dispersion degradation models. Technometrics 58:244–254
61. Tseng ST, Peng CY (2004) Optimal burn-in policy by using an integrated Wiener process. IIE Trans 36:1161–1170
62. Tseng ST, Tang J (2001) Optimal burn-in time for highly reliable products. Int J Ind Eng 8:329–338
63. Tseng ST, Tang J, Ku IH (2003) Determination of burn-in parameters and residual life for highly reliable products. Nav Res Logist 50:1–14
64. Tseng ST, Balakrishnan N, Tsai CC (2009) Optimal step-stress accelerated degradation test plan for gamma degradation processes. IEEE Trans Reliab 58:611–618
65. Tseng ST, Tsai CC, Balakrishnan N (2011) Optimal sample size allocation for accelerated degradation test based on Wiener process. In: Balakrishnan N (ed) Methods and applications of statistics in engineering, quality control, and the physical sciences, chapter 27. Wiley, Hoboken, pp 330–343
66. Tseng ST, Yu HF (1997) A termination rule for degradation experiments. IEEE Trans Reliab 46:130–133
67. von Noortwijk JM (2009) A survey of the application of gamma processes in maintenance. Reliab Eng Syst Saf 94:2–21
68. Wang HW, Xu TX, Wang WY (2016) Remaining life prediction based on Wiener processes with ADT prior information. Qual Reliab Eng Int 32:753–765
69. Wang L, Pan R, Li X, Jiang T (2013) A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliab Eng Syst Saf 112:38–47
70. Wang W, Carr M, Xu W, Kobbacy AKH (2011) A model for residual life prediction based on Brownian motion with an adaptive drift. Microelectron Reliab 51:285–293
71. Wang X (2008) A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects. Stat Sin 18:1153–1163
72. Wang X, Balakrishnan N, Guo B (2016) Mis-specification analyses of nonlinear Wiener process-based degradation models. Comput Stat Simul Comput 45:814–832
73. Wang X, Jiang P, Guo B, Cheng Z (2014) Real-time reliability evaluation for an individual product based on change-point gamma and Wiener process. Qual Reliab Eng Int 30: 513–525
74. Wang X, Xu D (2010) An inverse Gaussian process model for degradation data. Technometrics 52:188–197
75. Wei M, Chen M, Zhou D (2013) Multi-sensor information based remaining useful life prediction with anticipated performance. IEEE Trans Reliab 62:183–198
76. Wu F, Niknam SA, Kobza JE (2015) A cost effective degradation-based maintenance strategy under imperfect repair. Reliab Eng Syst Saf 144:234–243
77. Wu S, Xie M (2007) Classifying weak, and strong components using ROC analysis with application to burn-in. IEEE Trans Reliab 56:552–561
78. Wu SJ, Chang CT (2002) Optimal design of degradation tests in presence of cost constraint. Reliab Eng Syst Saf 76:109–115

79. Ye ZS, Chen LP, Tang LC, Xie M (2014) Accelerated degradation test planning using the inverse Gaussian process. IEEE Trans Reliab 63:750–763
80. Ye ZS, Chen N (2014) The inverse Gaussian process as a degradation model. Technometrics 56:302–311
81. Ye ZS, Shen Y, Xie M (2012) Degradation-based burn-in with preventive maintenance. Eur J Oper Res 221:360–367
82. Ye ZS, Tang LC, Xu HY (2011) A distribution-based systems reliability model under extreme shocks and natural degradation. IEEE Trans Reliab 60:246–256
83. Ye ZS, Wang Y, Tsui KL, Pecht M (2013) Degradation data analysis using Wiener processes with measurement errors. IEEE Trans Reliab 62:772–780
84. Ye ZS, Xie M, Shen Y, Tang LC (2012) Degradation-based burn-in planning under competing risks. Technometrics 54:159–168
85. Ye ZS, Xie M, Tang LC, Chen N (2014) Semiparametric estimation of gamma processes for deteriorating products. Technometrics 56:504–513
86. Yu HF (2003) Designing an accelerated degradation experiment by optimizing the estimation of the percentile. Qual Reliab Eng Int 19:197–214
87. Yu HF (2006) Designing an accelerated degradation experiment with a reciprocal Weibull degradation rate. J Stat Plan Inference 136:282–297
88. Yu HF, Chiao CH (2002) An optimal designed degradation experiment for reliability improvement. IEEE Trans Reliab 51:427–433
89. Yu HF, Tseng ST (1999) Designing a degradation experiment. Nav Res Logist 46:689–706
90. Zhang Y, Liao H (2015) Analysis of destructive degradation tests for a product with random degradation initiation time. IEEE Trans Reliab 64:516–527
91. Zhang M, Gaudoin O, Xie M (2015) Degradation-based maintenance decision using stochastic filtering for systems under imperfect maintenance. Eur J Oper Res 245:531–541
92. Zhang M, Ye ZS, Xie M (2015) Optimal burn-in policy for highly reliable products using inverse Gaussian degradation process. Lect Notes Mech Eng 19:1003–1011
93. Zhang Q, Ye ZS, Yang J, Zhao Y (2016) Measurement errors in degradation-based burn-in. Reliab Eng Syst Saf 150:126–135
94. Zhang S, Zhou W, Qin H (2013) Inverse Gaussian process-based corrosion growth model for energy pipelines considering the sizing error in inspection data. Corros Sci 73:309–320

Part II
Modeling and Experimental Designs

Chapter 7
Degradation Test Plan for a Nonlinear Random-Coefficients Model

Seong-Joon Kim and Suk Joo Bae

Abstract Sample size and inspection schedule are essential components in degradation test plan. In practice, an experimenter is required to determine a certain level of trade-off between total resources and precision of the degradation test. This paper develops a design of cost-efficient degradation test plan in the context of a nonlinear random-coefficients model, while satisfying precision constraints for the failure-time distribution derived from the degradation testing data. The test plan introduces a precision metric to identify the information losses due to reduction of test resources, based on the cost function to balance the test plan. In order to determine a cost-efficient inspection schedule, a hybrid genetic algorithm is used to solve a cost optimization problem under test precision constraints. The proposed method is applied to degradation data of plasma display panels (PDPs). Finally, sensitivity analysis is provided to show the robustness of the proposed test plan.

Keywords Degradation test • Fisher information matrix • Nonlinear random-coefficients model • Reliability

7.1 Introduction

Intense global competition in the high-technology industry forces manufacturers to evaluate product reliability within shorter testing times, along with limited resources. Recently, (accelerated) degradation tests have replaced traditional (accelerated) life tests when the failure data are supplemented by degradation data as measurements of product wear available at one or more time points during the reliability test. Degradation data not only lead to improved reliability analysis

S.-J. Kim
Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju, Seoul, South Korea
e-mail: seongjoon.kim@chosun.ac.kr

S.J. Bae (✉)
Department of Industrial Engineering, Hanyang University, 222 Wangsimni-ro Seongdong-gu, Seoul, South Korea
e-mail: sjbae@hanyang.ac.kr

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_7

[11], but they also provide additional information related to failure mechanisms for testing units.

To conduct a degradation test, three experimental factors must be specified a priori: a threshold level of failure (or *soft failure*), an inspection schedule for degradation measurements including termination time, and the total number of testing units. In non-acceleration degradation test, Yu and Tseng proposed a degradation test plan to determine decision variables such as the sample size, inspection frequency, and termination time for a power degradation model [25], and a reciprocal Weibull degradation rate [26]. Marseguerra et al. [14] proposed a multi-objective genetic algorithm for designing degradation experiments. In accelerated degradation test (ADT), Boulanger and Escobar [4] proposed an optimal ADT plan for a nonlinear degradation process as a function of temperature. Yu and Tseng [24] suggested an online procedure to determine an appropriate termination time for ADT. Tseng and Wen [19] proposed a step-stress accelerated degradation test to reduce experimental costs in the assessment of the lifetime distribution of light emitting diodes (LEDs).

In practice, two additional issues must also be considered in a degradation experiment so that the degradation test can be efficient in terms of testing cost and experimental precision: limited experimental budget including both tangible and intangible costs (e.g., sample cost, operation cost, labor cost, etc.) and the estimation precision of interest in the experiment (e.g., pth quantile of failure-time distribution). The size of budget affects the decisions on the total number of testing units, the number of inspections, and termination time, hence it affects estimation precision of failure-time distribution. Accordingly, proper allocation of such limited resources helps keep a balance between testing cost and estimation precision of the experiment. Wu and Chang [21] suggested optimal combinations of the three decision variables minimizing the variance of quantiles of failure-time distribution from a degradation experiment, mainly based on the nonlinear mixed-effect model proposed by Lu and Meeker [10], the parameters of which were estimated using the least squares (LS) method.

However, most degradation test plans are derived based on simple degradation models like a linear degradation model or a power degradation model. Recently, a complicated degradation model (i.e., *bi-exponential model*) was proposed to describe nonlinear degradation paths of plasma display panels (PDPs) [2], and those of organic light emitting diodes (OLEDs) [17]. The bi-exponential model has precedence in analyzing longitudinal data from biological and pharmacokinetics systems [12, 13].

This research proposes an optimal degradation test plan which minimizes total experimental cost in the context of a nonlinear random-coefficients model. Under the degradation test plan, we seek to meet some precision constraints for failure-time distribution derived from the modeling of nonlinear degradation paths. Furthermore, we reflect industries' endeavor for meeting time-to-market by assigning higher values of cost at latter measurements in the testing period. We try to quantify information losses incurred by reducing the test resources through the probability of false decisions (i.e., Type I error and Type II error), and explicitly incorporate it into

both the cost functions and the precision constraints. Unlike existing degradation test plans, the test plan permits an experimenter to select proper quality level of the degradation test by adjusting Type I and Type II errors with acceptable estimation precision.

7.2 The Degradation Model

7.2.1 Nonlinear Random-Coefficients Model

Random-coefficients models provide a flexible and powerful tool for analyzing repeated-measurement data that arise in various fields of application. Random-coefficients models are intuitively appealing because they allow for various variance-covariance structures of the response vector. Each individual product can experience a different source of variation during fabrication. Therefore, a degradation model with a random degradation rate is more appropriate to capture item-to-item variability of the degradation process for individual items. By introducing a random-coefficients model, we can easily incorporate individual variability into the degradation model.

A nonlinear random-coefficients (NRC) model for the jth response on the ith individual item is defined as

$$y_{ij} = \eta\left(t_{ij}, \boldsymbol{\theta}_i\right) + \epsilon_{ij}, \qquad j = 1, \ldots, n_i,\ i = 1, \ldots, M \tag{7.1}$$

where y_{ij} is the jth response for the ith individual, t_{ij} is jth measurement time for the ith individual, $\eta(\cdot)$ is a nonlinear function of t_{ij} and parameter vector $\boldsymbol{\theta}_i$, and ϵ_{ij} is a s-normally distributed random error term. We assume that $\eta(\cdot)$ is a twice differentiable continuous function. By letting $\mathbf{y}_i$, and $\boldsymbol{\epsilon}_i$ be the $(n_i \times 1)$ vectors of responses and random within-individual errors for individual i, respectively, the model (7.1) can be written as

$$\mathbf{y}_i = \boldsymbol{\eta}\left(\boldsymbol{\xi}_i, \boldsymbol{\theta}_i\right) + \boldsymbol{\epsilon}_i, \tag{7.2}$$

where $\boldsymbol{\eta}\left(\boldsymbol{\xi}_i, \boldsymbol{\theta}_i\right)$ is the $(n_i \times 1)$ vector of mean responses for the ith individual item, which depends on $\boldsymbol{\xi}_i = \left(t_{i1}, \ldots, t_{in_i}\right)^T$ and the $(p \times 1)$ individual-specific model parameter $\boldsymbol{\theta}_i$. We assume that $\boldsymbol{\epsilon}_i$ has a multivariate normal distribution with mean of zeros and variance-covariance matrix $\boldsymbol{\Sigma}_i$. In many practical applications, within-individual heteroscedasticity and correlation are dominated by between-individual variation, thus may be negligible. Therefore, we define $\boldsymbol{\Sigma}_i = \sigma^2 \boldsymbol{I}_{n_i}$, the common specification of uncorrelated within-individual errors with constant variance. The parameter vector $\boldsymbol{\theta}_i$ can be considered as the sum of the fixed and random effects vectors because it can vary from item to item. This is incorporated into the model as an additive model by writing $\boldsymbol{\theta}_i$ as

$$\boldsymbol{\theta}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i, \qquad \mathbf{b}_i \sim \mathscr{MN}_s\left(\mathbf{0}, \boldsymbol{D}\right) \tag{7.3}$$

where $\boldsymbol{\beta}$ and $\mathbf{b}_i$ are the $(r \times 1)$ and $(s \times 1)$ vectors of the fixed and random effects, respectively, $\mathbf{A}_i$ and $\mathbf{B}_i$ are known design matrices of size $(p \times r)$ and $(p \times s)$ corresponding to $\boldsymbol{\beta}$ and $\mathbf{b}_i$, respectively, and $\boldsymbol{D}$ is a $(s \times s)$ variance-covariance matrix of the random effects. Let $\boldsymbol{\Psi} = \left(\boldsymbol{\beta}^T, d_1, \ldots, d_s, d_{12}, \ldots, d_{s-1,s}, \sigma^2\right)^T$ be the vector of all model parameters, where d_k represents the variance of the kth component for $k = 1, \ldots, s$, and d_{ij} represents the covariance component (i, j) in the variance-covariance matrix of the random effects.

7.2.2 The Fisher Information Matrix

The Cramer-Rao theorem provides a lower bound on the variance-covariance matrix of unbiased parameter estimates. The (elementary) Fisher information matrix is given by

$$\mathscr{I}\left(\boldsymbol{\xi}_i|\boldsymbol{\Psi}\right) = -E\left[\frac{\partial^2 l\left(\boldsymbol{\xi}_i|\boldsymbol{\Psi}\right)}{\partial \boldsymbol{\Psi}\, \partial \boldsymbol{\Psi}^T}\right], \qquad i = 1, 2, \ldots, M \tag{7.4}$$

where $l\left(\boldsymbol{\xi}_i|\boldsymbol{\Psi}\right)$ is the log-likelihood with respect to the inspection schedule vector $\boldsymbol{\xi}_i$.

When the function η is nonlinear in $\mathbf{b}_i$, the likelihood function does not have a closed-form expression, nor does the resulting Fisher information matrix. Several approximation methods have been proposed to estimate the likelihood function of the NRC model. Mentre et al. [15] developed the expression of the Fisher information matrix using a first-order Taylor expansion of the model around the random effects. Bae and Kvam [1] introduced four different methods to approximate the likelihood function of the NRC model: a first-order (FO) method [18], Lindstrom and Bates' algorithm [9], adaptive importance sampling, and adaptive Gaussian quadrature. Delyon et al. [5] and Kuhn and Laville [8] applied a stochastic approximation expectation-maximization (SAEM) algorithm to approximate the likelihood function of the NRC model.

In this work, we choose the FO method, which linearizes the NRC model around the expectation of random effects with a closed-form expression for the Hessian, and the simulation-based first-order conditional estimate (FOCE) method [3], which linearizes the NRC model around empirical estimates of random effects.

7.2.2.1 The FO Method

The linearization of the NRC model was first proposed by Sheiner et al. [18] in the parameter estimation problem. Through a first-order Taylor expansion of the NRC model around a mean value $\bar{\mathbf{b}}$ of the random effects, the degradation model (7.2) can be written as

$$\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}}) + \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}})^T}{\partial \mathbf{b}}(\mathbf{b} - \bar{\mathbf{b}}) + \boldsymbol{R}(\mathbf{b}), \tag{7.5}$$

where $\boldsymbol{R}(\mathbf{b})$ is the remainder of the Taylor series. The marginal expectation and variance of $\mathbf{y}$ around $\bar{\mathbf{b}} = \mathbf{0}$ can be approximated as, respectively,

$$\boldsymbol{\mu} \approxeq \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta}), \qquad \text{and} \qquad \mathbf{V} \approxeq \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta})^T}{\partial \mathbf{b}} \boldsymbol{D} \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta})}{\partial \mathbf{b}} + \sigma^2 \boldsymbol{I}_n,$$

and, the approximated log-likelihood function is then given as

$$l(\boldsymbol{\xi}|\boldsymbol{\Psi}) \approxeq -\frac{1}{2}\left[n \cdot \ln(2\pi) + \ln|\mathbf{V}| + (\mathbf{y} - \boldsymbol{\mu})^T \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\mu})\right]. \tag{7.6}$$

The Fisher information matrix can be derived from the approximated log-likelihood function (7.6).

Let $\boldsymbol{\Xi}$ be the population inspection scheme with $Q(\leq M)$ distinct inspection schedules, $\{\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_Q\}$. Each of these schedules $\boldsymbol{\xi}_q$ consists of n_q sampling times and they are conducted in M_q test units for $q = 1, \ldots, Q$. A general expression of the population inspection scheme $\boldsymbol{\Xi}$ can be denoted by

$$\boldsymbol{\Xi} = \left\{[\boldsymbol{\xi}_1, M_1]; \ldots; [\boldsymbol{\xi}_Q, M_Q]\right\} \quad \text{where} \quad M = \sum_{q=1}^{Q} M_q.$$

A population Fisher information matrix $\mathscr{I}(\boldsymbol{\Xi}|\boldsymbol{\Psi})$ that incorporates the (elementary) Fisher information matrix is expressed as

$$\mathscr{I}(\boldsymbol{\Xi}|\boldsymbol{\Psi}) = \sum_{q=1}^{Q} M_q \times \mathscr{I}\left(\boldsymbol{\xi}_q|\boldsymbol{\Psi}\right). \tag{7.7}$$

7.2.2.2 Simulation-Based FOCE Method

Because we do not have actual degradation data for individual items in the phase of degradation test design, empirical estimates of random effects cannot be evaluated. Thus, the Fisher information matrix is computed via Monte Carlo simulation to mimic the FOCE method. Simulated random effects $\bar{\mathbf{b}}_s$ are generated from $\mathscr{MN}_s(\mathbf{0}, \boldsymbol{D})$, and the Fisher information matrix $\mathscr{I}\left(\boldsymbol{\xi}, \bar{\mathbf{b}}_s|\boldsymbol{\Psi}\right)$ is computed based on the following expectation and variance, respectively,

$$\boldsymbol{\mu}_s \approxeq \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}}_s) - \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}}_s)^T}{\partial \mathbf{b}} \bar{\mathbf{b}}_s,$$

$$\mathbf{V}_s \approxeq \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}}_s)^T}{\partial \mathbf{b}} \boldsymbol{D} \frac{\partial \boldsymbol{\eta}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\beta} + \mathbf{B}\bar{\mathbf{b}}_s)}{\partial \mathbf{b}} + \sigma^2 \boldsymbol{I}_n.$$

The approximated Fisher information matrix is computed by

$$\mathscr{I}\left(\boldsymbol{\xi}_q, \bar{\mathbf{b}}_s | \boldsymbol{\Psi}\right) \approxeq \frac{1}{S} \sum_{s=1}^{S} \mathscr{I}\left(\boldsymbol{\xi}_q, \bar{\mathbf{b}}_s | \boldsymbol{\Psi}\right),$$

then the population Fisher information matrix is obtained by

$$\mathscr{I}\left(\boldsymbol{\Xi}, \bar{\mathbf{b}}_s | \boldsymbol{\Psi}\right) = \sum_{q=1}^{Q} M_q \times \mathscr{I}\left(\boldsymbol{\xi}_q, \bar{\mathbf{b}}_s | \boldsymbol{\Psi}\right). \tag{7.8}$$

7.3 Failure-Time Distribution

After pre-specifying the critical level η_c, the failure-time T is defined as the time at which the mean degradation path $\eta(t, \boldsymbol{\theta})$ reaches the critical threshold η_c. For a monotonically decreasing degradation path (MDDP), a failure occurs when the degradation level decreases below the threshold η_c, i.e., $\eta(t, \boldsymbol{\theta}) \leq \eta_c$, and when $\eta(t, \boldsymbol{\theta}) \geq \eta_c$ for a monotonically increasing degradation path (MIDP). Suppose that the η is irreversible, continuous, and twice differentiable for $\boldsymbol{\beta}$ and $\mathbf{b}$, thus the η is a strictly monotonic function over time t. Furthermore, by assuring that $\eta(t, \boldsymbol{\theta}) = \eta_c$ has a unique and finite solution for any given parameters, $P[\lim_{t\to 0} \eta(t, \boldsymbol{\theta}) \leq \eta_c] = 0$ and $P[\lim_{t\to\infty} \eta(t, \boldsymbol{\theta}) \leq \eta_c] = 1$ for an MDDP, and $P[\lim_{t\to 0} \eta(t, \boldsymbol{\theta}) \leq \eta_c] = 1$ and $P[\lim_{t\to\infty} \eta(t, \boldsymbol{\theta}) \leq \eta_c] = 0$ for an MIDP. Based on those assumptions, there exists an inverse function $\eta^{-1}(t, \boldsymbol{\theta})$.

Under the NRC model (7.1), the failure-time distribution at a given degradation level and the distribution of degradation at a specific time are closely related. Let $F_\eta(y|t)$ be the conditional cumulative degradation distribution of $\eta(t, \boldsymbol{\theta})$ at a given time t, and $F_T(t|y)$ be the conditional cumulative failure-time distribution of T given degradation level y. The relationship between $F_\eta(y|t)$ and $F_T(t|y)$ is [20]:

$$F_\eta(y|t) \equiv \Pr[\eta(t, \boldsymbol{\theta}) \leq y] = \Pr\left[T \leq \eta^{-1}(y, \boldsymbol{\theta})\right] \equiv F_T(t|y). \tag{7.9}$$

Using the relationship (7.9), the failure-time distribution for the MDDP at a threshold level η_c is

$$\begin{aligned} F_T(t|\eta_c) &= \Pr[T \leq t|\eta_c] = \Pr\left[T \leq \eta^{-1}(\eta_c, \boldsymbol{\theta})\right] \\ &= \Pr[Y \leq \eta_c] = F_\eta(\eta_c|t) = \Phi\left(\frac{\eta_c - \mu(t; \boldsymbol{\theta})}{\sqrt{V(t; \boldsymbol{\theta})}}\right), \end{aligned} \tag{7.10}$$

where $\mu(t;\boldsymbol{\theta})$ and $V(t;\boldsymbol{\theta})$ are the mean and variance of the degradation data at a specific time t, respectively. Here, $\Phi(\cdot)$ is the cumulative distribution of a standard normal. For the MIDP, the failure-time distribution is

$$F_T(t|\eta_c) = 1 - F_\eta(\eta_c|t) = 1 - \Phi\left(\frac{\eta_c - \mu(t;\boldsymbol{\theta})}{\sqrt{V(t;\boldsymbol{\theta})}}\right).$$

The precision of the pth quantile (denoted by t_p) of the failure-time distribution is the main concern of reliability analysis. From (7.10), t_p for the MDDP is obtained by solving the following equation:

$$\eta_c = \mu(t;\boldsymbol{\theta}) + \Phi^{-1}(p)\sqrt{V(t;\boldsymbol{\theta})} \equiv h(t;p,\boldsymbol{\theta}), \tag{7.11}$$

and similarly, $\eta_c = \mu(t;\boldsymbol{\theta}) + \Phi^{-1}(1-p)\sqrt{V(t;\boldsymbol{\theta})} \equiv h(t;p,\boldsymbol{\theta})$ for the MIDP.

For $h(t;p,\boldsymbol{\theta})$, the pth quantile of the degradation at t for a strictly monotonic function, there exists a unique inverse function $h^{-1}(\eta_c;p,\boldsymbol{\theta})$ satisfying $t_p = h^{-1}(\eta_c;p,\boldsymbol{\theta})$ for a given $\boldsymbol{\theta}$. As a measure of design precision, the asymptotic variance of the pth quantile for the failure-time distribution is

$$Var(t_p) \equiv \sigma_p^2(\boldsymbol{\Xi}) = \frac{\partial h^{-1}(\eta_c;p,\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \mathscr{I}^{-1}(\boldsymbol{\Xi}|\boldsymbol{\Psi}) \frac{\partial h^{-1}(\eta_c;p,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \tag{7.12}$$

where $\mathscr{I}(\boldsymbol{\Xi}|\boldsymbol{\Psi})$ is the $(2p \times 2p)$ upper-left population Fisher information matrix. The precision of t_p depends on the information matrix via the population inspection scheme $\boldsymbol{\Xi}$.

7.4 Optimal Degradation Test Plan Under Cost Functions

The accuracy of reliability estimation for the products depends highly on the experimental design, especially when degradation paths are nonlinear. This section proposes an efficient degradation test plan in terms of cost while meeting experimental precision requirements. To reduce the experimental cost, we have generally allocated less experimental time and resources than the maximum capacity values. This incurs a loss of information, resulting in an increase in the prediction variability. Limited experimental resources and the demands of qualified experiments force us to solve a cost-effective optimization problem, which requires a balance between cost and precision.

7.4.1 Specification of the Degradation Test

The experimental cost and precision of the pth quantile of the failure-time distribution are two main concerns that require a trade-off balance. The important factors in conducting a degradation test are the population inspection schedule $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_Q\}$ where $\boldsymbol{\xi}_q = (t_{q1}, \ldots, t_{qn_q}),\ q = 1, \ldots, Q$ and the number of testing units $\{M_1, \ldots, M_Q\}$ where $M = \sum_{q=1}^{Q} M_q$. Because the precision of the pth quantile has a direct connection to the information loss, we relate the information loss to Type I error and Type II error. The question is how to determine the appropriate decision rule associated with the precision of the pth quantile. The use of an ideal experimental plan with all of available testing units and without cost consideration will allow for the determination as to whether to accept various cost-effective plans by comparing them to the ideal experimental plan in terms of cost and information loss. We assume that the pth quantile and its precision in the ideal plan are true values. To derive the cost-effective plan, we use the following assumptions:

- The maximum number of available units M_{max}, number of elementary design Q, and corresponding number of inspection times $n_q,\ q = 1, \ldots, Q$ are predetermined by the experimenter.
- Without cost consideration, an ideal inspection scheme $\boldsymbol{\Xi}_{\mathrm{I}}$ satisfying the D-optimality criterion and the M_{max} testing units will achieve the maximum attainable precision $\mathrm{Var}\left(\tilde{t}_p\right) = \sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_I)$ for unbiased expectation of the pth quantile, $\tilde{t}_p$. Let T_{I} be the true pth quantile obtained from the ideal plan.
- We select M of the M_{max} testing units and determine the cost-effective inspection scheme $\boldsymbol{\Xi}_C$. Because of measurement errors and sampling errors, the observed pth quantile T_{C} from the cost-effective plan differs from the true value T_{I}.
- For a given experimental variability limit ω, if the pth quantile for the cost-effective plan is outside of the acceptance region $\left(\tilde{t}_p - \omega,\ \tilde{t}_p + \omega\right)$, then we conclude that the experiment is rejected; otherwise, the experiment is accepted.

t_p in (7.11) and $\mathrm{Var}\left(t_p\right)$ in (7.12) are used to establish the formulation for quantifying the information loss. The pth quantile of the failure-time distribution for the ideal experiment (T_{I}) follows a normal distribution with mean $\tilde{t}_p$ and variance $\sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_I)$. Because we hypothetically assume that M units are randomly selected from the M_{max} testing units, the pth quantile for the cost-effective plan T_{C} depends on T_{I}. Let us assume that the conditional distribution of T_{C} given $T_{\mathrm{I}} = t_{\mathrm{I}}$, say $g\left(t_{\mathrm{C}}|t_{\mathrm{I}}\right)$, is a normal with mean t_{I} and variance $\sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_C)$. This implies that the pth quantile estimated from the cost-effective experiment is unbiased, and its variance $\sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_C)$ is independent of T_{I}. For clarity, we simply express $\sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_I)$ and $\sigma_{\mathrm{p}}^2(\boldsymbol{\Xi}_C)$ as σ_{pI}^2 and σ_{pC}^2 hereinafter, respectively. The joint density function of T_{C} and T_{I} is

$$f\left(t_{\mathrm{C}}, t_{\mathrm{I}}\right) = g\left(t_{\mathrm{C}}|t_{\mathrm{I}}\right) \cdot k\left(t_{\mathrm{I}}\right),$$

where

$$g\left(t_{\mathrm{C}}|t_{\mathrm{I}}\right)=\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{pC}}}\exp\left(-\frac{\left(t_{\mathrm{C}}-t_{\mathrm{I}}\right)^{2}}{2\sigma_{\mathrm{pC}}^{2}}\right),\quad\text{and}\quad k\left(t_{\mathrm{I}}\right)=\frac{1}{\sqrt{2\pi}\sigma_{\mathrm{pI}}}\exp\left(-\frac{\left(t_{\mathrm{I}}-\tilde{t}_{p}\right)^{2}}{2\sigma_{\mathrm{pI}}^{2}}\right).$$

Define Type I error as α and Type II error as β. Type I error is the probability of falsely rejecting the experiment which has to be accepted. In this context, the α is evaluated as

$$\begin{aligned}\alpha\left(\omega,\sigma_{\mathrm{pC}}\right)&=\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}\int_{-\infty}^{\tilde{t}_{p}-\omega}f\left(t_{\mathrm{C}},t_{\mathrm{I}}\right)dt_{\mathrm{C}}dt_{\mathrm{I}}+\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}\int_{\tilde{t}_{p}+\omega}^{\infty}f\left(t_{\mathrm{C}},t_{\mathrm{I}}\right)dt_{\mathrm{C}}dt_{\mathrm{I}}\\&=2\left[\Phi\left(\frac{\omega}{\sigma_{\mathrm{pI}}}\right)-0.5\right]-\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}k\left(t_{\mathrm{I}}\right)\cdot p_{g}\left(\omega,\sigma_{\mathrm{pC}}\right)dt_{\mathrm{I}},\end{aligned}\tag{7.13}$$

where

$$p_{g}\left(\omega,\sigma_{\mathrm{pC}}\right)=\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}g\left(t_{\mathrm{C}}|t_{\mathrm{I}}\right)dt_{\mathrm{C}}=\Phi\left(\frac{\tilde{t}_{p}+\omega-t_{\mathrm{I}}}{\sigma_{\mathrm{pC}}}\right)-\Phi\left(\frac{\tilde{t}_{p}-\omega-t_{\mathrm{I}}}{\sigma_{\mathrm{pC}}}\right).$$

Similarly, Type II error is the probability of falsely accepting the experiment which has to be rejected. Here, the β is evaluated as

$$\begin{aligned}\beta\left(\omega,\sigma_{\mathrm{pC}}\right)&=\int_{-\infty}^{\tilde{t}_{p}-\omega}\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}f\left(t_{\mathrm{C}},t_{\mathrm{I}}\right)dt_{\mathrm{C}}dt_{\mathrm{I}}+\int_{\tilde{t}_{p}+\omega}^{\infty}\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}f\left(t_{\mathrm{C}},t_{\mathrm{I}}\right)dt_{\mathrm{C}}dt_{\mathrm{I}}\\&=\int_{-\infty}^{\infty}k\left(t_{\mathrm{I}}\right)\cdot p_{g}\left(\omega,\sigma_{\mathrm{pC}}\right)dt_{\mathrm{I}}-\int_{\tilde{t}_{p}-\omega}^{\tilde{t}_{p}+\omega}k\left(t_{\mathrm{I}}\right)\cdot p_{g}\left(\omega,\sigma_{\mathrm{pC}}\right)dt_{\mathrm{I}}.\end{aligned}\tag{7.14}$$

The information losses caused by reducing the experimental time and resources are quantified by the probabilities of false decision, α and β.

7.4.2 Cost Functions

The total cost of an experiment can be assessed in terms of experimental cost and the cost of information loss. Experimental costs are related to direct expenses required to conduct a degradation test such as sample costs and inspection (measurement) costs. The termination time of the degradation test, also affecting the experimental cost, is a type of user-defined factor depending on the testing situation. The cost of information loss is incurred by wrong decisions (i.e., false rejection and false acceptance costs). We use the probabilities of false rejection and false acceptance to predict and quantify the information loss.

7.4.2.1 Experimental Cost

An important issue in the degradation test is how to reduce testing time. The operation cost can be formulated as a function of termination time in the inspection schedule. Park et al. [16] considered both production delay cost and opportunity costs which might be incurred by testing for longer time and assigned exponentially increasing penalty cost after a critical time point to the total testing cost. Following their approach to reflect real characteristics of experimental circumstances, we introduce the following operational cost function

$$C_T(t) = \begin{cases} c_1 + c_2 \cdot t & 0 \le t \le t_a \\ [c_1 + c_2 \cdot t_a] \cdot e^{c_3(t-t_a)} & t > t_a, \end{cases} \tag{7.15}$$

where c_1 is a fixed cost, c_2 is an operating cost per unit time and c_3 is a delay-penalty coefficient for the critical time t_a. Generally, if time-to-market (denoted by t_a) for new products cannot be satisfied, the opportunity cost will increase nonlinearly. Combining all factors related to experimental cost results in the function

$$\mathrm{EC}(\boldsymbol{\Xi}) = M\left(C_s + \sum_{q=1}^{Q} C_T(t_{qn_q})\right) + M\sum_{q=1}^{Q} n_q \cdot C_{\mathrm{I}} \tag{7.16}$$

where n_q is the number of inspections, C_{I} is an inspection cost, C_s is a sample cost, and t_{qn_q} is the final inspection time for schedule q.

7.4.2.2 Information Loss Cost

Reconsider the probabilities of false decision α and β, both of which are functions of ω and σ_{pC}, representing the information loss caused by reducing experimental time and resources. If the experiment is rejected even though it should be accepted, then Type I error α is noted. The cost spent in conducting the experiment will then be *sunk cost* and other incidental costs that may arise to conduct a new experiment must also be considered. If the experiment is accepted even though it should be rejected, then Type II error β is noted. In this case, the reliability prediction is biased, incurring excessive quality costs when the experimental results are used in product design or developing a quality assurance policy. In this paper, we introduce penalty cost coefficients for wrong decisions, C_α and C_β, which are multiplied by the probability of false decision.

Finally, the total cost function considering the experimental cost and the cost of information loss is given by

$$\mathrm{TC}(\boldsymbol{\Xi}, \omega) = \mathrm{EC}(\boldsymbol{\Xi}) \times \left[1 + \alpha(\omega, \sigma_{\mathrm{pC}}) \cdot C_\alpha + \beta(\omega, \sigma_{\mathrm{pC}}) \cdot C_\beta\right]. \tag{7.17}$$

This establishes a trade-off between experimental resources and experimental precision via the pth quantile for the failure-time distribution.

7.4.3 The Cost Optimization Problem

The objective of the problem is to determine the optimal combination of decision variables that minimizes the total cost. The optimization model can be formulated as

$$\begin{aligned} \text{Minimize} \quad & \text{TC}\left(\boldsymbol{\Xi}, \omega\right) \\ \text{Subjectto} \quad & \alpha\left(\omega, \sigma_{\text{pC}}\right) \leq \alpha_{\text{B}}, \\ & \beta\left(\omega, \sigma_{\text{pC}}\right) \leq \beta_{\text{B}}, \end{aligned} \tag{7.18}$$

where α_{B} and β_{B} are permissible Type I error and Type II error, respectively. Note that σ_{pC} is a function of $\boldsymbol{\Xi}_{\text{C}}$. ω is pre-determined by the experimenter. Without loss of generality, assume $Q = 1$ to simplify the model; that is, $\boldsymbol{\Xi} = [\boldsymbol{\xi}, M]$. However, the total cost function has a highly complex solution space in evaluating the Fisher information matrix in the nonlinear random-coefficients model, Thus, determination of the optimal combination of decision variables $\boldsymbol{\xi}^* = \left(t_1^*, \ldots, t_n^*\right)$ and M^* is a type of nonlinear, mixed continuous/integer optimization problem, hence we propose the following procedure:

Step 0: Initialize the problem parameters.
Step 1: Calculate the ideal design $\boldsymbol{\Xi}_I = [\boldsymbol{\xi}_I, M_{max}]$ to evaluate the information loss of a cost-effective design. Set $M = 2$.
Step 2: Find a minimum total cost design $\boldsymbol{\Xi}_C = [\boldsymbol{\xi}_C, M]$ subject to constraints for a given number of samples M. Let $G(M) = \text{TC}\left(\boldsymbol{\Xi}_C, \omega\right)$.
Step 3: Update $M = M + 1$. If $M \leq M_{max}$ return to **Step 2**, else go to **Step 4**.
Step 4: Find the cost-effective design $\boldsymbol{\Xi}_C^* = [\boldsymbol{\xi}_C^*, M^*]$ that has a minimum total cost. That is, $G(M^*) = \min_{2 \leq M \leq M_{max}} G(M)$.

To determine the cost-effective design with experimental quality constraints for a fixed number of samples, we use a modified elite-based hybrid genetic algorithm (GA), which was first proposed by Yen et al. [23]. The hybrid algorithm utilizes a local search technique within a elite-based genetic algorithm to compensate for the drawback of each method. In the proposed hybrid algorithm, a local search is applied to top elite chromosomes that are guided by a genetic algorithm. We modified the hybrid GA proposed by Yen et al. [23] to adapt for use in this research problem. This modified method uses the Nelder-Mead simplex method and a slightly different hybrid rule. Please see Kim and Bae [7] for details of the modified hybrid simplex-genetic algorithm.

7.5 Practical Application: PDP Example

In the PDP degradation analysis in [2], brightness is the main characteristic concern for the PDP degradation and industry standards identify a failure when the brightness of the display device falls below 50% of its initial brightness. Bae et al.

[2] proposed a bi-exponential model to describe the degradation paths of PDP brightness by incorporating two distinct causes of degradation phenomena, i.e., contamination effects of the impurities on the PDP electrode and the inherent degradation characteristic of phosphors. The bi-exponential degradation model is of the form

$$y(t) = \beta_1 \exp\left(-\exp\left(\beta_2\right) \cdot t\right) + \beta_3 \exp\left(-\exp\left(\beta_4\right) \cdot t\right) \quad t \geq 0.$$

Based on the analytical results in Bae et al. [2], we set the parameters in the bi-exponential degradation model as

$$\begin{aligned} y_{ij} = &\left(19 + b_{1i}\right) \exp\left(-\exp\left(-6 + b_{2i}\right) \cdot t_{ij}\right) \\ &+ \left(80 + b_{3i}\right) \exp\left(-\exp\left(-10 + b_{4i}\right) \cdot t_{ij}\right) + \epsilon_{ij}, \end{aligned} \tag{7.19}$$

where $\widehat{\boldsymbol{D}} = \text{Diag}\left(17, 0, 21, 0\right)$, and the variance of random error is $\hat{\sigma}^2 = 1$. We round the estimated parameter values in [2] to integers to easily compare and analyze the design results.

7.5.1 The Cost-Effective optimal plan for PDP Degradation Test

PDP is generally too expensive to allocate a large number of units, especially during the developing phase. The maximum number of available units M_{max} is assumed to be 20. As aforementioned previously, there is one elementary inspection schedule, $Q = 1$. Ermakov and Melas [6] showed useful properties of D-optimal designs for exponential regression models such as the bi-exponential model. First, the number of inspection points of a D-optimal design is equal to the number of parameters in the model. They called this a *saturated optimal design*. Second, a saturated D-optimal design is always unique and has equal weights at its inspection points. The number of inspection times n_q is assumed to be 4, which is the same as the number of fixed effects. Set $p = 0.1$, $\eta_c = 50$, $\underset{\sim}{\alpha}_{\text{B}} = 0.1$ and $\underset{\sim}{\beta}_{\text{B}} = 0.01$, and $\tilde{t}_{0.1}$ is estimated to be 8878.6 under prior model parameters. An ideal inspection scheme $\boldsymbol{\Xi}_1$ is $[\xi_1 = (0, 388.8, 2397.6, 25762.8), \; M_{\max} = 20]$ and its corresponding σ_{pI} is 409.9. An experimenter may impose restrictions on σ_{pC} to calculate ω. The experimental variability limit ω can then be determined from (7.13) and (7.14) given the parameters $\underset{\sim}{\alpha}_{\text{B}}$, $\underset{\sim}{\beta}_{\text{B}}$, and σ_{pC}. If the experimenter allows the σ_{pC} can be twice as large as σ_{pI}, i.e., $\sigma_{\text{pC}} = 2\sigma_{\text{pI}} \approx 820$, then $\omega = \max\{\omega | \underset{\sim}{\alpha}(\omega, 820) = 0.1, \underset{\sim}{\beta}(\omega, 820) = 0.01\} = \max\{1507.335, 928.231\}$. Thus, we select 1500 as ω for an acceptable variability range. Initial cost values are set as: $\left(C_s, C_1, C_{\underset{\sim}{\alpha}}, C_{\underset{\sim}{\beta}}, c_1, c_2, c_3, t_a\right) = (100, 10, 4, 40, 0, 0.2240, 9.3165 \times 10^{-4}, 4464)$. For example, if the time-to-market limitation t_a is 6 months, we can set

$t_a = 24 \cdot 31 \cdot 6 = 4464$. Similarly, $c_2 = 1000/(24 \cdot 31 \cdot 6) = 0.224$ denotes the operation cost of the experiment which linearly increases to 10 times the sample cost after 6 months, and $c_3 = \log 2/(24 \cdot 31)$ represents a double penalty imposed every month after t_a. In practice, the penalty costs associated with the mis-classification errors are unknown. When such information is not available, the receiver operating characteristic (ROC) analysis [22] can be used to select optimal decisions. However, if the mis-classification costs $C_{\underset{\sim}{\alpha}}$ and $C_{\underset{\sim}{\beta}}$ are assumed to be unknown, the cost optimization problem (7.18) becomes a complex problem to be solved. Therefore, we first assumed that the mis-classification costs are known a priori. The ROC analysis for the mis-classification costs might be done after determining the values of $\boldsymbol{\Xi}$ to search for the area under the ROC curve (see Wu and Xie [22] for more details).

Using the proposed degradation test plan under the FO method, the total cost of degradation testing was calculated as 10,267.64, including the cost-effective inspection scheme $\{\xi^*_{\text{FO}} = (0, 222.5, 1376.1, 4036.6),\ M^*_{\text{FO}} = 7\}$. Under the FOCE method via the simulation of 100 individuals, the total cost and corresponding inspection scheme are 10,267.17 and $\{\xi^*_{\text{FOCE}} = (0, 359.69, 1327.24, 4002.26),\ M^*_{\text{FOCE}} = 7\}$, respectively. The input parameters and resulting decision variables under the cost-effective degradation test plan are summarized in Table 7.1. The contour plots for the total cost evaluated using the FO and FOCE method, with respect to t_2 and t_3, are given in Fig. 7.1a, b, respectively, for fixed t_1 and t_4. Note that the total cost and the required number of testing samples are similar under the two different approximation methods.

To evaluate the effectiveness of the proposed degradation plan, we compared it with three heuristic inspection strategies that have been commonly used in practice: equal degradation (ED) plan, equal log-spacing (EL) plan, and equal spacing (ES) plan. The first two heuristic plans were introduced in [4] to derive an optimal design in accelerated degradation test. Table 7.2 summarizes the results of termination times and total costs under fixed numbers of testing samples at the initial cost values. Note that all of the results satisfy permissible error levels, $\underset{\sim}{\alpha}_{\text{B}} = 0.1$ and $\underset{\sim}{\beta}_{\text{B}} = 0.01$, except for cells marked with an asterisk. We confirmed that $\underset{\sim}{\beta}(\omega, \sigma_{\text{pC}})$ from all of the inspection schemes lies within the permissible Type II error $\underset{\sim}{\beta}_{\text{B}}$. The proposed degradation test plan with the FOCE method provides the shortest test duration,

Table 7.1 Input parameters and decision results for a PDP example

Input parameters	Decision results		
$(C_s, C_I, C_\alpha, C_\beta, c_1, c_2, c_3, t_a)$, $(\omega, \alpha_B, \alpha_B)$	Plan type	$\boldsymbol{\Xi} = (\boldsymbol{\xi}, M)$	Total cost
(100, 10, 4, 40, 0, 0.2240, 9.317e-4, 4464), (1500, 0.1, 0.01)	Proposed	FO={(0, 222.5, 1376.1, 4036.6) , 7}	10,268
		FOCE={(0, 359.7, 1327.2, 4002.3) , 7}	10,267
	ED	{(0, 295.2, 1126.5, 4118.0) , 7}	10,588
	EL	{(0, 19.9, 396.9, 7906.1) , 6}	209,495
	ES	{(0, 1468.9, 2937.8, 4406.8) , 9}	14,390

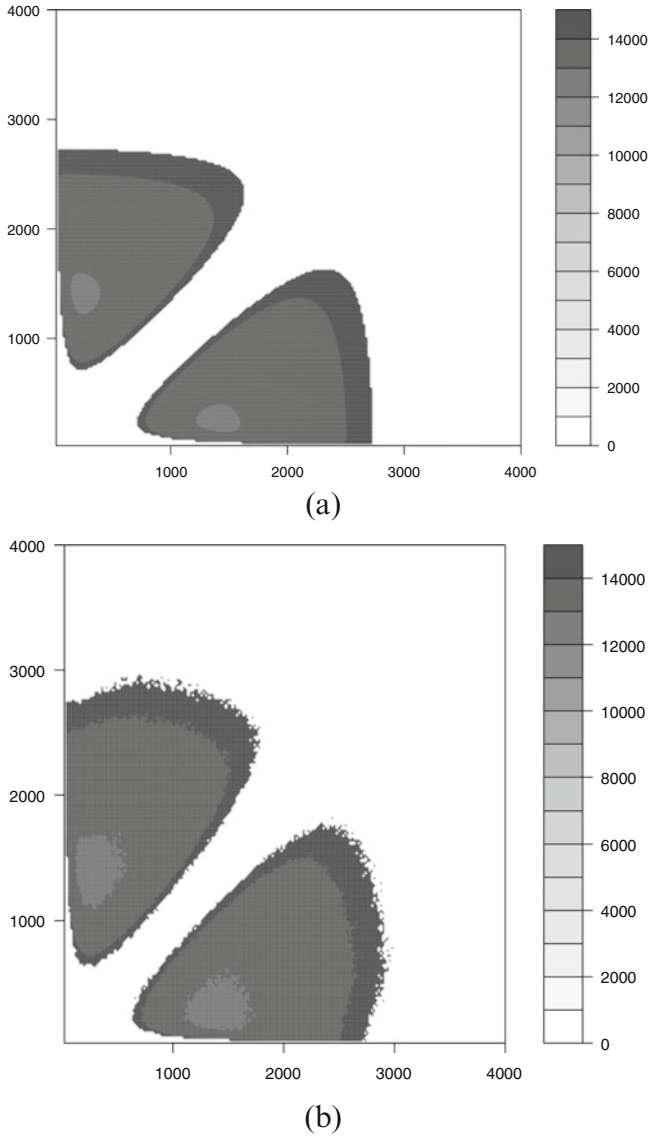

Fig. 7.1 Contour plot for the total cost evaluated using (**a**) the FO and (**b**) the FOCE method

while the EL plan requires the longest test duration to compensate for information losses, and incurs the largest total cost by exceeding the critical time $t_a = 4464$. Table 7.2 also indicates that both the FO and FOCE methods provide similar total cost and termination time according to the number of testing samples. From a design point of view, we consider only the FO method in the following sensitivity analysis because the FOCE method requires computationally intensive simulation work.

Table 7.2 Corresponding total costs and termination times for various inspection plans under fixed numbers of testing samples (underlined cost values represent the case that the evaluated $\alpha(\omega, \sigma_C)$ from each plan reaches at a permissible risk level α_B)

	Total cost					Termination time				
	Proposed plan					Proposed plan				
M	FO	FOCE	ED	EL	ES	FO	FOCE	ED	EL	ES
5	243,724.1	(*)	3,330,007.3	(*)	608,296.0	8268.72	(*)	11,077.22	(*)	9252.12
6	12,830.66	12,648.18	12,926.61	209,353.88	39,772.92	4811.58	4790.38	4820.38	7906.09	6118.04
7	10,267.64	10,267.17	10,447.07	210,702.88	20,916.13	4036.56	4002.26	4118.04	7746.56	5248.67
8	10,771.57	10,765.92	11,132.17	217,710.85	15,189.81	3654.07	3635.33	3797.46	7637.60	4747.05
9	11,429.17	11,468.48	11,947.40	226,313.71	13,111.22	3410.82	3433.26	3655.50	7552.15	4406.53
10	12,151.53	12,166.20	12,767.07	235,417.63	13,776.30	3236.91	3267.28	3618.30	7480.79	4155.07
11	12,890.45	12,932.03	13,577.77	244,622.73	14,474.61	3183.09	3093.36	3569.06	7419.13	3958.73
12	13,619.62	13,672.47	14,381.62	253,749.21	15,190.74	3145.51	3117.90	3529.10	7364.56	3799.98
13	14,340.48	14,345.86	15,179.73	262,715.75	15,915.90	3106.66	3086.25	3489.30	7315.45	3667.70
14	15,054.43	15,124.17	15,972.86	271,486.29	16,645.22	3069.14	2975.85	3454.37	7270.72	3555.85
15	15,762.07	15,797.10	16,761.40	280,042.33	17,370.08	3026.73	3051.29	3414.74	7229.54	3512.41
16	16,464.19	16,524.39	17,545.76	288,386.73	18,085.19	2992.35	2918.52	3381.30	7191.39	3468.54
17	17,161.27	17,243.59	18,326.16	296,515.81	18,791.53	2958.51	2857.86	3347.01	7155.75	3427.46
18	17,853.69	17,857.46	19,102.77	304,439.89	19,489.88	2929.26	2928.64	3314.95	7122.31	3390.08
19	18,541.59	18,582.06	19,875.74	312,167.72	20,180.89	2898.20	2845.24	3289.16	7090.82	3351.80
20	19,225.55	19,244.62	20,645.15	319,706.22	20,865.11	2867.00	2861.89	3259.20	7061.02	3317.45

Table 7.3 Cost-effective degradation plan under various cost values (C_s, t_a)

$(C_s,\ t_a)$	Total cost	$\xi^* = (t_1, t_2, t_3, t_4)$	M^*
(10, 2976)	12,296.03	(0, 215.4, 1369.5, 2996.4)	12
(10, 4464)	9382.74	(0, 221.5, 1375.4, 4036.6)	7
(10, 5952)	9382.72	(0, 222.5, 1372.4, 4036.6)	7
(100, 2976)	13,762.37	(0, 216.0, 1284.7, 3103.4)	11
(100, 4464)	10,267.64	(0, 222.5, 1376.1, 4036.6)	7
(100, 5952)	10,264.07	(0, 239.5, 1407.6, 4811.6)	6
(1000, 2976)	25,776.76	(0, 218.8, 1345.5, 3654.0)	8
(1000, 4464)	19,116.99	(0, 298.6, 1154.8, 4205.4)	7
(1000, 5952)	17,849.24	(0, 239.6, 1407.4, 4811.6)	6

7.5.2 Sensitivity Analysis

It is necessary to estimate the unknown model parameters and determine the optimization parameters (i.e., the values of cost functions, ω, and permissible levels) a priori to plan the cost-effective degradation testing scheme. We consider three cases for sensitivity analysis to examine the impacts of those parameters on the proposed degradation test plan.

7.5.2.1 The Effect of Cost Factors

The factors C_s and t_a are key in the trade-off between the number of samples and the inspection schedule in the degradation experiment. We varied the values of C_s as 10, 100, and 1000, and the values of critical time t_a as 4, 6, and 8 months to investigate their effects on the required size of testing samples and termination time. Table 7.3 lists the cost-effective degradation plans for PDPs under various values of the cost factors (C_s, t_a) with a total of four inspection points. The results show that the termination time t_4 is more sensitive to the changes in the critical time t_a than is the sample cost C_s. It can be also observed that the required size of testing samples is more affected by the t_a than is C_s. Note that, as the critical time t_a increases, the requisite testing duration increases, while the required size of testing samples decreases.

7.5.2.2 The Effect of Estimated Model Parameters

In the bi-exponential model, the most influential parameters in the evaluation of failure-time distribution are β_4, b_4, and σ^2 because they determine the long-term degradation paths in the model. We consider nine different scenarios with respect to β_4, b_4 and σ^2 to examine the sensitivity of the degradation inspection scheme. We determine the ranges of those parameters to be $-10.25 \leq \beta_4 \leq -9.75$, $0 \leq b_4 \leq 0.01$, and $0.99 \leq \sigma^2 \leq 1.01$, respectively. Table 7.4 compares results of the

Table 7.4 Cost-effective degradation test plan under various model parameter values with respect to (β_4, b_4, σ^2)

β_4	b_4	σ^2	Plan type	ξ^*				M^*	Total cost	Relative cost bias
				t_1	t_2	t_3	t_4			
−0.25	0.01	−0.01	Proposed	0	431.52	1338.18	4570.81	20	37,017.26	260.52%
			ED	0	283.916	1095.23	4628.56	20	38,833.18	271.71%
			EL	0	21.23	450.79	9570.97	20	3,469,093.69	1557.05%
			ES	0	1945.34	3890.68	5836.02	20	110,946.05	746.19%
−0.25	0	0	Proposed	0	219.11	1433.40	4439.50	15	25,064.34	144.11%
			ED	0	278.44	1047.76	4451.50	16	26,798.28	156.51%
			EL	0	21.07	443.99	9355.50	18	2,530,565.43	1108.75%
			ES	0	1688.61	3377.21	5065.82	20	55,728.58	325.04%
−0.25	0.01	0.01	Proposed	0	432.82	1341.74	4626.99	20	38,801.39	277.90%
			ED	0	285.51	1109.50	4679.82	20	40,548.57	288.13%
			EL	0	21.25	451.50	9593.77	20	3,545,213.36	1593.41%
			ES	0	1970.06	3940.13	5910.19	20	118,642.17	804.89%
0	0.01	0.01	Proposed	0	426.42	1303.09	4489.01	8	13,090.88	27.50%
			ED	0	310.31	1258.25	4506.39	8	13,278.84	27.11%
			EL	0	19.98	399.19	7975.84	8	298,126.58	42.40%
			ES	0	1503.51	3007.03	4510.54	11	18,320.72	39.73%
0	0	0.01	Proposed	0	224.40	1377.62	4058.21	7	10,315.53	0.47%
			ED	0	295.89	1132.44	4136.49	7	10,487.95	0.39%
			EL	0	19.92	396.99	7910.00	6	210,115.87	0.36%
			ES	0	1476.80	2953.60	4430.40	9	14,316.50	9.19%

(continued)

Table 7.4 (continued)

β_4	b_4	σ^2	Plan type	ξ^*				M^*	Total cost	Relative cost bias
				t_1	t_2	t_3	t_4			
0	0.01	−0.01	Proposed	0	424.68	1299.56	4436.32	7	12,755.08	24.23%
			ED	0	308.23	1239.34	4453.39	8	12,798.06	22.50%
			EL	0	19.97	398.79	7963.87	8	294,819.19	40.82%
			ES	0	1486.25	2972.50	4458.76	11	17,615.32	34.35%
0.25	0	0.01	Proposed	0	214.70	1246.67	2968.19	5	5543.61	46.01%
			ED	0	284.17	985.87	3208.50	5	5721.89	45.23%
			EL	0	18.20	331.23	6028.31	5	31,043.56	85.17%
			ES	0	1192.40	2384.79	3577.19	5	6588.73	49.75%
0.25	0	−0.01	Proposed	0	217.57	1238.70	2942.54	5	5503.63	46.40%
			ED	0	283.45	981.17	3193.15	5	5686.59	45.57%
			EL	0	18.19	330.99	6021.86	5	30,865.42	85.26%
			ES	0	1181.61	2363.21	3544.82	5	6536.04	50.15%
0.25	0.01	0	Proposed	0	359.74	1200.69	3396.47	5	6306.81	38.58%
			ED	0	293.89	1051.20	3414.87	5	6335.61	39.36%
			EL	0	18.44	340.19	6274.63	5	38,803.05	81.47%
			ES	0	1338.84	2677.68	4016.52	5	7279.14	44.48%

proposed degradation plan and the traditional heuristic plans in various scenarios, given the settings of cost values in Sect. 5.2.1. Let TC^* be the total cost evaluated by the propose degradation plan using the model parameter values in (7.19). The relative cost bias is defined by

$$\frac{|TC^*_{spe} - TC^*_{ori}|}{TC^*_{ori}} \times 100\ (\%)$$

where TC^*_{ori} and TC^*_{spe} are the total cost from the cost-effective degradation test plan with the original parameter values in (7.19) and that with the specified parameter values in Table 7.4, respectively. The proposed degradation test plan provides more cost-effective results than does the traditional heuristic test plans. As for the relative cost biases, the EL and ES plans are highly sensitive to the changes in the model parameter values. Especially, the requisite sample size M^* in the two heuristic test plans is highly sensitive to β_4.

7.5.2.3 The Effect of *p*th Quantiles

The pth quantile for the failure-time distribution and its variance are used to establish the formulation for quantifying the information loss. The cost-effective plans for various pth quantiles of interest are given in Table 7.5. The results indicate that the total cost from the degradation testing plan and requisite sample size greatly increase when the p deviates from 0.5. Meanwhile, the inspection schedule ξ^* is fairly robust to the p values. Summarizing the results in Tables 7.4 and 7.5, the cost-effective degradation plan mainly compensates for information losses by increasing the size of samples rather than by changing the inspection schedule.

Table 7.5 Cost-effective degradation test plan under various pth quantiles

	ξ^*					
p	t_1	t_2	t_3	t_4	M^*	Total cost
0.01	0	227.77	1278.40	3163.81	19	26,385.85
0.05	0	233.41	1365.50	4033.08	8	13,426.19
0.1	0	222.52	1376.13	4036.56	7	10,267.64
0.2	0	218.54	1407.89	4438.54	5	7942.39
0.5	0	213.22	1409.97	4255.48	5	7581.44
0.7	0	214.48	1434.90	4535.42	5	8462.32
0.9	0	217.57	1423.63	4270.08	8	12,287.67
0.95	0	217.83	1438.06	4462.35	9	14,393.80
0.99	0	220.54	1433.04	4308.67	14	22,157.05

7.6 Conclusion

Determining the number of samples and the inspection schedule under limited testing time and resources is an important issue in a degradation test. In this article, we propose a cost-effective degradation test plan under the precision constraint for the pth quantile of the failure-time distribution in the context of a nonlinear random-coefficients model. By relating degradation distribution to its failure-time distribution, the pth quantile for the failure-time distribution and its precision are directly calculated. We propose the method to quantify the information losses incurred by reducing testing resources. An optimization framework is then established to minimize total cost while keeping a balance between total cost and precision of the failure-time distribution. Unlike existing works on degradation test plann, the proposed approach has the following characteristics:

1. We introduce an ideal experimental plan satisfying D-optimality as a standard of comparison so that we can easily determine the precision limit of the cost-effective degradation test plan through the probability of false decisions.
2. In minimizing the total experimental cost while meeting certain precision constraints, we explicitly consider the time-to-market requirement in the cost function, as well as the trade-off between termination time and sample size.
3. A proposed degradation test plan for a nonlinear random-coefficients model does not depend on specific types of degradation models. In the proposed framework, we do not devise the optimal design for only the bi-exponential model used in this research, but also for other degradation models widely used in practice. In particular, for complicated degradation models with random-coefficients, the proposed framework is a novel approach for the design with cost-minimization objective and precision constraints.

References

1. Bae SJ, Kvam PH (2004) A nonlinear random-coefficients model for degradation testing. Technometrics 46:460–469
2. Bae SJ, Kim S-J, Kim MS, Lee BJ, Kang CW (2008) Degradation analysis of nano-contamination in plasma display panels. IEEE Trans Reliab 57:222–229
3. Beal SL, Sheiner LB (1992) NONMEM user's guide. University of California, NONMEM Project Group, San Francisco
4. Boulanger M, Escobar LA (1994) Experimental design for a class of accelerated degradation tests. Technometrics 36:260–272
5. Delyon B, Lavielle M, Moulines E (1999) Convergence of a stochastic approximation version of the EM algorithm. Ann Stat 27:94–128
6. Ermakov SM, Melas VB (1995) Design and analysis of simulation experiments. Springer, New York
7. Kim S-J, Bae SJ (2013) Cost-effective degradation test plan for a nonlinear random-coefficients model. Reliab Eng Syst Safe 110:68–79

8. Kuhn E, Lavielle M (2005) Maximum likelihood estimation in nonlinear mixed effects models. Comput Stat Data Ann 49:1020–1038
9. Lindstrom MJ, Bates DM (1990) Nonlinear mixed effects models for repeated measures data. Biometrics 46:673–687
10. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
11. Lu CJ, Meeker WQ, Escobar LA (1996) A comparison of degradation and failure-time analysis methods for estimating a time-to-failure distribution. Stat Sin 6:531–546
12. Matis J, Wehrly T (1979) Stochastic models of compartmental systems. Biometrics 35:199–220
13. Metzler CM (1971) Usefulness of the two-compartment open model in pharmacokinetics. J Am Stat Assoc 66:49–53
14. Marseguerra M, Zio E, Cipollone M (2003) Designing optimal degradation tests via multi-objective genetic algorithms. Reliab Eng Syst Safe 79:87–94
15. Mentre F, Mallet A, Baccar D (1997) Optimal design in random-effects regression models. Biometrika 84:429–442
16. Park JI, Baek SH, Jeong MK, Bae SJ (2009) Dual features functional support vector machines for fault detection of rechargeable batteries. IEEE Trans Syst Man Cybern C 39:480–485
17. Park JI, Bae SJ (2010) Direct prediction methods on lifetime distribution of organic light-emitting diodes from accelerated degradation tests. IEEE Trans Reliab 59:74–90
18. Sheiner LB, Rosenberg B, Melmon KL (1972) Modelling of individual pharmacokinetics for computer-aided drug dosage. Comput Biomed Res 5:441–459
19. Tseng ST, Wen ZC (2000) Step-stress accelerated degradation analysis for highly reliable products. J Qual Technol 32:209–216
20. Wu SJ, Shao J (1999) Reliability analysis using the least squares method in nonlinear mixed-effect degradation. Stat Sin 9:855–877
21. Wu SJ, Chang C (2002) Optimal design of degradation tests in presence of cost constraint. Reliab Eng Syst Safe 76:109–115
22. Wu SM, Xie M (2007) Classifying weak, and strong components using ROC analysis with application to burn-in. IEEE Trans Reliab 3:552–561
23. Yen J, Liao JC, Lee B, Randolph D (1998) A hybrid approach to modeling metabolic systems using a genetic algorithm and simplex method. IEEE Trans Syst Man Cybern B 28:173–191
24. Yu HF, Tseng ST (1998) On-line procedure for terminating an accelerated degradation test. Stat Sin 8:207–220
25. Yu HF, Tseng ST (1999) Designing a degradation experiment. Nav Res Log 46:689–706
26. Yu HF, Tseng ST (2004) Designing a degradation experiment with a reciprocal Weibull degradation rate. Qual Technol Quant M 1:47–63

Chapter 8
Optimal Designs for LED Degradation Modeling

Tzong-Ru Tsai, Yuhlong Lio, Nan Jiang, Hon Keung Tony Ng, and Ding-Geng (Din) Chen

Abstract The lifetime information of highly reliable products is usually very difficult to be obtained within an affordable amount experimental time through using traditional life testing methods. Because the benefit of lower manufacturing cost for many highly reliable products, manufacturers can offer more highly reliable products for implementing an accelerated degradation test under constant-stress loading conditions. In this chapter, lumen degradation measurements of high power light emitting diodes based on their cumulative damage are studied. The measurements are collected by the constant-stress accelerated degradation testing method with the stress loadings of ambient temperature and drive current. Each cumulative damage process is model by a Wiener process, of which drift parameter depends on the two stress loadings. General statistical inference of the model parameters and percentiles of the lifetime distribution of light emitting diodes is addressed, and approximate lower confidence bounds of the lifetime percentiles are evaluated using the Fisher information of maximum likelihood estimators. Based on the obtained maximum likelihood estimates of the model parameters, an optimal strategy for implementing a constant-stress accelerated degradation test is established. The optimal strategy can reach a compromised decision between the

T.-R. Tsai (✉)
Department of Statistics, Tamkang University, Tamsui District, New Taipei City 25137, Taiwan
e-mail: tzongru@gms.tku.edu.tw

Y. Lio • N. Jiang
Department of Mathematical Sciences, University of South Dakota, 57069, Vermillion, SD, USA
e-mail: Yuhlong.Lio@usd.edu; Nan.Jiang@usd.edu

H.K.T. Ng
Department of Statistical Science, Southern Methodist University, 75275, Dallas, TX, USA
e-mail: ngh@mail.smu.edu

D.-G. (Din) Chen
School of Social Work & Department of Biostatistics, University of North Carolina, 27599, Chapel Hill, NC, USA

Department of Statistics, University of Pretoria, Pretoria, South Africa
e-mail: dinchen@email.unc.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_8

experimental budget and estimation precision of reliability analysis. An algorithm is provided to search for the proposed optimal strategy. The proposed method is illustrated with an example of light emitting diodes.

Keywords Cumulative exposure model • Fisher information • Gamma process • Inverse Gaussian distribution • Wiener process

8.1 Introduction

Because of the rapid advances of manufacturing technology and sustained quality improvement effort, modern products have high reliability and prolonged lifetimes. For highly reliable products or devices, lifetime information usually is difficult to obtain by using traditional life testing methods, such as censored life testing experiments or truncated life tests, within an affordable amount time. In such case, engineers and experimenters are encouraged to use accelerated life test methods to overcome this difficulty. However, it is still possible that only a few failures of highly reliable products could be observed at the end of a life test even though the products are tested under accelerated conditions. Therefore, a modern accelerated degradation test (ADT) is introduced as an alternative to assess the lifetime characteristics of highly reliable products. ADT allows engineers to assess the potential life span of highly reliable products based on the degradation information because the degradation measurements provide lifetime characteristics of highly reliable products. Hence, both degradation and failure information obtained from ADTs could be used in reliability analysis.

To implement an ADT for highly reliable products, the first step is to choose proper stress variables for the test. Then, engineers need to select proper stress loading methods and a degradation model to start the degradation test and gather the degradation information over time. Temperature and voltage (or current) are two commonly used stress variables for many ADTs. Traditionally, the Arrhenius law model, power law model or exponential law model could be applied to connect the stress variables and the degradation parameters. A comprehensive discussion on the use of stress variables can be found in [14]. The stress loading method determines how the stress conditions should be applied for test units over time. The stress loading can be broadly classified into constant-stress loading, step-stress loading and progressive-stress loading conditions. A degradation model describes the behavior of the degradation characteristics as a function of time. The general degradation path model and the stochastic process model are two typical degradation models in practical ADT applications. Detailed information for these two degradation models can be found in [12, 36]. The practice of using the constant-stress ADT (CSADT) method or step-stress ADT (SSADT) method for reliability inference has been studied by many authors over the past few decades, for example, [8, 10, 12, 19–23, 25–27, 29–35]. The applications of SSADT method could save test resource. However, inferences on the lifetime distribution parameters under the

SSADT method usually developed under the assumption of cumulative exposure model, CEM.

The CEM assumes that the residual life of a test unit in a SSADT depends only on the cumulative exposure the unit has experimented with no memory of how this exposure was accumulated, see [6, 15]. This assumption is valid only when the degradation of the component quality characteristics can be well controlled under specific accelerated conditions. Modern products are more reliable through well manufacturing efforts. When an ADT involves highly reliable products as test units, their quality characteristics may degrade slowly even undergoing accelerated conditions. Engineers are forced to apply higher stress levels to accelerate the degradation speed of test units. Tsai et al. [26] mentioned that the quality degradation of highly reliable products could not be well controlled under higher stress loading conditions in an ADT. Hence, the CEM assumption cannot be validated in this situation. It becomes unrealistic to ignore the exposure effect that has accumulated from the former set-up stress conditions under the step-stress loading design. Therefore, the CEM assumption would be doubtful and might make the reliability analysis of highly reliable products inaccurate.

Light emitting diode (LED) is one of the highly reliable products developed in the past few decades. Significant advances have been reached for the manufacturing technologies of LED. Because of the low-energy consumption benefit, LEDs have been successfully applied as an indicator for lighting devices recently. Nowadays, LED lighting products are successfully replacing legacy lighting products. However, the reliability of LED is difficult to be evaluated due to the long lifetime of LED. The lumen of LED is highly dependent upon the LED lifetime and hence it can be a good surrogate variable for implementing the reliability evaluation of LED. Manufacturers or engineers are interested in studying the reliability evaluation method using the degradation information based on the lumen of LED.

When high power LED is used under an ADT experiment, engineers would like to use high stress conditions to accelerate lumen degradation speed for accumulating degradation information. It becomes unrealistic to ignore the cumulative exposure effect from previous stress levels over time. Moreover, the experimental setting for statistical inference about possible interaction between stress variables is difficult in a SSADT practically. The degradation of test unit is usually more sensitive to the changes of high accelerated conditions, but less sensitive to the changes of low accelerated conditions. Reliability analysis based on the degradation and failure information under higher accelerated conditions only may be biased due to the concern of over-stress. If the test time and the cost of test equipments are affordable, it is better to implement an ADT with lower and higher accelerated conditions separately, such that the degradation measurements contain sufficient information for reliability analysis.

Another problem raised from some existing ADT methods is to predict the potential lifetimes of highly reliable products through extrapolation. Detailed information about this problem was provided by [4, 5]. Since an ADT is usually terminated much earlier than the mean-time-to-failure (MTTF) of test units, engineers need to use extrapolation with the degradation information to predict the MTTF for a

given threshold of degradation level. Extrapolation is easy to implement, but it could result in a serious prediction error. Another feasible method to tackle this prediction problem is to estimate a lifetime percentile using the maximum likelihood estimation method based on a stochastic process model.

Traditionally, the statistical models for such a stochastic process ignore the threshold level of degradation and only involves the threshold information to evaluate the lower confidence bound of a lifetime percentile, see [10]. However, the threshold information is important and associated with the estimation quality. It is more reasonable to include the threshold information to obtain reliable MLE of the distribution parameters, and then using these MLEs to develop an approximate lower confidence bound of a lifetime percentile for reliability analysis.

To conduct an ADT, a well planned strategy is necessary to reach a compromised decision between the experiment budget and the estimation precision on the reliability inference. Some recent studies about the optimal ADT design methods can be found in [9–11, 24, 25, 31–33, 37]. Onar and Padgett [17] suggested a general approach based on locally penalized D-optimality, called LPD-optimality, for ADT designs. The LPD-optimality approach can simultaneously minimize the variances of the model parameter estimators. Using Bayesian Markov chain Monte Carlo method, [18] provided an efficient estimation method to obtain the MLEs of the multiple-step SSADT models based on Wiener and gamma processes, respectively. The ADT designs studied by [17, 18] were developed under one accelerating variable SSADT model. However, highly reliable products are often subject to multiple loading stresses and a pitfall in SSADT model is the memoryless assumption of the CEM. To incorporate several accelerating variables into ADT methods, [21, 22] provided hyper-cuboidal volume approaches for different CSADT models with several accelerating variables, but these methods did not cover the interaction effect of stress variables.

In these days, the manufacturing cost of LED has been significantly reduced. The low manufacturing cost of LED enables engineers to have a potential to use more LEDs for conducting CSADTs. The CSADT and SSADT are two major ADTs in practice. A SSADT usually uses less test units than a CSADT, however, a SSADT works under the memoryless assumption of the CEM. The CEM assumption could not be realistic for multiple-variable ADTs. Following the same approach by [26], we discuss an inference approach with the CSADT that incorporates both accelerating variables of ambient temperature and current as well as the interaction effect term, for the modeling of the lumen degradation of LED under the Wiener process. A general statistical inference procedure involving the threshold information is introduced and the MLEs of distribution parameters and percentiles are obtained. Moreover, an approximate lower confidence bound of a lifetime percentile is studied. The method is easily implemented and allows a flexible testing schedule with inconsistent inspection times or different termination times among experimental runs. The proposed method does not require the use of the extrapolation to predict the failure time of test unit. Planning ADT strategies under the CSADT model are also discussed in this chapter, such that the asymptotic variance of the $100p$th lifetime percentile estimator is minimized subject to a budget constraint.

The rest of this chapter is organized as follows. The CSADT model under the Wiener process is provided in Sect. 8.2. In Sect. 8.3, an optimal design of the CSADT are established by using the maximum likelihood estimation method and Fisher information. Moreover, an algorithm is provided to search an optimal ADT plan. An example regarding the lumen degradation of LED is used in Sect. 8.4 for illustrating the application of the proposed method. Finally, concluding remarks are given in Sect. 8.5.

8.2 Constant-Stress Accelerated Degradation Model

Let the damage process start at x_0 and be terminated at level x at time t, which can be denoted by level x_t under a loading level L. The damage increment $x_t - x_0$ is assumed to be a Wiener process with positive diffusion β^2 and drift ν_L, which is a function of L. It can be shown that $(x_t - x_0)$ has a normal distribution with mean $\nu_L t$ and variance $\beta^2 t$. The probability density function (pdf) of x_t can be defined by

$$f(x_t; x_0, t) = \frac{1}{\beta\sqrt{t}}\phi\left(\frac{x_t - x_0 - \nu_L t}{\beta\sqrt{t}}\right), \; x_t > 0. \tag{8.1}$$

The failure time, defined as the first passage time (S) for x_S over the threshold C, has an inverse Gaussian distribution with pdf

$$f_S(s; \nu, \beta) = \frac{C - x_0}{\beta\sqrt{2\pi s^3}} \exp\left\{-\frac{(C - x_0 - \nu_L s)^2}{2\beta^2 s}\right\}, \; s > 0. \tag{8.2}$$

Let μ_L denote the MTTF of test units under loading level L. Then μ_L can be expressed as $\mu_L = (C - x_0)/\nu_L$. Using the results from [13], which extend the conditioning argument of [3] for the position of a Gaussian process given the process below the threshold C, it can be shown that the pdf of x_t conditional on $x_t < C$ at $t \geq 0$ could be represented as the pdf of a truncated Wiener process as the following [19]:

$$f_X(x_t; t) = \frac{1}{\beta\sqrt{t}}\phi\left(\frac{x_t - x_0 - \nu_L t}{\beta\sqrt{t}}\right)\left[1 - \exp\left\{-\frac{2(C - x_0)(C - x_t)}{\beta^2 t}\right\}\right], \quad x_0 < x_t < C. \tag{8.3}$$

Assume that n units are randomly selected and put on a degradation tests simultaneously at the initial time $t_0 = 0$ until the termination times prescribed to n experimental units, respectively, are reached. Let the degradation model on the quality characteristics follow a Wiener cumulative damage process with $x_0 = 0$, simultaneously. A failure of test unit is defined as the accumulated damage passes over the threshold C. Let the number of failed units be p and the number of survival

units be q at the end of ADT, i.e., $n = p + q$. When an ADT is well developed, it is commonly assumed that the drift parameter of the Wiener process is affected only by the stress conditions. The higher the stress loading level, the higher the value of ν_L. In other words, a shorter MTTF for products is found under a higher stress loading level, or failures are expected to occur more frequently under the higher stress loading level.

Assumptions for implementing the ADT

(A1) The absolute ambient temperature (L'_{1i}) and the drive current (L'_{2i}) are used as the acceleration variables for the ADT to accelerate the degradation of LED lumen over time for $i = 1, 2, \cdots, k$.

(A2) n_i units are allocated to the ith stress loading level L'_i of the ADT, where $L'_i = (L'_{1i}, L'_{2i})$ and $i = 1, 2, \cdots, k$.

(A3) The two components, (L'_{1i}, L'_{2i}), of each stress loading level L'_i, are standardized, respectively, according to

$$L_{1i} = \frac{1/L'_{10} - 1/L'_{1i}}{1/L'_{10} - 1/L'_{1\mathrm{M}}} \text{ and } L_{2i} = \frac{\log(L'_{2i}) - \log(L'_{20})}{\log(L'_{2\mathrm{M}}) - \log(L'_{20})},\ i = 1, 2, \cdots, k,$$

where L'_{10} and L'_{20} are normal stress loading levels, and L'_{1M} and L'_{2M} are the highest stress loading levels before standardization. Therefore, $L_{10} = L_{20} = 0$, $L_{1\mathrm{M}} = L_{2\mathrm{M}} = 1$, $0 < L_{1i} \leq 1$ and $0 < L_{2i} \leq 1$, $i = 1, 2, \cdots, k$.

(A4) For the jth unit among q_i survival units under the stress loading level L_i, the accumulative damage levels are observed at times $t_{ijh}(t_{ij(h-1)} < t_{ijh})$, $h = 1, 2, \cdots m_{ij}$ and denoted by $x_{ij1}, x_{ij2}, \cdots x_{ijm_{ij}}$, which follow a Wiener process with the drift parameter ν_{L_i} and diffusion parameter β for $j = 1, 2, \cdots, q_i$, $i = 1, 2, \cdots, k$.

(A5) The drift parameter can be expressed in terms of the standardized loading levels via the generalized Eyring model (GEM),

$$\nu_{L_i} = \xi \exp\{\gamma_1 L_{1i} + \gamma_2 L_{2i} + \gamma_3 L_{1i} L_{2i}\},\ i = 1, 2, \cdots, k, \tag{8.4}$$

where $\xi = \exp(\gamma_0)$.

(A6) Let $t_{ij0} = 0$ and $x_{ij0} = 0$ for all test units under ADT. The cumulative damage for a failed unit is defined in terms of the percentage loss of LED lumen.

The GEM in (8.4) involves the Arrhenius law model, power law model and exponential law model as special cases if only one acceleration variable is utilized in ADT. Note that the GEM in (8.4) contains the interaction term [2, 26] and it is not a special case of the hyper-cuboidal volume approach in [21].

8.3 Statistical Inference

The lifetime of a test unit is defined as the first passage time for the cumulative damage over a given threshold C. Assume that p_i failed units have been observed with lifetimes, s_{il}, $l = 1, 2, \cdots, p_i$, respectively, in the run i and q_i units still survive at the termination time of run i for $i = 1, 2, \cdots, k$. Hence, $n_i = p_i + q_i$, $i = 1, 2, \cdots, k$. Let $n_{tot} = \sum_{i=1}^{k} n_i$, $\Theta = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \beta)$, $\triangle x_{ijh} = x_{ijh} - x_{ij(h-1)}$, $\triangle t_{ijh} = t_{ijh} - t_{ij(h-1)}$ and $y_{ijh} = (C - x_{ijh})(C - x_{ij(h-1)})/\triangle t_{ijh}$ for all i, j, h. The log-likelihood function for the failed units and damage measurements are respectively given by

$$\ell_F(\Theta) = \sum_{i=1}^{k} \sum_{l=1}^{p_i} \log f_S(s_{il}; \nu_L, \beta), \tag{8.5}$$

and

$$\ell_D(\Theta) \propto \sum_{i=1}^{k} \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} \log f_X(x_{ijh}; x_{ijh-1}, \triangle t_{ijh}). \tag{8.6}$$

The log-likelihood function based on all failed units and survival units $\ell(\Theta) = \ell_F(\Theta) + \ell_D(\Theta)$ can be expressed by

$$\begin{aligned} \ell(\Theta) \propto & -n_{tot} \log \beta - \sum_{i=1}^{k} \sum_{l=1}^{p_i} \frac{(C_0 - \nu_{L_i} s_{il})^2}{2\beta^2 s_{il}} \\ & - \sum_{i=1}^{k} \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} \left[\frac{(\triangle x_{ijh} - \nu_{L_i} \triangle t_{ijh})^2}{2\beta^2 \triangle t_{ijh}} - \log\left(1 - e^{-2y_{ijh}/\beta^2}\right) \right]. \end{aligned} \tag{8.7}$$

The justification of (8.7) is that the degradation scheme for ceasing observation depends only on the unit's previous degradation measurements (i.e. stopping time). For more information, reader can refer [1, 7, 21].

The MLEs of γ_0, γ_1, γ_2, γ_3 and β, labeled by $\hat{\gamma}_i$ for $i = 0, 1, 2, 3$ and $\hat{\beta}$, respectively, can be obtained by solving the equations of $\partial \ell(\Theta)/\partial \gamma_i = 0$ for $i = 0, 1, 2, 3$ and $\partial \ell(\Theta)/\partial \beta = 0$, simultaneously, where

$$\frac{\partial \ell(\Theta)}{\partial \gamma_0} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{j=1}^{p_i} (C - \nu_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - \nu_{L_i} \triangle t_{ijh}) \right\} \nu_{L_i}, \tag{8.8}$$

$$\frac{\partial \ell(\Theta)}{\partial \gamma_1} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{j=1}^{p_i} (C - \nu_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - \nu_{L_i} \triangle t_{ijh}) \right\} \nu_{L_i} L_{1i}, \tag{8.9}$$

$$\frac{\partial \ell(\Theta)}{\partial \gamma_2} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{j=1}^{p_i} (C - v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - v_{L_i} \triangle t_{ijh}) \right\} v_{L_i} L_{2i}, \tag{8.10}$$

$$\frac{\partial \ell(\Theta)}{\partial \gamma_3} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{j=1}^{p_i} (C - v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - v_{L_i} \triangle t_{ijh}) \right\} v_{L_i} L_{1i} L_{2i}, \tag{8.11}$$

and

$$\frac{\partial \ell(\Theta)}{\partial \beta} = -\frac{n_{tot}}{\beta} + \frac{1}{\beta^3} \sum_{i=1}^{k} \left(\sum_{l=1}^{p_i} \frac{(C - v_{L_i} s_{il})^2}{s_{il}} + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} \left[\frac{(\triangle x_{ijh} - v_{L_i} \triangle t_{ijh})^2}{\triangle t_{ijh}} - \frac{4 y_{ijh}}{e^{2 y_{ijh}/\beta^2} - 1} \right] \right). \tag{8.12}$$

Suppose the expected Fisher information matrix of the MLEs is $\mathbf{I}(\boldsymbol{\Theta})$, the asymptotic variances and covariances of the MLEs can be obtained from the entries of $\mathbf{I}^{-1}(\boldsymbol{\Theta})$. Unfortunately, the exact closed forms of the entries of the inverse of the expected Fisher information matrix are difficult to obtain. Therefore, we consider the observed Fisher information matrix, $\hat{\mathbf{I}}(\hat{\boldsymbol{\Theta}}) = \{-\partial^2 \log L(\boldsymbol{\Theta}) / \partial \theta_i \partial \theta_j\}|_{\boldsymbol{\Theta} = \hat{\boldsymbol{\Theta}}}$, $i, j = 1, 2, 3, 4, 5$. Given the values of p_i and q_i, $i = 1, 2, \cdots, k$, we have

$$\hat{\mathbf{I}}(\hat{\boldsymbol{\Theta}}) = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & 0 \\ a_2 & a_5 & a_6 & a_7 & 0 \\ a_3 & a_6 & a_8 & a_9 & 0 \\ a_4 & a_7 & a_9 & a_{10} & 0 \\ 0 & 0 & 0 & 0 & a_{11} \end{bmatrix}. \tag{8.13}$$

The entries of $\hat{\mathbf{I}}(\hat{\boldsymbol{\Theta}})$ and the entries of the Fisher information matrix $\mathbf{I}(\boldsymbol{\Theta})$ can be evaluated through using the following equations:

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_0^2} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{l=1}^{p_i} (C - 2 v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - 2 v_{L_i} \triangle t_{ijh}) \right\} v_{L_i},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_0 \partial \gamma_1} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{l=1}^{p_i} (C - 2 v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - 2 v_{L_i} \triangle t_{ijh}) \right\} v_{L_i} L_{1i},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_0 \partial \gamma_2} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{l=1}^{p_i} (C - 2 v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - 2 v_{L_i} \triangle t_{ijh}) \right\} v_{L_i} L_{2i},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_0 \partial \gamma_3} = \frac{1}{\beta^2} \sum_{i=1}^{k} \left\{ \sum_{l=1}^{p_i} (C - 2 v_{L_i} s_{il}) + \sum_{j=1}^{q_i} \sum_{h=1}^{m_{ij}} (\triangle x_{ijh} - 2 v_{L_i} \triangle t_{ijh}) \right\} v_{L_i} L_{1i} L_{2i},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_1^2} = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{\sum_{l=1}^{p_i}(C - 2\nu_{L_i}s_{il}) + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh})\right\}\nu_{L_i}L_{1i}^2,$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_1 \partial \gamma_2} = \frac{\partial \ell^2(\Theta)}{\partial \gamma_0 \partial \gamma_3},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_1 \partial \gamma_3} = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{\sum_{l=1}^{p_i}(C - 2\nu_{L_i}s_{il}) + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh})\right\}\nu_{L_i}L_{1i}^2L_{2i},$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_2^2} = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{\sum_{l=1}^{p_i}(C - 2\nu_{L_i}s_{il}) + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh})\right\}\nu_{L_i}L_{2i}^2,$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_2 \partial \gamma_3} = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{\sum_{l=1}^{p_i}(C - 2\nu_{L_i}s_{il}) + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh})\right\}\nu_{L_i}L_{1i}L_{2i}^2,$$

$$\frac{\partial \ell^2(\Theta)}{\partial \gamma_3^2} = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{\sum_{l=1}^{p_i}(C - 2\nu_{L_i}s_{il}) + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh})\right\}\nu_{L_i}L_{1i}^2L_{2i}^2,$$

$$\frac{\partial \ell^2(\Theta)}{\partial \beta^2} = \frac{n_{tot}}{\beta^2} - \frac{3}{\beta^4}\sum_{i=1}^{k}\left(\sum_{l=1}^{p_i}\frac{(C - \nu_{L_i}s_{il})^2}{s_{il}} + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}\frac{(\triangle x_{ijh} - \nu_{L_i}\triangle t_{ijh})^2}{\triangle t_{ijh}}\right)$$
$$+\frac{1}{\beta^4}\sum_{i=1}^{k}\sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}\frac{4y_{ijh}}{e^{2y_{ijh}/\beta^2} - 1}\left(3 - \frac{4y_{ijh}}{\beta^2}\frac{e^{2y_{ijh}/\beta^2}}{e^{2y_{ijh}/\beta^2} - 1}\right).$$

Since $E(s_{il}) = \mu_{L_i} = C/\nu_{L_i}$, and $E(\triangle x_{ijh}) = \nu_{L_i}\triangle t_{ijh}$. Then $E(C - 2\nu_{L_i}s_{il}) = -C$, and $E(\triangle x_{ijh} - 2\nu_{L_i}\triangle t_{ijh}) = -\nu_{L_i}\triangle t_{ijh}$. The entries of $\hat{\mathbf{I}}(\hat{\boldsymbol{\Theta}})$ and the entries of the Fisher information matrix $\mathbf{I}(\boldsymbol{\Theta})$ can be obtained as

$$a_1 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial \gamma_0^2}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_iC + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}\nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i},$$

$$a_2 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial \gamma_0 \partial \gamma_1}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_iC + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}}\nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i},$$

$$a_3 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_0\partial\gamma_2}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{2i},$$

$$a_4 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_0\partial\gamma_3}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i}L_{2i},$$

$$a_5 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_1^2}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i}^2,$$

$$a_6 = a_4,$$

$$a_7 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_1\partial\gamma_3}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i}^2L_{2i},$$

$$a_8 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_2^2}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{2i}^2,$$

$$a_9 = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_2\partial\gamma_3}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i}L_{2i}^2,$$

$$a_{10} = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\gamma_3^2}\right) = \frac{1}{\beta^2}\sum_{i=1}^{k}\left\{p_i C + \sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} \nu_{L_i}\triangle t_{ijh}\right\}\nu_{L_i}L_{1i}^2L_{2i}^2,$$

$$a_{11} = -E\left(\frac{\partial \ell^2(\Theta)}{\partial\beta^2}\right) = -\frac{n_{tot}}{\beta^2} + \frac{3}{\beta^4}\sum_{i=1}^{k}\left(\sum_{l=1}^{p_i} E\left[\frac{(C-\nu_{L_i}s_{il})^2}{s_{il}}\right] + \sum_{j=1}^{q_i} m_{ij}\beta^2\right)$$
$$-\frac{1}{\beta^4}\sum_{i=1}^{k}\sum_{j=1}^{q_i}\sum_{h=1}^{m_{ij}} E\left[\frac{4y_{ijh}}{e^{2y_{ijh}/\beta^2}-1}\left(3-\frac{4y_{ijh}}{\beta^2}\frac{e^{2y_{ijh}/\beta^2}}{e^{2y_{ijh}/\beta^2}-1}\right)\right].$$

Analogy to inference procedure proposed by [16], the $100p$th percentile of test units can be approximately obtained by

$$s_L(p) = \frac{\left[z_p\beta + \sqrt{z_p^2\beta^2 + 4C\nu_L}\right]^2}{4\nu_L^2}, \tag{8.14}$$

where z_p is the $100p$th percentile of the standard normal distribution. Because $\nu_{L_0} = \exp(\gamma_0)$ for the normal use condition, the MLE of $s_{L_0}(p)$ for the normal use condition can be given by

$$\hat{s}_{L_0}(p) = \frac{\left[z_p\hat{\beta} + \sqrt{z_p^2\hat{\beta}^2 + 4Ce^{\hat{\gamma}_0}}\right]^2}{4e^{2\hat{\gamma}_0}}. \tag{8.15}$$

The asymptotic variance of $\hat{s}_{L_0}(p)$ can also be obtained by using the Fisher information matrix and denoted by

$$\tilde{\sigma}^2(\hat{s}_{L_0}(p)) = \kappa^T \boldsymbol{I}^{-1}(\Theta)\kappa, \tag{8.16}$$

where

$$\kappa^T = \left[\frac{\partial s_{L_0}(p)}{\partial \gamma_0}, 0, 0, 0, \frac{\partial s_{L_0}(p)}{\partial \beta}\right],$$

$$\frac{\partial s_{L_0}(p)}{\partial \gamma_0} = \frac{1}{2e^{2\gamma_0}}\left[z_p\beta + \sqrt{z_p^2\beta^2 + 4Ce^{\gamma_0}}\right] \times \left[2Ce^{\gamma_0}\left(z_p^2\beta^2 + 4Ce^{\gamma_0}\right)^{-\frac{1}{2}} - (z_p\beta + \sqrt{z_p^2\beta^2 + 4Ce^{\gamma_0}})\right],$$

and

$$\frac{\partial s_{L_0}(p)}{\partial \beta} = \frac{1}{2e^{2\gamma_0}}\left[z_p\beta + \sqrt{z_p^2\beta^2 + 4Ce^{\gamma_0}}\right]\left[z_p(1 + z_p\beta\left(z_p^2\beta^2 + 4Ce^{\gamma_0}\right)^{-\frac{1}{2}})\right].$$

8.4 Optimal Strategy

To implement an ADT, engineers are interested in developing an optimal strategy that determines the sample size and termination time of each run. The length of the measuring time interval for all units in the run i, denoted by δ_i, can be determined according to the operation schedule. In practice, a constant length of the measuring time interval with $\delta_i = \delta$ is preferred due to administrative convenience. Based on the previous ADT experience for collecting the LED data set used by [26], the total cost contains three components, fixed cost, total operating cost and variable cost. The fixed cost is $c_0 \times n_{tot}$, where c_0 is the fixed cost per unit. Total operating cost can be presented by $c_{op}\sum_{i=1}^{k} t_i$, where c_{op} is the operating cost per unit time. The variable cost is based on the fact that using different stress loading on the test units in the laboratory incurs a different cost. It is assumed that an addition of one degree in temperature beyond the normal use temperature costs c_{L_1} dollars per unit

time and an addition of one milliampere in current costs c_{L_2} dollars per unit time for each test unit. Thus, variable cost in the ADT experiment can be presented as $c_{L_1}\sum_{i=1}^{k} t_i(L'_{1i}-L'_{10})+c_{L_2}\sum_{i=1}^{k}(n_i\times t_i\times L'_{2i})$. In reference of the fixed cost, operating cost and variable cost for the total cost function, the total cost can be expressed as

$$\mathrm{TC} = c_0 \times n_{tot} + c_{op}\sum_{i=1}^{k} t_j + c_{L_1}\sum_{i=1}^{k} t_i(L'_{1i} - L'_{10}) + c_{L_2}\sum_{i=1}^{k}(n_i \times t_i \times L'_{2i}). \quad (8.17)$$

In some cases, the variable cost may contain different components from the proposed cost model. For example, the variable cost could include operating cost, labor cost and cost incurred under different combinations of stress levels. The total cost model needs to be re-formulated if any composition of fixed cost, operating cost or variable cost is different from the proposed one. For some experiments, the unit cost of testing electronic device could vary with different mA levels due to the difficulty to control the lower current in the mA level. It requires more sophisticated and precision instruments to generate the current and maintain the stability of the current at a lower mA level. In such cases, the unit current cost could be changed with respect to the current level. A multiple-level current cost function is suggested for those situations.

Oftentimes, the scheduled time to finish the reliability analysis of highly reliable products via ADT is tight because manufacturers would like to promote products into the market as early as possible. Hence, an upper bound of experimental time, labeled by t_U, would be preassigned. Let $\mathbf{n} = (n_1, n_2, \cdots, n_k)$, $\mathbf{t} = (t_1, t_2, \cdots, t_k)$ and $\tilde{\sigma}^2(\mathbf{n}, \mathbf{t}) \equiv \tilde{\sigma}^2(\hat{s}_{pL_0})$. The optimal settings of $(\mathbf{n}^*, \mathbf{t}^*)$ can be determined such that $\tilde{\sigma}(\mathbf{n}, \mathbf{t})$ is minimized, subject to a given total budget, say ϕ_0. It is reasonable to assume that all units in each experimental run need to be measured at least once. Then, the optimal test plan $(\mathbf{n}^*, \mathbf{t}^*)$ can be obtained through solving a constraint optimization problem, such as System (8.18). Without loss of generality, let t_U be a positive multiple of the length of the measuring time interval, δ. The optimum scheduled experiment times for all k runs can be set up such that $\delta \le t_1, t_2, \cdots, t_k \le t_U$. For example, if an engineer wants to have the ADT containing 6 runs for the LED example in Sect. 8.5 to be implemented and ended by 36 weeks with $\delta = 6$ weeks or 1008 h to collect the degradation information. In this case, t_U is 36 weeks or 6048 h. The optimal experimental times schedule for the ADT can be determined from the set $\{1008 \le t_1, t_2, \cdots, t_k \le 6048\}$ through solving the following constraint optimization problem:

$$\begin{aligned} &\text{Minimize } \tilde{\sigma}(\mathbf{n}, \mathbf{t}) \quad (8.18)\\ &\text{Subject to}\\ &\quad \mathrm{TC} \le \phi_0,\\ &\quad \delta \le t_1, t_2, \cdots, t_k \le t_U,\\ &\quad n_i \ge 1,\ i = 1, 2, \cdots, k,\\ &\quad n_1 + \cdots + n_k = n_{tot}. \end{aligned}$$

Let B_t be the set of all possible combinations of $(t_1, t_2, \cdots, t_k)$ such that the conditions of $\{\delta \le t_1, t_2, \cdots, t_k \le t_U\}$ and TC $\le \phi_0$ are satisfied. The interval $[\delta, t_U]$ is divided into subintervals with equal length δ. Then, the partition points are $\delta, 2\delta, \cdots, t_U$. The collection of all possible subsets of $\{\delta, 2\delta, \cdots, t_U\}$ is labeled as B_t, defined by $B_t = \{\mathbf{t} | \delta \le t_1, t_2, \cdots, t_k \le t_U$ and TC $\le \phi_0\}$. Global searching from all possible solutions in order to obtain an optimal strategy on $(\mathbf{n}^*, \mathbf{t}^*)$ is time consuming and very difficult. We provide an simple algorithm to find the optimal ADT plan on $(\mathbf{n}^*, \mathbf{t}^*)$ if the equal number of highly reliable units are allocated for all experimental runs, that is, $n_i^* = n^*$, $i = 1, 2, \cdots, k$. Such design is easily operated by engineers to implement an ADT. Based on this setting, the optimal ADT plan $(\mathbf{n}^*, \mathbf{t}^*)$ can be reduced to $(n^*, \mathbf{t}^*)$. For each $\mathbf{t}^{(i)} = \left(t_1^{(i)}, \cdots, t_k^{(i)}\right)$ in B_t, an upper bound of sample size, denoted by $n_{U,\mathbf{t}^{(i)}}$, can be obtained by using the inequality, TC $< \phi_0$, in System (8.18). We then have

$$n_{U,\mathbf{t}^{(i)}} = \left\lfloor \frac{\phi_0 - c_{op} \sum_{j=1}^{k} t_j^{(i)} - c_{L_1} \sum_{j=1}^{k} t_j^{(i)} (L'_{1j} - L'_{10})}{k \times c_0 + c_{L_2} \sum_{j=1}^{k} (t_j^{(i)} \times L'_{2j})} \right\rfloor, \tag{8.19}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. Let

$$\Psi(n_{\mathbf{t}^{(i)}}^* | \mathbf{t}^{(i)}) = \min_{n=1,2,\cdots,n_{U,\mathbf{t}^{(i)}}} \tilde{\sigma}(n, \mathbf{t}^{(i)}),\ i = 1, 2, \cdots, n_{B_t}, \tag{8.20}$$

and

$$\Psi(n^*, \mathbf{t}^*) = \min_{i=1,2,\cdots,n_{B_t}} \Psi(n_{\mathbf{t}^{(i)}}^* | \mathbf{t}^{(i)}), \tag{8.21}$$

where $n_{\mathbf{t}^{(i)}}^*$ is the optimal sample size at times $\mathbf{t}^{(i)}$. Then the optimal ADT plan, $(n^*, \mathbf{t}^*)$, can be obtained through the following algorithm:

Algorithm

1. Determine the set B_t.
2. Find $n_{U,\mathbf{t}^{(i)}}$ and $\Psi(n_{\mathbf{t}^{(i)}}^* | \mathbf{t}^{(i)})$ via Eq. (8.19) for each $\mathbf{t}^{(i)}$ in B_t and Eq. (8.20), respectively.
3. The optimal ADT plan $(n^*, \mathbf{t}^*)$ is the solution such that Eq. (8.21) is satisfied.

8.5 Illustration Example

An ADT experiment with accelerating variables, the absolute ambient temperature in Celsius degree (°C) and the driven current in milliampere (mA), had been conducted on transistor outline can (TO-can) packaged LEDs from 2010 to 2011 in a laboratory of Taiwan. Measurements of luminous flux of LED source were collected

Table 8.1 The experimental design of ADT for LED

i	Runs	°C	mÅ	L_1	L_2	n_i	p_i	q_i
1	AC0	25	350	0.0000	0.0000	10	0	10
2	AC1	45	650	0.4377	1.0000	10	0	10
3	AC2	60	650	0.7315	1.0000	10	0	10
4	AC3	75	450	1.0000	0.4060	10	0	10
5	AC4	75	550	1.0000	0.7301	7	1	6
6	AC5	75	650	1.0000	1.0000	10	2	8

using KEITHLEY 2430 pulse source current meter with integrating sphere OL500 and a spectroradiometer CAS140B during the ADT. In this ADT experiment, the normal use condition was set at 25°C and 350 mA for LED units. Six experimental runs with the following stress loading combinations of two variables, (25, 350), (45, 650), (60, 650), (75, 450), (75, 550) and (75, 650) in (°C, mA), shown in Table 8.1 were used for the ADT. The collected lumen degradation data set had been studied by [26, 27]. Because the confidentiality agreement with the research institute, the original degradation measurements could not be reproduced. In this section, manipulative lumen degradation measurements for LED units under the same setting of the ADT experiment in the laboratory of Taiwan were generated for modeling and optimal strategy illustration. The manipulative lumen degradation measurements, under each stress loading combination of absolute ambient temperature and driven current shown in Table 8.1, are displayed in Fig. 8.1. In this study, the parameters of data inputs are $k = 6$, $p_i = 0$, and $q_i = 10$ for $i = 1, 2, 3, 4$; and $p_5 = 1$, $p_6 = 2$, $q_5 = 6$ and $q_6 = 8$. All stress loading levels are taken as the settings given in Table 8.1.

8.5.1 Statistical Inference

It should be noticed that with nonlinear Model (8.4) and Model (8.7), the likelihood Eq. (8.8), (8.9),(8.10), (8.11) and (8.12) are no longer a generalized linear model. Hence, even if the capacity of test equipment is sufficient to generate more observations, there is no statistical test for the interaction effects available for model Eq. (8.4). After converting the ambient temperatures into the absolute temperatures, the loading levels for absolute temperature and drive current are standardized according to the assumption (A3), respectively. The corresponding standard values for each pairs of stress-loadings from the six combinations are presented in Table 8.1. The normal use loading condition pairs is labeled by AC0 and the other five pairs of stress loading conditions are denoted by AC1-AC5, respectively. Moreover, according to the discussions in [27], the parameters used for modeling are γ_i, for $i = 0, 1, 2, 3$ and β. All degradation measurements are used to find MLEs for model parameters. Define r as the indicator for the $100(1 - r)\%$ loss of the total luminous flux of LED source, where $r = 0.5$ and 0.7 are considered here.

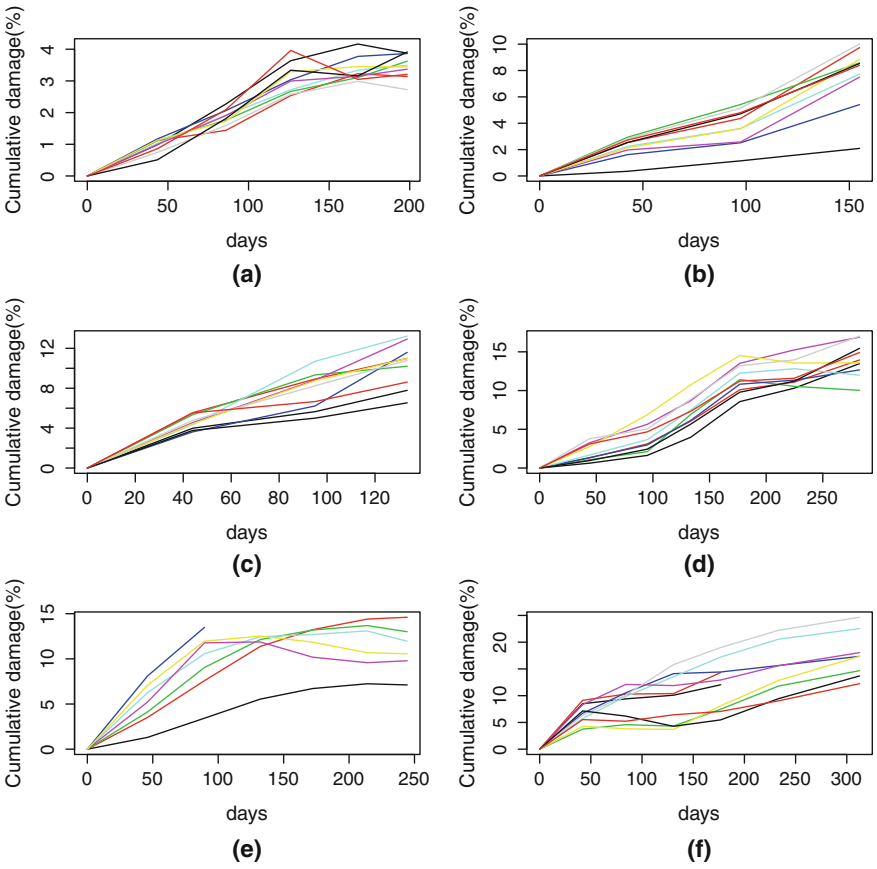

Fig. 8.1 Cumulative damage measurements of LED for experimental runs: (**a**) AC0, (**b**) AC1 (**c**) AC2, (**d**) AC3 (**e**) AC4, (**f**) AC5. Each line represents a degradation of a unit

Table 8.2 The MLE and estimated MTTF

r	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\beta}$	$\hat{s}_{L_0}(0.1)$	MTTF
0.7	−4.0476	0.8045	0.9898	−0.3956	0.2999	24,390.79	41,226.24
0.5	−4.0476	0.7456	0.7954	0.0325	0.3954	40,206.78	68,710.53

Based on the manipulative lumen degradation measurements for LED units under six runs displayed in Fig. 8.1, the MLEs, $\hat{\gamma}_i$, of γ_i for $i = 0, 1, 2, 3$ and the MLE, $\hat{\beta}$, of β were obtained and presented in Table 8.2. The tenth percentile of the lifetime distribution, $s_{L_0}(0.1)$, and the MTTF under the normal use loading condition can be estimated in hours by using the MLEs $\hat{\gamma}_i$, $i = 0, 1, 2, 3$ and $\hat{\beta}$ and they are also displayed in Table 8.2. It can be seen that under the normal use loading condition, the MLE $\hat{s}_{L_0}(0.1) = 24390.79$ h for $r = 0.7$ and $\hat{s}_{L_0}(0.1) = 40206.78$ h for $r = 0.5$. The MTTF of LEDs is estimated by 41226.24 h for $r = 0.7$ and 68710.53 h for $r = 0.5$.

8.5.2 Optimal Strategy

In practice, an ADT is often implemented to meet the time schedule for the introduction of the products in the market with an optimal level of the estimation precision from the reliability inference subject to a budget constraint. Engineers would like to reach a compromised design of the ADT. In this study, we consider that the length of the measuring time interval is taken as $\delta = 1008$ h, and the ADT will be ended at the 6048 h. The scheduled measuring times will then be in the interval [1008, 6048]. To implement the searching procedure for an optimal ADT strategy provided in Sect. 8.4, the interval [1008, 6048] is divided into 5 subintervals of length 1008. Hence, the partition points are 1008, 2016, 3024, 4032, 5040 and 6048. It follows that $B_t = \{\mathbf{t} | 1008 \leq t_1, \cdots, t_6 \leq 6048 \text{ and } \text{TC} \leq \phi_0\}$.

It is assumed that a LED is classified as failure if 30% luminous flux is lost from the initial amount. The model parameters, γ_0, γ_1, γ_2, γ_3 and β in the lifetime distribution are replaced by their MLEs, $\hat{\gamma}_0$, $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$ and $\hat{\beta}$, respectively. In addition, the costs of experimentation are taken as $c_{L_2} = 0.001$, $c_{L_1} = d_1 \times c_{L_2}$, $d_1 = 5, 10$, $c_0 = 50$, $c_{op} = 0.6$ with lot sizes $N_D = 1000, 5000$, experiment budgets $\phi_0 = 10{,}000, 20{,}000$, and the reference lifetime percentiles of $p = 0.1$ and $p = 0.5$ for the ADT, respectively. Optimal ADT strategies are obtained via the algorithm proposed in Sect. 8.4. The results are presented in Table 8.3. From Table 8.3, it can be seen that $\tilde{\sigma}(n^*, \mathbf{t}^*)$ is decreased if a larger experiment budget is available for the ADT. Hence, the optimal ADT strategy would be established so that a total cost is close to the experiment budget if it is possible. Table 8.3 shows that increasing the

Table 8.3 Optimal ADT strategies on 10th and 50th percentiles

p	N_D	d_1	ϕ_0	n^*	t_1^*	t_2^*	t_3^*	t_4^*	t_5^*	t_6^*	TC	$\tilde{\sigma}(n^*, \boldsymbol{t}^*)$
0.1	1000	5	10,000	31	252	168	126	126	42	42	9990.25	146.81
			20,000	64	252	84	252	42	42	42	19,996.95	102.18
		10	10,000	31	252	168	126	42	42	42	9973.30	146.82
			20,000	64	252	252	42	42	42	42	19,998.44	102.18
	5000	5	10,000	31	252	168	126	126	42	42	9990.15	146.81
			20,000	64	252	84	252	42	42	42	19,996.95	102.18
		10	10,000	31	252	168	126	42	42	42	9973.30	146.82
			20,000	64	252	252	42	42	42	42	19,998.44	102.18
0.5	1000	5	10,000	31	252	126	42	210	42	42	9957.76	333.69
			20,000	64	252	168	42	126	42	84	19,998.12	232.24
		10	10,000	31	252	84	42	126	84	42	9955.28	333.69
			20,000	64	252	84	42	126	42	84	19,983.25	232.24
	5000	5	10,000	31	252	126	42	210	42	42	9957.76	333.69
			20,000	64	252	168	42	126	42	84	19,998.12	232.24
		10	10,000	31	252	84	42	126	84	42	9955.29	333.69
			20,000	64	252	84	42	126	42	84	19,983.25	232.24

sample size in each experiment run could decrease the value of $\tilde{\sigma}(n^*, \mathbf{t}^*)$ if a higher experiment budget is allowed.

The optimal ADT strategies provide a more accurate estimate (i.e., $\tilde{\sigma}(n^*, \mathbf{t}^*)$) for the $100p$th percentile with smaller p compared to larger p. The difference of total costs for estimating the 10th and the 50th percentiles is insignificant. Finally, Table 8.3 reveals that the total cost with $d_1 = 5$ is slightly larger than the total cost with $d_1 = 10$ under the condition of $\phi_0 = 10{,}000$. However, the total cost with $d_1 = 5$ is little bit smaller than that with $d_1 = 10$ under the budget condition of $\phi_0 = 20{,}000$. The difference between two total costs for different values of d_1 is insignificant. Hence, the effect of increasing the variable cost on unit absolute ambient temperature 5 times or higher over the variable cost on unit driven current is insignificant.

8.6 Conclusion and Discussion

In this chapter, modeling the lumen degradation of LED via Wiener diffusion process are discussed. The drift parameter in Wiener process is expressed as an exponential function in terms of two standardized stress variables via a GEM, which is different from the one proposed by [19] and slightly different from the hyper-cuboidal volume approach by [21] for modeling fatigue crack growth and cumulative damage. It should be mentioned that the degradation process can be very slow for highly reliable product such as the LED system. Therefore, any two consecutive degradation measurements from a unit of LED had been taken at two widely separated time schedules in the laboratory. The flaw in using the Wiener assumption is that the process increment follows a normal distribution and could possibly generate negative damage value, see [20]. However, it could be convinced that the probability of negative increment can be negligible based on the setting of degradation measurement time schedules for each test unit in the laboratory. Regardless of the pitfall of Wiener process due to negative increment, the current modeling of the lumen degradation using Wiener process has provided satisfactory results on the estimations of lower percentiles of lifetimes and MTTF based on the discussions with LED engineers. Next, a drawback of using a Wiener process is the possibility of getting a negative cumulative damage. A planning strategy under Gamma stochastic processes with warranties can be applied. More detail had been investigated by [28]. R source codes are provided in the Appendix to search an optimal ADT plan with plug-in MLEs of the model parameters. User can use these codes to generate the optimal ADT plans in the Table 8.3.

Optimal strategies under different combinations of cost components and two percentiles of concern have been constructed and summarized in Table 8.3 for reference. The proposed method can provide a reasonable planning strategy for the accelerated degradation testing procedure proposed by [26]. Nowadays, warranties are important for marketing highly reliable products since a good warranty policies make the products more competitive in the market. Considering a warranty cost

for the planning strategy of the accelerated degradation testing experiment is an important consideration for manufacturers. Investigations related to warranty analysis will be an interesting future research direction.

Appendix

```
#####################################################
##  The optimal design of an ADT using plug-in MLEs
#####################################################

rm(list=ls(all=TRUE))
library(SuppDists)

##
## Define parameters and cost components
##

m=k=6               ## all test units are survival in 6 runs
C=(1-0.70)*100      ## failure defined by 30% luminous loss
pp=0.1              ## 100*0.1 percentile
tim.u=24            ## time unit per day
phi_0=10000
US=1
d1=5
c_L2=0.001/US
c_L1=d1*c_L2
c_0=50/US
N_D=1000
c_op=0.6/US
pi=c(0,0,0,0,0.1,0.2)
delta=2
L11o=c(25,45,60,75,75,75)+273.15
L22o=log(c(350,650,650,450,550,650))
L11=(1/L11o[1]-1/L11o)/(1/L11o[1]-1/L11o[6]);
L22=(L22o-L22o[1])/(L22o[6]-L22o[1])
L1=L11[1:6]; L2=L22[1:6]

zp=qnorm(pp)    ## (100*pp)th percentile of N(0,1).
ta=c(1008,2016,3024,4032,5040,6048)/tim.u

##
## The MLEs
##

r0=-4.0476; r1=0.8045; r2=0.9898; r3=-0.3956
be=0.2999          ## for failure defined by 30% luminous loss
## be=0.3954       ## for failure defined by 50% luminous loss
lam=C^2/be^2
```

```
nu=exp(r0+r1*L1+r2*L2+r3*L2*L1)

muLi=C/nu; muL0=muLi[1]

## Define functions

tE=expand.grid(ta,ta,ta,ta,ta,ta)
t=matrix(NA,(length(ta)^6),6)
for(i in 1:(length(ta)^6)){
 for(j in 1:6) {t[i,j]=tE[i,j]}
}

##
########## Function of AS(n,t)
##
AS=function(nn,tt){
 p=round(nn*pi)
 q=nn-p
 a1=(1/(be^2))*( sum(p*C*nu)+sum(nu^2*tt*q) )
 a2=(1/(be^2))*( sum(p*C*nu*( L1 ))+sum(nu^2*tt*q*( L1 )) )
 a3=(1/(be^2))*( sum(p*C*nu*( L2 ))+sum(nu^2*tt*q*( L2 )) )
 a4=(1/(be^2))*( sum(p*C*nu*( L1*L2 ))
      +sum(nu^2*tt*q*( L1*L2 )) )
 a5=(1/(be^2))*( sum(p*C*nu*( L1^2 ))
      +sum(nu^2*tt*q*( L1^2 )) )
 a6=a4
 a7=(1/(be^2))*( sum(p*C*nu*( L1^2*L2 ))
      +sum(nu^2*tt*q*( L1^2*L2 )) )
 a8=(1/(be^2))*( sum(p*C*nu*( L2^2 ))
      +sum(nu^2*tt*q*( L2^2 )) )
 a9=(1/(be^2))*( sum(p*C*nu*( L1*L2^2 ))
      +sum(nu^2*tt*q*( L1*L2^2 )) )
 a10=(1/(be^2))*( sum(p*C*nu*( L1^2*L2^2 ))
      +sum(nu^2*tt*q*( L1^2*L2^2 )) )

 EInv=array()
 for(i in 1:m){
  integrandEInv=function(x) {((C-nu[i]*x)^2/x)
                *dinvGauss(x,muLi[i],lam)}
  EInv[i]=integrate(integrandEInv,lower = 0,upper=Inf)$value}

 mij=ceiling(tt/delta)
 N=sum(p)+sum(mij*q)
 a11=-N/be^2+3/be^4*sum(p*EInv+q*mij*be^2)
#---
 F.mat=array(c(a1,a2,a3,a4,0,a2,a5,a6,a7,0,a3,a6,a8,a9,0,a4,
       a7,a9,a10,0,0,0,0,0,a11), dim=c(5,5))
 F.mat=round(F.mat,4)

 a.inv=solve(F.mat)
 pc=( (zp*be+sqrt(zp^2*be^2+4*C*exp(r0)))*
     ((2*C*exp(r0))*(zp^2*be^2+4*C*exp(r0))^(-1/2)
      -(zp*be+sqrt(zp^2*be^2+4*C*exp(r0)))) )/(2*exp(2*r0))
 pb=((zp*be+sqrt(zp^2*be^2+4*exp(r0)*C))
```

```
      *(zp*(1+zp*be*(zp^2*be^2+4*exp(r0)*C)^(-1/2))))
      /(2*exp(2*r0))
 Avar=matrix(c(pc,0,0,0,pb),1,5)%*%a.inv
      %*%matrix(c(pc,0,0,0,pb),5,1)
 sqrt(Avar)
}

##
## The function to reach an optimal plan
##    of (n_{tot},t1,\ldots,tk,TC,sigma(n,t))
##
OA=function(phi_0){
 ans2=array(Inf)
 for(B in 1:nrow(t)){
  tk=min(t[B,])
  nU=floor((phi_0-c_op*sum(t[B,])-c_L1*sum(t[B,]
     *(L11o-(25+273.15))))/(k*c_0+c_L2*sum(t[B,]*L22o)))
  ans1=array(0)
 for(inu in 1:nU){
   {if(inu<=0){next}}
   ans1[inu]=AS(inu,t[B,])
  }
 minn1=which.min(ans1)
 varCost=sum(t[B,]*(c_L1*(L11o-(25+273.15))
         +minn1*c_L2*L22o))
 TC=k*minn1*c_0+c_op*sum(t[B,])+varCost
 ans2=rbind(ans2,c(minn1,t[B,],TC,ans1[minn1]))
 }
 dataMN=which(ans2[,9]==min(ans2[,9]))
 ans2[dataMN,]  ## print all solution sets
}

##
## Print the output
##

cat("p=",pp, "N_D=", N_D, "d1=", d1, "\n", "\n")
OA(phi_0)
```

References

1. Andersen PK, Borgan O, Gill RD, Keiding N (1992) Statistical models based on counting processes. Springer, New York
2. Barbosa EP, Colosimo EA, Louzada-Neto F (1996) Accelerated life tests analyzed by a piecewise exponential distribution via generalized linear models. IEEE Trans Reliab 45:619–623
3. Cox DR, Miller HD (1965) The theory of stochastic processes. John Wiley & Sons, New York

4. Duan F, Cooper S, Marathe A, Zhang J, Jayanarayanan SK (2006) Impact of monitoring voltage on the lifetime extrapolation during the accelerated degradation tests. Integrated Reliability Workshop, pp 139–141
5. Escobar LA, Meeker WQ (2006) A review of accelerated test models. Stat Sci 21:552–577
6. Komori Y (2006) Properties of the Weibull cumulative exposure model. J Appl Stat 33:17–34
7. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
8. Liao HT, Elsayed EA (2004) Reliability prediction and testing plan based on an accelerated degradation rate model. Int J Mater Prod Technol 21:402–422
9. Liao CM, Tseng S-T (2006) Optimal design for step-stress accelerated degradation tests. IEEE Trans Reliab 55:59–66
10. Lim H, Yum B-J (2011) Optimal design of accelerated degradation tests based on Wiener process models. J Appl Stat 38:309–325
11. Liu X, Tang L-C (2010) A Bayesian optimal design for accelerated degradation tests. Qual Reliab Eng Int 26:863–875
12. Lu C-J, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
13. Lu J (1995) Degradation processes and related reliability models. Unpublished PhD Thesis, McGill Universrity, Canada
14. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. John Wiley & Sons, New York
15. Miller R, Nelson B (1983) Optimal simple step stress plans for accelerating life testing. IEEE Trans Reliab R-32:59–65
16. Onar A, Padgett WJ (2000) Accelerated test methods with the inverse Gaussian distribution. J Stat Plan Inference 89:119–133
17. Onar A, Padgett WJ (2004) A penalized local D-optimality approach to design for accelerated test models. J Stat Plan Inference 119:411–420
18. Pan Z, Balakrishnan N (2010) Multiple-steps step-stress accelerated degradation modeling based on Wiener and gamma processes. Commun Stat-Simul Comput 39:1384–1402
19. Padgett WJ, Tomlinson MA (2004) Inference from accelerated degradation and failure data based on Gaussian process models. Lifetime Data Anal 10:191–206
20. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Anal 11:511–527
21. Park C, Padgett WJ (2006) Stochastic degradation models with several accelerating variables. IEEE Trans Reliab 55:379–390
22. Park C, Padgett WJ (2007) Cumulative damage models for failure with several accelerating variables. Qual Technol Quant Manag 4:17–34
23. Park C, Padgett WJ (2008) Cumulative damage models based on gamma processes. In: Ruggeri F, Faltin F, Kenett R (eds) Encyclopedia of statistics in quality and reliability. John Wiley & Sons, New York, pp 1–5
24. Peng J-Y (2012) A note on optimal allocations for the second elementary symmetric function with applications for optimal reliability design. Nav Res Logist 59:278–284
25. Tsai C-C, Tseng S-T, Balakrishnan N (2011) Optimal burn-in policy for high reliable products using gamma degradation process. IEEE Trans Reliab 60:234–245
26. Tsai T-R, Lin C-W, Sung Y-L, Chou P-T, Chen C-L, Lio YL (2012) Inference from lumen degradation data under Wiener diffusion process. IEEE Trans Reliab 61:710–718
27. Tsai T-R, Lio YL, Jiang N (2014) Optimal decisions on the accelerated degradation test plan under the Wiener process. Qual Technol Quant Manag 11:461–470
28. Tsai T-R, Sung W-Y, Lio YL, Chang SI, Lu J-C (2016) Optimal two-variable accelerated degradation test plans for gamma degradation processes. IEEE Trans Reliab 65:459–468
29. Tseng S-T, Wen ZC (2000) Step-stress accelerated degradation analysis for highly reliable products. J Qual Technol 32:209–216
30. Tseng S-T, Peng C-Y (2007) Stochastic diffusion modeling of degradation data. Data Sci 5:315–313

31. Tseng S-T, Balakrishnan N, Tsai C-C (2009) Optimal step-stess accelerated degradation test plan for gamma degradation process. IEEE Trans Reliab 58:611–618
32. Tseng S-T, Tsai C-C, Balakrishnan N (2011) Optimal sample size allocation for accelerated degradation test based on Wiener process. In: Balakrishnan N (ed) Methods and applications of statistics in engineering, quality control, and the physical sciences. Wiley, New York, pp 330–343
33. Tseng ST, Lee IC (2016) optimum allocation rule for accelerated degradation tests with a class of exponential-dispersion degradation models. Technometrics 58:244–254
34. Weaver BP, Meeker WQ (2014) Methods for planning repeated measures accelerated degradation tests. Appl Stoch Model Bus Ind 30:658–671
35. Weaver BP, Meeker WQ, Escobar LA, Wendelberger J (2013) Methods for planning repeated measures degradation studies. Technometrics 55:122–134
36. Wu S-J, Chang C-T (2002) Optimal design of degradation tests in presence of cost constraint. Reliab Eng Syst Saf 76:109–115
37. Ye Z-S, Xie M, Tang L-C, Shen Y (2012) Degradation-based burn-in planning under competing risk. Technometrics 54:159–168

Chapter 9
Gamma Degradation Models: Inference and Optimal Design

N. Balakrishnan, Chih-Chun Tsai, and Chien-Tai Lin

Abstract During the past two decades, degradation analysis has been widely used to assess the lifetime information of highly reliable products. Usually, random effect models and/or Wiener processes are well suited for modelling stochastic degradation. But in many situations, such as materials that lead to fatigue, it is more appropriate to model the degradation data by a gamma degradation process which exhibits a monotone increasing pattern. This article surveys the theoretical aspects as well as the application of gamma processes in degradation analysis. Some statistical properties of degradation models based on gamma processes under different tests are also given. Furthermore, the corresponding optimal designs for conducting the degradation experiments efficiently are reviewed. Finally, some extensions and their applications of gamma-process degradation model are presented.

Keywords Accelerated degradation test • Burn-in test • Degradation data • Gamma process • Inspection and maintenance • Quality characteristics • Step-stress accelerated degradation tests

9.1 Introduction

These days, people worldwide have a good quality of life partly due to high quality of consumable/usable products that come from a highly competitive global market. At the same time, demand for digital multimedia products has grown rapidly, with the focus shifting to 3C (computing, communication, and consumer) and IA (information application) related products, and the parts in electrical cars. It takes a long time for the company to build up the reputation for their products, but just needs a short period of time to be branded as unreliable makers after shipping a flawed

N. Balakrishnan (✉)
Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada
e-mail: bala@mcmaster.ca

C.-C. Tsai • C.-T. Lin
Department of Mathematics, Tamkang University, Tamsui, Taiwan
e-mail: chihchuntsai@mail.tku.edu.tw; chien@mail.tku.edu.tw

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_9

product. It would certainly become a catastrophic problem if failures of products cause any injury and/or loss of life and consequent major lawsuits that arise as a result. For this reason, a careful assessment of the product reliability information, such as the mean time to failure (MTTF) or the p-th percentile of the lifetime within a reasonable life-testing time, becomes of great importance.

Traditional assessment of lifetimes of products uses conventional life tests that record only time-to-failure (complete or censored data) in order to quantify the life characteristics of products by assuming specific parametric lifetime distributions, such as Weibull and lognormal distributions. But in many situations, it is very difficult to collect such failure data under normal operating conditions due to very long lifetimes of modern day products. One way of obtaining additional information about the reliability of the products is to accelerate the lifetimes of units by testing them at higher levels of stress (e.g., elevated temperatures or voltages). This results in the so-called accelerated life tests (ALT). Some key references for theory and applications of ALT include Nelson [21], Bagdonavicius and Nikulin [2], Yang [46], Hsu and Fan [9, 10], and Balakrishnan and Ling [3].

Unfortunately, for some highly reliable products, the above methods provide little assistance as very few failures are likely to occur within the test duration. In such a situation, if there exists a quality characteristic (QC) whose degradation over time can be related to reliability, then the product's lifetime can be estimated well through the use of such degradation data. Generally, degradation models are constructed based on the formulations of random effects or Wiener process. Detailed discussions on these degradation models and their applications can be found in Singpurwalla [33], Chao [5], Meeker and Escobar [20], Doksum and Hóyland [7], Doksum and Normand [8], Whitmore and Schenkelberg [45], Yu and Tseng [51], Tseng et al. [39], Tang and Su [34], Peng and Tseng [29, 30], Si et al. [32], and Ye et al. [47, 48].

But, in many cases like in the case of crack growth of fatigue data shown in Lu and Meeker [18], the degradation paths of some specific products may display a monotone behavior which violate the standard assumptions made in Wiener processes. In order to properly model the degradation paths with monotone increasing patterns, we must rely on other stochastic processes. For the stochastic modelling of monotonic and gradual deterioration, the gamma process is quite appropriate.

In this chapter, we first review and address some degradation models based on gamma processes. As different ways of conducting the degradation experiments, degradation tests (DT), accelerated degradation tests (ADT), and step-stress accelerated degradation tests (SSADT) are then discussed. Moreover, the corresponding optimal designs are addressed in such a way that timely information of product's reliability can be provided efficiently to customers. Finally, some extensions and their applications of gamma degradation models are presented.

The rest of this chapter is organized as follows. In Sect. 9.2, we introduce the definition of gamma process and then present expressions of the distribution of the product's lifetime under the gamma degradation model as well as under a gamma process with random effects. In Sect. 9.3, we review some optimal designs and methods for estimating the parameters of the model based on a gamma process

under degradation tests, accelerated degradation tests, and step-stress accelerated degradation tests. In Sect. 9.4, we describe some extensions and their applications of gamma degradation models. Finally, some concluding remarks are made in Sect. 9.5.

9.2 Degradation Model Based on Gamma Process

We first present the definition of a gamma process and the expressions of the distribution of the product's lifetime under a gamma degradation model as well as under a gamma process with random effects.

9.2.1 Definition of Gamma Process

Let $L(t)$ denote the degradation path of the product, where $t \geq 0$, and $L(0) = 0$. Then, the degradation path of the product is said to follow a gamma process if the following properties hold:

(1) non-overlapping increments are independent, that is, the increments $\Delta L(t) = L(t + \Delta t) - L(t)$ are independent;
(2) for a given monotone increasing function $\eta(t)$ with $\eta(0) = 0$, and for fixed t and Δt,

$$\Delta L(t) \sim Ga(\beta, \Delta\eta(t)), \tag{9.1}$$

where $\Delta\eta(t) = \eta(t + \Delta t) - \eta(t)$ and β are the shape (function) and scale parameters of the gamma distribution, respectively, having its probability density function (pdf) as

$$f_{\Delta L(t)}(y) = \frac{y^{\Delta\eta(t)-1} e^{-y/\beta}}{\beta^{\Delta\eta(t)}\Gamma(\Delta\eta(t))}, \quad y > 0. \tag{9.2}$$

It is then evident that

$$\mathrm{E}(L(t)) = \eta(t)\beta$$

and

$$\mathrm{Var}(L(t)) = \eta(t)\beta^2. \tag{9.3}$$

9.2.2 Distribution of Product's Lifetime

The product's lifetime T can be suitably defined as the first passage time when $L(t)$ crosses a failure threshold level ω; that is,

$$T = \inf\{t|L(t) \geq \omega\}. \tag{9.4}$$

Because $L(t)$ is strictly increasing in t, the cumulative distribution function (cdf) of T in (9.4) can be expressed as

$$F_T(t) = P(T \leq t) = P(L(t) \geq \omega) = \frac{\Gamma(\eta(t), \omega/\beta)}{\Gamma(\eta(t))}, \tag{9.5}$$

where $\Gamma(a, z)$ is the upper incomplete gamma function defined by

$$\Gamma(a, z) = \int_z^\infty x^{a-1} e^{-x} dx.$$

When the mean degradation path is linear, that is, $\eta(t) = \alpha t$, Park and Padgett [26] obtained the exact pdf of the lifetime as

$$f_T(t) = \frac{dF_T(t)}{dt} = \alpha \left[\psi(\alpha t) - \log(\omega/\beta)\right] \left[1 - \frac{\Gamma(\alpha t, \omega/\beta)}{\Gamma(\alpha t)}\right] + \frac{\alpha}{\Gamma(\alpha t)} \frac{(\omega/\beta)^{\alpha t}}{(\alpha t)^2} \left(1 + \sum_{k=1}^{\infty} \left(\frac{\alpha t}{\alpha t + k}\right)^2 \frac{(-\omega/\beta)^k}{k!}\right),$$

where $\psi(z) = \frac{d}{dz} \log(\Gamma(z))$ is the digamma function.

However, the computation of this pdf is too complicated for practical examples. Therefore, Park and Padgett [26] proposed the following cdf and pdf of the Birnbaum-Saunders distribution to approximate the exact cdf and pdf of the lifetime of products:

$$F_T(t) = \Phi\left[\frac{1}{\alpha^*}\left(\sqrt{\frac{t}{\beta^*}} - \sqrt{\frac{\beta^*}{t}}\right)\right],$$

$$f_T(t) = \frac{1}{2\alpha^* t}\left(\sqrt{\frac{t}{\beta^*}} + \sqrt{\frac{\beta^*}{t}}\right) \phi\left[\frac{1}{\alpha^*}\left(\sqrt{\frac{t}{\beta^*}} - \sqrt{\frac{\beta^*}{t}}\right)\right], \quad t > 0,$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, and $\alpha^* = \sqrt{\beta/\omega}$ and $\beta^* = \omega/(\alpha\beta)$ are the shape and scale parameters of the Birnbaum-Saunders distribution.

In degradation modelling and analysis, there is substantial heterogeneity between the degradation paths of different individuals or units, which can not be duly

accounted for just by conditioning on explanatory variables. It is common in this case to incorporate unit-specific random effects or covariates to model such variability; for example, see Lu and Meeker [18] for growth curve models with random coefficients. Lawless and Crowder [14] constructed a tractable gamma process by incorporating random effects into the scale parameter β in (9.1). Specifically, they assumed that β^{-1} follows a gamma distribution $Ga(r^{-1}, \delta)$ with pdf

$$g(v) = \frac{v^{\delta-1} r^{\delta} e^{-rv}}{\Gamma(\delta)}, \quad v > 0. \tag{9.6}$$

In this setting, they showed that $\delta L(t)/(r\eta(t))$ has an F-distribution whose cdf is denoted by $F_{2\eta(t),2\delta}(x)$, and $L(t)/(r + L(t))$ has a beta distribution $Beta(\eta(t), \delta)$.

From these results, Tsai et al. [37] gave the expression of the cdf of the lifetime T as

$$F_T(t) = 1 - F_{2\eta(t),2\delta}\left(\frac{\delta\omega}{r\eta(t)}\right) = \frac{B(\frac{\omega}{\omega+r}; \eta(t), \delta)}{B(\eta(t), \delta)}, \tag{9.7}$$

where $B(x; a, b) = \int_x^1 z^{a-1}(1-z)^{b-1}dz$ is the upper incomplete beta function, and $B(a, b)$ is the complete beta function, i.e., $B(a, b) = B(0, a, b)$. Moreover, by assuming that $\eta(t) = \alpha t$, they derived an exact expression for the pdf of the product's lifetime as

$$f_T(t) = \alpha W\left(\frac{\omega}{\omega + r}, \alpha t, \delta\right),$$

where

$$W(x, a, b) = \frac{x^a}{aB(a, b)}\left[\left(\psi(a) - \psi(a + b) - \ln(x) + \frac{1}{a}\right){}_2F_1(\{a, 1 - b\}, \{1 + a\}; x)\right.$$
$$\left. - \frac{x(1 - b)}{(1 + a)^2}{}_3F_2(\{2 - b, 1 + a, 1 + a\}, \{2 + a, 2 + a\}; x)\right],$$

with ${}_gF_h$ being the confluent hypergeometric function defined by

$${}_gF_h(\{c_1, \cdots, c_g\}, \{d_1, \cdots, d_h\}; x) = \sum_{k=0}^{\infty} \frac{(c_1)_k \cdots (c_g)_k}{(d_1)_k \cdots (d_h)_k} \frac{x^k}{k!},$$

Pochhammer symbol $(c)_k = \Gamma(c + k)/\Gamma(c)$, and $(c)_0 = 0$.

9.3 Design and Inference of Degradation Experiment

We now review some optimal designs and methods for estimating the parameters of the model based on gamma process under degradation tests, accelerated degradation tests, and step-stress accelerated degradation tests.

9.3.1 Degradation Tests

Suppose n units are randomly selected for conducting a degradation test, and the measurements of each unit are made at times $t_0, t_1, \ldots, t_m$, with $t_0 = 0$. Let $L_i(t_j)$ denote the sample path of the i-th tested unit at time t_j, where $1 \le i \le n$ and $1 \le j \le m$. Let $\boldsymbol{\theta}$ be the unknown model parameter vector, and $Y_{ij} = L_i(t_j) - L_i(t_{j-1})$, $\Delta\eta(t_j) = \eta(t_j) - \eta(t_{j-1})$, for $1 \le i \le n$ and $1 \le j \le m$. Then, by the s-independent increment property of the gamma process, the likelihood function under gamma process, without random effect, can be expressed as

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n}\prod_{j=1}^{m} \frac{y_{ij}^{\Delta\eta(t_j)-1}}{\Gamma\left(\Delta\eta(t_j)\right)\beta^{\Delta\eta(t_j)}} \exp\left(-\frac{y_{ij}}{\beta}\right). \tag{9.8}$$

Similarly, from (9.6), the likelihood function for gamma process, with random effects, is given by

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}) &= \prod_{i=1}^{n}\left(\int_0^\infty \prod_{j=1}^{m} \frac{y_{ij}^{\Delta\eta(t_j)-1}\nu^{\Delta\eta(t_j)}}{\Gamma\left(\Delta\eta(t_j)\right)} \exp\left(-\nu y_{ij}\right) \frac{\nu^{\delta-1}r^{\delta}}{\Gamma(\delta)} \exp\left(-r\nu\right) d\nu\right) \\ &= \frac{r^{n\delta}\{\Gamma(\delta+\eta(t_m))\}^n \prod_{i=1}^{n}\prod_{j=1}^{m} y_{ij}^{\Delta\eta(t_j)-1}}{\left(\Gamma(\delta)\prod_{j=1}^{m}\Gamma(\Delta\eta(t_j))\right)^n \left(\prod_{i=1}^{n}(y_{im}+r)\right)^{\eta(t_m)+\delta}},\end{aligned} \tag{9.9}$$

where $y_{im} = L_i(t_m)$.

It is clear that optimization of these likelihood functions requires the knowledge of the parametric form of $\eta(\cdot)$, such as an exponential law or a power law. The work of van Noortwijk [41] provided an overview of parametric inferential methods for gamma process including methods of moments, maximum likelihood, and Bayesian inference. Recently, Lu et al. [19] made use of the Genz transform and a quasi-Monte Carlo method to maximize the likelihood function. They found that the performance of estimation in this case is much more effective than under the direct maximization method. Tsai et al. [35] used the approach of White [44] to investigate the effect on the estimated product's MTTF when the true model comes from a

gamma degradation process, but is wrongly assumed to be a Wiener degradation process. Their numerical results demonstrated that the effects on the accuracy and precision of the product's MTTF prediction for a laser data, taken from Meeker and Escobar [20], become critical when the shape and scale parameters of the gamma degradation process become large. Park and Kim [28] used the gamma process model for describing the nature of gradual and continuous performance degradation of light emitting diode (LED) lamps by assuming the shape function to be $\eta(t) = at^b$. Their experiments indicated that the service and warranty lifetime at the high temperature test conditions were estimated to be about a third of what they would be under normal temperature test conditions.

In many cases, the physical degradation process can be considered as monotone while the observed process is a perturbation of the degradation process which yields the monotone phenomenon. In this setting, Bordes et al. [4] considered a perturbed gamma degradation process $L(t)$ as

$$L(t) = Y(t) + \tau B(t), \tag{9.10}$$

where $Y(t)$ follows a homogeneous gamma process as given in (9.1) and $B(t)$ is a standard Brownian motion which is assumed to be independent of the gamma process. When $\tau = 0$, this model simply reduces to a gamma degradation process. Moreover, when $\mathrm{E}(Y(t))$ tends to a positive constant and $\mathrm{Var}(Y(t))$ tends to 0, this model converges weakly to a Brownian motion with positive drift. For this reason, the role of Brownian motion in model (9.10) can be interpreted as measurement error. Based on independent copies of the perturbed gamma process observed at irregular instants, Bordes et al. [4] then used the method of moments for estimating the unknown parameters of the model and discussed the asymptotic properties of these estimators.

Usually, as pointed out by Pulcini [31], the measurement errors are assumed to be statistically independent of the hidden process when the process is modeled by a gamma process. However, in some situations, it is plausible that the measurement errors are statistically dependent on the actual degradation (or usage) level of the hidden process. For this reason, Pulcini [31] proposed a new perturbed gamma process $L(t)$, which includes the gamma process $Y(t)$ together with a statistically state-dependent measurement error $\varepsilon(t)$, in the following form:

$$L(t) = Y(t) + \varepsilon(t), \tag{9.11}$$

where the error $\varepsilon(t)$ is assumed to be normally distributed with mean zero and standard deviation that depends on the actual level, $\sigma(Y(t))$. Based on the trend of the empirical estimate of the variance-to-mean ratio of the perturbed process, a graphical check is then used to recognize both the nature of the measurement errors and a suitable functional form for the standard deviation of its distribution. In addition, an estimation procedure, based on the maximum likelihood method and Monte Carlo integration, is suggested. When the inspections are invasive or destructive and the error standard deviation depends on the actual level in a linear

way, an approximate maximum likelihood estimation procedure that avoids the use of numerical or Monte Carlo integration has also been proposed.

If suitable prior information is not available about the parametric form of the shape function, a semiparametric estimation is preferred. Because $\eta(\cdot)$ is estimated nonparametrically, direct optimization of the likelihood function becomes extremely difficult as there are too many parameters to estimate. When the observation times differ from unit to unit, the number of parameters would become even larger. Wang [42] discussed semiparametric inference for the model of Lawless and Crowder [14]. A pseudo-likelihood method, which ignores the dependence between the degradation measurements at successive observation times, is proposed to estimate the unknown parameters. Although the maximum pseudo-likelihood estimator is consistent, there is a substantial loss in efficiency as compared to the full likelihood method. Ye et al. [50] showed that the maximum likelihood estimates (MLE) of the parameters can be obtained through the expectation-maximization (EM) algorithm by converting these two estimation problems to missing data problems, and then made use of the nonparametric percentile bootstrap method to construct confidence intervals for $\boldsymbol{\theta}$.

9.3.2 Accelerated Degradation Tests

For some highly reliable products, degradation path may degrade very slowly and it is therefore impossible to obtain a good estimate within a reasonable test duration. To overcome this problem, the reliability information at the normal use condition can be extrapolated by collecting degradation data under higher stress levels such as elevated temperatures or voltages. This is called an accelerated degradation test (ADT).

Suppose there are $N = \sum_{l=1}^{k} n_l$ test units available for conducting an ADT with k stress levels, $S_0 \leq S_1 \leq \cdots \leq S_k$, where S_0 is the use condition, and n_l items are randomly assigned for a degradation test at stress level S_l, for $1 \leq l \leq k$. For stress level S_l, the inspections are made m_l times and the degradation measurements of each unit are available at times $t_1, \cdots, t_{m_l}$. Then, the product's lifetime in this test can be suitably defined as the first passage time when the degradation path under use condition $L(t|S_0)$ crosses the threshold level ω. Hence, the product's lifetime T, under use condition S_0, can be expressed as

$$T = \inf\{t|L(t|S_0) \geq \omega\}. \tag{9.12}$$

For $1 \leq i \leq n_l$, $1 \leq j \leq m_l$, and $1 \leq l \leq k$, let $L_i(t_j|S_l)$ denote the sample degradation path of ith test unit at time t_j under the stress level S_l, and assume that the degradation path $L(t|S_l)$ under stress level S_l follows a gamma process with shape function $\eta(t; S_l) = \alpha_l t$ and scale parameter β, where the parameters α_l and stress level S_l are assumed to have a relationship given by

$$\ln(\alpha_l) = \gamma_0 + \gamma_1 A(S_l). \tag{9.13}$$

Two commonly used expressions for $A(S_l)$ are as follows:

$$A(S_l) = \begin{cases} 1/(273.15 + S_l), & \text{Arrhenius model,} \\ \ln S_l, & \text{Inverse-power model.} \end{cases} \tag{9.14}$$

Park and Padgett [26] studied the accelerated degradation models for a single accelerating variable to a general class of models that included Gaussian process models, geometric Brownian motion models, and gamma process models as special cases. In their class of models, both failures and degradation measurements are incorporated for inferential purposes. It should be noted that some popular acceleration functions such as power rule model, Arrhenius reaction rate model, inverse-log model, exponential model, and inverse-linear model can handle only one accelerating variable. Other specific link functions have been developed for two accelerating variables, such as the generalized Eyring model (GEM), but these are of somewhat limited use for specific physical situations (Park and Padgett [27]). Hence, by assuming that there are h accelerating variables $\ell_1, \ell_2, \ldots, \ell_h$, Park and Padgett [27] proposed the following generalized acceleration function

$$\kappa(\ell_1, \ell_2, \ldots, \ell_h) = \theta_0 \prod_{i=1}^{h} \{\zeta(\ell_i)\}^{\theta_i},$$

where $\zeta(\cdot)$ is any monotone function, and $\theta_0, \ldots, \theta_h$ are unknown model parameters, to investigate stochastic degradation models. They also provided the exact likelihood functions for the degradation paths, and discussed maximum likelihood estimation of the model parameters. The MLEs are shown to outperform the estimates of Park and Padgett [26] described earlier.

Under the assumption that the relationship between the parameter α_l and the stress level S_l is an Arrhenius model, Guan and Tang [11] discussed the optimal constant-stress ADT (CSADT) plans for the test stress levels and the proportion of units allocated to each stress level based on D-optimality and V-optimality criteria. More precisely, they sought to find a test plan $\boldsymbol{\xi} = (x_1, x_2, \ldots, x_k, \pi_1, \pi_2, \ldots, \pi_k)$ which consists of k standardized test stress levels $x_1, x_2, \ldots, x_k$, where

$$x_l = \left(\frac{1}{273.15 + S_0} - \frac{1}{273.15 + S_i}\right) \Big/ \left(\frac{1}{273.15 + S_0} - \frac{1}{273.15 + S_k}\right), \tag{9.15}$$

such that the experimental region of $x_l, l = 1, \ldots, k$, is in the range $[0, 1]$, and the corresponding proportions of test units $\pi_1 = n_1/N, \pi_2 = n_2/N, \ldots, \pi_k = n_k/N$ satisfying $\sum_{l=1}^{k} \pi_l = 1$, maximize the determinant of the Fisher information matrix $I(\boldsymbol{\xi})$ (the negative of the expected value of the second derivative of the total log-likelihood function), and minimize the approximate variance of the estimated MTTF at the use stress S_0, $\text{AVar}(\widehat{\text{MTTF}}_0|\boldsymbol{\xi})$. Applying the general equivalence

theorem (Whittle [43]), they verified that these two optimized test plans are globally optimum. In addition, they used a simple grid search method to study ADT compromise test plans with three stress levels for practical applications.

Lim [16] developed the optimal CSADT plan for a single accelerating variable with Arrhenius model, power model, or exponential model when $k = 2$. Under the constraint that the total experimental cost does not exceed a pre-fixed budget, the decision variables such as the number of measurements, measurement times, test stress levels and number of units allocated to each stress level are optimally determined by minimizing the approximate variance of the MLE of the p-th percentile of the lifetime distribution at the use condition. In other words, by assuming that $m_l = m \quad \forall l$, and $t_j - t_{j-1} = \Delta t \quad \forall j$, the optimal test plan $\boldsymbol{\xi}^* = (x_1^*, x_2^*, n_1^*, n_2^*, \Delta t^*)$ is obtained by

$$\min_{\boldsymbol{\xi}} \text{AVar}(\hat{t}_p|\boldsymbol{\xi}), \tag{9.16}$$

subject to

$$\text{TC}(\boldsymbol{\xi}) = C_{op} m \Delta t + C_{mea}\, Nm + C_{it}\, N \le C_b, \tag{9.17}$$

where $\text{TC}(\boldsymbol{\xi})$ is the total cost of conducting the ADT experiment, $\text{AVar}(\hat{t}_p|\boldsymbol{\xi})$ denotes the approximate variance of the estimated p-th percentile of the lifetime, C_b denotes the total budget for conducting the experiment, and C_{op}, C_{mea} and C_{it} denote the unit cost of operation, measurement and item, respectively. From the simplicity in the constraint structure and the integer restriction on the decision variables except the stress levels, Lim [16] presented a computational algorithm for determining the optimal solution. In addition, he developed a compromise plan with three stress levels for checking the adequacy of the assumed acceleration function.

Tsai et al. [38] investigated the optimal CSADT strategy with the loadings of two stress variables (temperature and current) and the GEM in the gamma process. Let k be the total number of combinations of two stress variables. They assumed that the relationship between α_l and standardized lth stress-loading level (x_{1l}, x_{2l}) can be expressed as

$$\alpha_l = \exp(\gamma_0 + \gamma_1 x_{1l} + \gamma_2 x_{2l} + \gamma_3 x_{1l} x_{2l}),$$

for $l = 1, \ldots, k$. Their optimal ADT procedure was established to minimize the asymptotic variance of the MLE of the MTTF of a product subject to a given budget. That is, the optimal test plan $\boldsymbol{\xi}^* = (n_1^*, n_2^*, \ldots, n_k^*, t_{m1}^*, t_{m2}^*, \ldots, t_{mk}^*)$ can be obtained as

$$\min_{\boldsymbol{\xi}} \text{AVar}(\widehat{\text{MTTF}}_0|\boldsymbol{\xi}), \tag{9.18}$$

subject to

$$\mathrm{TC}(\boldsymbol{\xi}) \leq C_b, \delta \leq t_{m1}, t_{m2}, \ldots, t_{mk} \leq t_U, \sum_{l=1}^{k} n_l \leq n_u,$$

where t_U is the upper bound of the measurement time to implement an ADT, δ is measurement frequency in all runs, n_u is an upper bound on the total sample size, $\mathrm{TC}(\boldsymbol{\xi})$ is the total operating cost which includes the fixed cost, operating cost, and variable cost, and C_b is as defined in (9.17). To achieve the optimal ADT plan on the sample size and termination time for each run in the ADT at a constant measurement frequency, they proposed a grid search-type algorithm by splitting the space of termination times into a set of discrete points such that the optimal ADT plan can be found over a domain with a finite number of combinations from candidate parameters to prevent the computational algorithm from diverging. Using Monte Carlo simulations, they studied the sensitivity of the MLE of model parameters to the sample size for different numbers of measurement times. Similar simulation results can also be found in Chiang et al. [6].

Ling et al. [17] discussed accelerated degradation analysis that characterizes the health and quality of systems with monotonic and bounded degradation. In particular, they studied the degradation of light intensity of LED lamps based on a gamma process, time-scale transformation, and a power link function for associating covariates. They developed inferences for different lifetime characteristics such as the reliability, the mean and median lifetimes, the conditional reliability, and the remaining useful life of systems under normal use conditions, and presented approximate confidence intervals for these quantities of interest by using the observed Fisher information matrix. Their methods can be used for some other applications such as maintenance of roads, railways, bridges, buildings, and industrial plants.

9.3.3 Step-Stress Accelerated Degradation Tests

Although ADT is an efficient life-test method, it may not be possible to have enough test units for conducting such a life test for a newly developed product, or an expensive product. In this case, a CSADT will no longer be useful for assessing the lifetime distribution of highly reliable products at use condition. Hence, a step-stress ADT (SSADT) is suggested to handle this problem, especially when the available test items are few.

In a SSADT experiment, an item is first tested, subject to a pre-fixed stress level for a specified length of time. If it does not fail, it is tested again at a higher stress level for another fixed length of time. The stress on the specimen is thus increased step by step until an appropriate termination time is reached. Specifically, the experiment can be described as follows. Let S_0 be the normal stress level and $S_1, \ldots, S_m$ denote m higher stress levels such that $S_0 \leq S_1 \leq \cdots \leq S_m$. Suppose there are n test units subject to a degradation test (with a measurement frequency per f units time) under stress S_1, and the duration time of degradation test under the

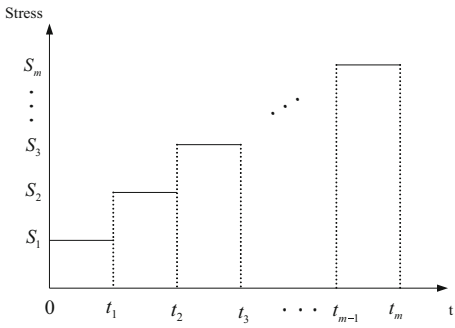

Fig. 9.1 The relationship between stress and time

stress S_1 is t_1; next, we increase the stress level to S_2, and the duration time under S_2 is up to t_2, and so on until the stress is up to S_m, and the experiment is terminated at time t_m. Thus, the testing stress, S, of an SSADT experiment can be expressed as

$$S = \begin{cases} S_1 & \text{if } 0 \le t < t_1, \\ \vdots & \vdots \\ S_m & \text{if } t_{m-1} \le t < t_m. \end{cases}$$

Figure 9.1 shows the relationship between stress and time.

Tseng et al. [40] first dealt with the optimal SSADT plan (including the optimal settings for the sample size, measurement frequency, and termination time) for a gamma degradation process. In their work, they assumed that the increments under stress S_i $\forall i = 1, \cdots, m$, $\Delta L(t|S_i) \sim Ga(\beta, \alpha_i \Delta t)$, and that there exists a relationship between α_i and stress S_i as given in (9.13).

Let $L_{ss}(t)$ be the degradation path of a SSADT with a gamma degradation model. Then, under the stress S_1, and for $t \in [0, t_1)$, we have

$$L_{ss}(t) = L(t|S_1) \sim Ga(\beta, \alpha_1 t), \quad 0 \le t < t_1.$$

When we increase the stress up to S_2 at time t_1, by the additive property with respect to shape parameter of the gamma distribution, the degradation path for $t \in [t_1, t_2)$ becomes

$$L_{ss}(t) = L(t_1|S_1) + L\left((t - t_1)|S_2\right) \sim Ga\left(\beta, \alpha_2(t - t_1) + \alpha_1 t_1\right).$$

Similarly, for general $j \ge 1$, $t_{j-1} < t < t_j$, and $t_0 = 0$, we obtain

$$\begin{aligned} L_{ss}(t) &= L(t_1|S_1) + L\left((t_2 - t_1)|S_2\right) + \cdots + L\left((t - t_{j-1})|S_j\right) \\ &\sim Ga\left(\beta, \alpha_j(t - t_{j-1}) + \sum_{i=1}^{j-1} \alpha_i(t_i - t_{i-1})\right). \end{aligned} \tag{9.19}$$

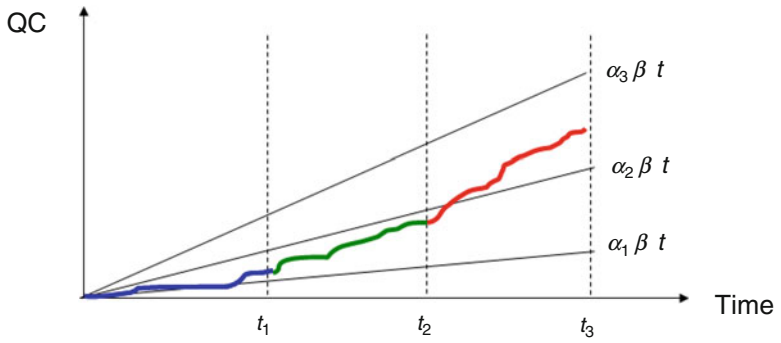

Fig. 9.2 Illustration of a degradation path of $L_{ss}(t)$ with $m = 3$

Figure 9.2 gives an illustration of a degradation path of $L_{ss}(t)$ with $m = 3$, when the mean degradation paths under each stress follow a gamma process with linear pattern.

For $1 \le i \le m$, let l_i denote the total number of measurements under stress S_i. Then, $t_i = f \times (\sum_{k=1}^{i} l_k)$. Tseng et al. [40] determined the optimal setting of $\boldsymbol{\xi}^* = (n^*, f^*, \{l_i^*\}_{i=1}^m)$ by solving the following optimization problem:

$$\min_{\boldsymbol{\xi}} \text{AVar}(\widehat{\text{MTTF}}_0 | \boldsymbol{\xi}), \tag{9.20}$$

subject to

$$\text{TC}(\boldsymbol{\xi}) = C_{op} f \sum_{i=1}^{m} l_i + C_{mea}\, n \sum_{i=1}^{m} l_i + C_{it}\, n \le C_b, \qquad n, f, l_1, \cdots, l_m \in \mathbb{N},$$

where $\text{AVar}(\widehat{\text{MTTF}}_0)$ denotes the approximate variance of $\widehat{\text{MTTF}}_0$, $\text{TC}(\boldsymbol{\xi})$ denote the total cost of conducting a SSADT experiment, and C_{op}, C_{mea}, C_{it}, and C_b are as defined in (9.17).

Pan and Balakrishnan [22] indicated that the number of products is relatively few in the SSADT experiment, and so the test control to the products should be done more carefully. Thus, by considering the differences in products, they suggested that multiple-steps step-stress accelerated degradation models based on Wiener and gamma processes, respectively, in which the times of stress level elevating are random and vary from product to product., i.e., the ith product's testing stress can be expressed as

$$S = \begin{cases} S_1 & \text{if } 0 \le t < t_{i,1}, \\ \vdots & \vdots \\ S_m & \text{if } t_{i,m-1} \le t < t_{i,m}, \end{cases}$$

for $i = 1, \ldots, n$. Note that, given pre-specified degradation values $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_{m-1})$, the stress will be changed to S_{l+1} when the degradation value of a product reaches ω_l, and the corresponding time is $t_{i,l}$. It can be seen that

$$\triangle t_{i,l} = t_{i,l} - t_{i,l-1} = \inf\{t | L(t|S_l) \geq \omega_l - \omega_{l-1}\}, \ \omega_0 = 0.$$

For a gamma process, the exact pdf of $\triangle t_{i,l}$ is quite complicated. So, by using the approach of Park and Padgett [26], they obtained the likelihood function approximately. Because this likelihood function is analytically intractable and is therefore difficult to use numerical search algorithms for determining the MLE directly, they applied the Bayesian Markov chain Monte Carlo (MCMC) method to estimate the parameters of the model efficiently.

Amini et al. [1] extended the model of Pan and Balakrishnan [22] to a more economic plan. They assumed that the times of stress level elevating in the new model are identical for all products, and that the new method does not require continuous inspection of the degradation by the experimenter or electronic sensors. The advantage of such a plan is that it uses only one chamber (oven) for testing all test units and consequently decreases the cost of measurement. In their SSADT model, they assumed that M measurements are made for each unit with a measurement frequency per f unit of time. Also, under the lth stress level S_l, $l = 1, \ldots, m$, the stress level is elevated as soon as a new measurement of the degradation of at least one of the test units exceeds the threshold value ω_l. Hence, the testing stress of the ith unit, $i = 1, \ldots, n$, can be expressed as follows:

$$S = \begin{cases} S_1 & \text{if } 0 \leq \kappa_1 f, \\ \vdots & \vdots \\ S_m & \text{if } \kappa_{m-1} f \leq t < Mf, \end{cases}$$

where

$$\kappa_j = \min\left\{M, \left\lfloor \frac{t_{(1),j}}{f} \right\rfloor\right\}, \ j = 1, \ldots, m-1,$$

with the notation $\lfloor x \rfloor$ standing for the ceiling of the real number x, and

$$t_{(1),j} = \min\{t_{i,j}; i = 1, \ldots, n\}, \ j = 1, \ldots, m-1.$$

Then, under the constraint that the total experimental cost does not exceed a pre-fixed budget, the optimal setting of $\boldsymbol{\xi}^* = (n^*, f^*, M^*, \omega_1^*, \ldots, \omega_{m-1}^*)$ can be obtained by solving the following optimization problem:

$$\min_{\boldsymbol{\xi}} \text{AVar}(\hat{t}_p | \boldsymbol{\xi}), \tag{9.21}$$

subject to

$$\mathrm{TC}(\boldsymbol{\xi}) = C_{op}\, fM + C_{mea}\, nM + C_{it}\, n \le C_b, \qquad n, f, M \in \mathbb{N},$$

where $\mathrm{AVar}(\hat{t}_p)$ denotes the approximate variance of the estimated pth percentile of the lifetime under use condition, and C_{op}, C_{mea}, C_{it}, and C_b are as defined in (9.17).

Li et al. [15] proposed the Bayesian optimal design of SSADT based on a gamma process and relative entropy. Under the framework of Bayesian theory, relative entropy can be used to measure the information gained between the prior and the posterior distribution. The goal is to maximize the expectation of relative entropy between the prior and the posterior distribution. That is, for fixed m, n and f, the optimal stress levels and inspection number scheme $\boldsymbol{\xi}^* = (\{S_i^*, l_i^*\}_{i=1}^m)$ can be determined by solving the following optimization problem:

$$\min_{\boldsymbol{\xi}} \mathrm{E}_x(I_1 - I_0), \tag{9.22}$$

subject to

$$S_0 < S_1 < S_2 < \ldots < S_m \le S_u, l_1 \ge l_2 \ge \ldots \ge l_m,$$

where I_0 is the information contained in the prior distribution, I_1 is the total amount of information obtained from the posterior distribution, E_x is the mathematical expectation of the sample space information (the increment of degradation paths), and S_u is the upper bound of stress levels. Compared with the asymptotic theory or the D optimization method, they found that the relative entropy could comprehensively utilize sample information and prior information from the products under ADT, which can make the results more accurate.

9.4 Extensions and Some Applications

Other than the models discussed above, there are many other models available for degradation modelling in the literature. In this section, we review the models with more than one quality characteristic, and also some applications of gamma processes in inspection models, maintenance decisions, and burn-in test.

9.4.1 Multiple Quality Characteristics

Many highly reliable products have multiple components with more than one quality characteristic that are dependent on each other due to their complex structure. In such situations, a bivariate or multivariate degradation model is needed to estimate the reliability of products. This analysis is needed not only for design and technical purposes, but also for making managerial decisions.

Pan and Balakrishnan [23] introduced the degradation model with two dependent quality characteristics by assuming that their degradation can be modeled by gamma processes. They used a bivariate Birnbaum-Saunders distribution and its marginal distributions to approximate the reliability of the product, and applied the Bayesian MCMC method for developing the corresponding inference.

An extension of this work can be seen in Pan et al. [24, 25]. Pan et al. [25] discussed the lifetime distribution and associated inferential method of systems with multiple degradation measurements by assuming that all the degradation paths of the components are governed by gamma processes as shown in (9.1) with $\eta(t) = \alpha t$. They used a multivariate Birnbaum-Saunders distribution and its marginal distributions to approximate the reliability of the system. Later, by following the approach of Pan and Balakrishnan [23], Pan et al. [24] studied the reliability model for parallel and series systems with two dependent degrading components. They specifically used a bivariate Birnbaum-Saunders distribution to approximate the reliability of a parallel system, and used a bivariate Birnbaum-Saunders distribution and its marginal distributions to approximate the reliability of a series system.

9.4.2 Inspection Model and Maintenance Decision

The gamma process is largely used in deterioration modelling for determining optimal inspection and maintenance of the products. Kallen and van Noortwijk [12] gave a short review of inspection models based on the gamma process for the period 1987 to 2002. However, none of these works deal with imperfect inspections. So, they used gamma process to model the uncertain reduction of wall thickness due to corrosion and proposed an adaptive Bayesian decision model to determine these optimal inspection plans by minimizing the expected average cost per year. van Noortwijk [41] gave an overview of the most important statistical and mathematical properties as well as methods for estimation (such as maximum likelihood, method of moments, Bayesian updating, and expert judgement), approximation, and simulation of gamma processes. He also surveyed models for time-based and condition-based maintenance under gamma-process deterioration. Here, time-based preventive maintenance is carried out at regular intervals of time, whereas condition-based maintenance is carried out at times determined by inspecting or monitoring a structure's condition. Khanh et al. [13] constructed a deterioration model using a non-homogeneous gamma process with Gaussian noise as shown in (9.10). They considered a two-stage procedure to estimate the remaining lifetime distribution. In this procedure, they used the Gibbs sampling technique to approximate the hidden degradation states in the first stage, and applied the stochastic expectation-maximization (EM) algorithm to estimate the model parameters in the second stage. Then, based on the results of the remaining lifetime estimation, they proposed a maintenance decision rule to prevent system's failure.

9.4.3 Burn-In Test

Burn-in test is a manufacturing process applied to products to eliminate latent failures or weak components in the factory before the products get delivered to customers. For highly reliable products with monotone degradation, the optimal burn-in policies can be determined through a gamma process. Often, optimal burn-in policies are determined by employing one of four criteria: (i) maximization of the mean residual lifetime (MRL) of the product, (ii) achievement of a prescribed mission reliability, (iii) minimization of cost, or (iv) optimization of an objective function subject to some constraints.

Motivated by laser data, Tsai et al. [36] proposed a mixed gamma process to describe the degradation path of the product, which can be expressed as follows:

$$L(t) \sim \begin{cases} Ga(\beta_1, \eta_1(t)) \text{ for weak group,} \\ Ga(\beta_2, \eta_2(t)) \text{ for typical group,} \end{cases} \tag{9.23}$$

where $\eta_1(t)\beta_1 > \eta_2(t)\beta_2 > 0$, $\eta_1(t) \neq \eta_2(t) > 0$, and $\beta_1 \neq \beta_2 > 0$, $\forall t > 0$. Note that the condition $\eta_1(t)\beta_1 > \eta_2(t)\beta_2$ stipulates that an item in the weak group will degrade faster on an average than an item in the typical group at time t. Hence, it can be seen that an unit is classified as a typical unit if $L(t)$ is relatively small. Let $\zeta(t)$ denote a cutoff point for $L(t)$ between the typical and weak units. Then, they presented the following decision rule for classifying an unit as typical or weak:

$$R: \text{An item is classified into the typical group at time } t \text{ iff } L(t) \leq \zeta(t). \tag{9.24}$$

They then used a cost model to determine the optimal termination time of a burn-in test. Let n denote the total number of units subject to a burn-in test, and p denote the proportion of weak units. Then, the total misclassification cost can be expressed as

$$\mathrm{MC}(\zeta(t), t) = C_\varphi n(1-p)\varphi(t) + C_\vartheta np\vartheta(t),$$

where C_φ (C_ϑ) denotes the unit cost of misclassifying a typical (weak) item as a weak (typical) item, and $\varphi(t)$ ($\vartheta(t)$) is the probability of type-I (type-II) error, of misclassifying a typical (weak) item as a weak (typical) item. For a fixed t, they determined the optimal cutoff point $\zeta^*(t)$ by minimizing the total misclassification cost $\mathrm{MC}(\zeta(t), t)$; that is,

$$\zeta^*(t) = \arg\min_{\xi(t)} \mathrm{MC}(\xi(t), t), \tag{9.25}$$

In addition to the misclassification cost, they also considered test costs that include the cost of conducting the degradation test, and of measuring the data. Suppose $t = 0, t_1, \cdots, t_l$ are the check points of a burn-in test; then, the total number of data collection points at t_b is $b + 1$ for $1 \leq b \leq l$. Hence, the total cost of

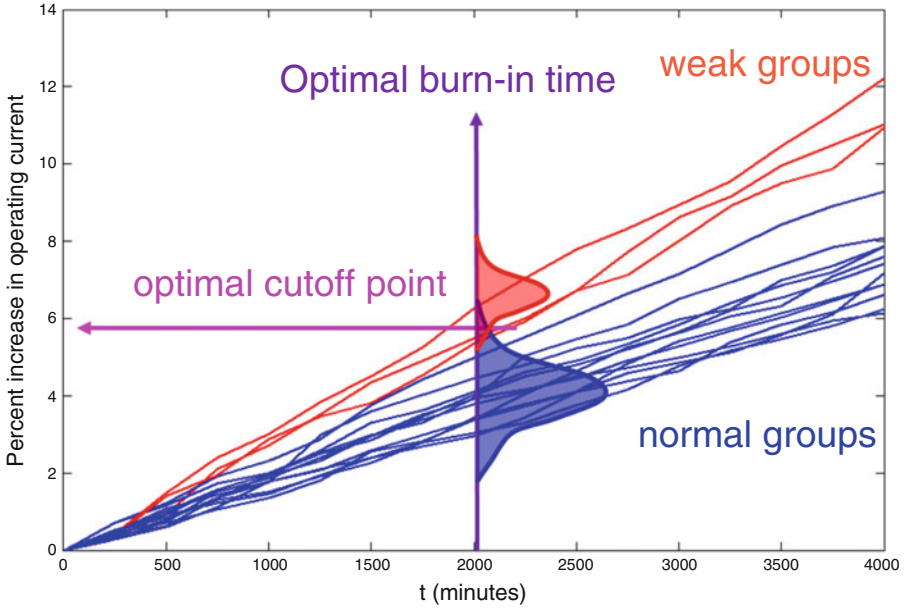

Fig. 9.3 Illustration of the procedure on burn-in test

conducting a burn-in test up to time t_b can be expressed as

$$TC(\zeta^*(t_b), t_b) = \mathrm{MC}(\zeta(t_b), t_b) + C_{op} n t_b + C_{mea} n(b+1),$$

where C_{op} and C_{mea} are as defined in (9.17). For a fixed burn-in time t_b, they determined the optimal burn-in time by substituting $\zeta^*(t_b)$ in (9.25) into $\mathrm{TC}(\zeta(t_b), t_b)$, and further determined the optimal burn-in time t_b^* by minimizing $\mathrm{TC}(\zeta^*(t_b), t_b)$, i.e.,

$$t_b^* = \arg \min_{\{t_b\}_{b=1}^{l}} \mathrm{TC}(\zeta^*(t_b), t_b).$$

An illustration of the two-step procedure on burn-in test is shown in Fig. 9.3.

Many products are prone to multiple failure modes. According to the failure mechanism, a failure mode can be either a degradation-threshold failure or a catastrophic failure. A degradation-threshold failure occurs when a measurable physical degradation reaches a critical threshold level, which is often specified by industrial standards, while a catastrophic failure causes instant product failure. Both kinds of failure modes may be subject to infant mortality. To identify and eliminate units with infant mortality, engineers often resort to burn-in by activating all infant mortality failure modes during the test for a certain duration; see, for example, Ye et al. [49].

Motivated by two real-life examples, Ye et al. [49] considered a general burn-in planning framework for products with independent competing risks. This framework is able to differentiate between normal and infant mortality failure modes and recommends degradation-based burn-in approaches. Particularly, they assumed that the degradation-threshold failure mode is subject to an infant mortality, and the degradation path $L(t)$ is a gamma process with random effect as shown in Sect. 2.2.3. In addition to this failure mode, the product is also prone to a catastrophic failure, which is a normal failure mode. Their decision rule during the burn-in test is as follows: After burn-in with a duration of b, the QC of tested unit is measured nondestructively. If the QC exceeds a predetermined cutoff level ζ_b, the unit is discarded. The per-unit burn-in cost includes setup cost C_s and the burn-in operational cost C_o per time unit of burn-in. If an unit has failed at the end of the burn-in period (e.g., due to the catastrophic failure), it is discarded with a cost of C_{dis}. Otherwise, its degradation is measured with a measurement cost of C_{mea}. If the degradation exceeds the cutoff level ζ_b, the unit is rejected with a disposal cost of C_d. An accepted unit will be put into field operation. If it fails within the mission time t_U, some handling and administrative cost of C_f is incurred. Otherwise, a gain of K is generated.

Based on this framework, they developed three degradation-based burn-in models (burn-in model for products with only a single (degradation-threshold) failure mode, a normal failure mode but only the degradation-threshold mode is activated during burn-in, and a catastrophic failure mode that is normal but has to be activated during burn-in), and derived the optimal cutoff degradation level ζ_b^* and burn-in time b^* by minimizing the associated mean total cost for each model.

9.5 Concluding Remarks

Since the introduction of gamma process in the area of reliability in 1975, it has been used extensively to model the degradation process. This article surveys the theoretical aspects as well as the application of gamma processes in degradation analysis. Some of the inferences and optimal designs for conducting the degradation experiments efficiently under different tests (DT, ADT, and SSADT) are reviewed. By aggregating the information of merits of each model, this review paper provides the key notes about circumstances and conditions for choosing optimal test plans. The developments on models under gamma-process deterioration with more than one quality characteristic, and the application of gamma processes in inspection models, maintenance decisions, and burn-in test with multi-component and multi-failure-mode models including their statistical dependencies, have also been highlighted.

References

1. Amini M, Shemehsavar S, Pan Z (2016) Optimal design for step-stress accelerated test with random discrete stress elevating times based on gamma degradation process. Qual Reliab Eng Int 32:2391–2402
2. Bagdonavicius V, Nikulin M (2002) Accelerated life models, modeling and statistical analysis. Chapman & Hall, Boca Raton
3. Balakrishnan N, Ling MH (2014) Best constant-stress accelerated life-test plans with multiple stress factors for one-shot device testing under a weibull distribution. IEEE Trans Reliab 63:944–952
4. Bordes L, Paroissin C, Salami A (2016) Parametric inference in a perturbed gamma degradation process. Commun Stat Theory Methods 45:2730–2747
5. Chao MT (1999) Degradation analysis and related topics: some thoughts and a review. Proc Nat Sci Counc Ser A23:555–566
6. Chiang JY, Sung WY, Tsai TR, Lio Y (2015) Sensitivity analysis of sample allocation and measurement frequency under a degradation test with gamma process. ICIC Express Lett Part B Appl 6:737–742
7. Doksum KA, Hóyland A (1992) Model for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. Technometrics 34:74–82
8. Doksum KA, Normand S-LT (1995) Gaussian models for degradation processes – part I: methods for the analysis of biomarker data. Lifetime Data Anal 1:135–144
9. Fan TH, Hsu TM (2012) Accelerated life tests of a series system with masked interval data under exponential lifetime distributions. IEEE Trans Reliab 61:798–808
10. Fan TH, Hsu TM (2014) Constant stress accelerated life test on a multiple-component series system under Weibull lifetime distributions. Commun Stat Theory Methods 43:2370–2383
11. Guan Q, Tang YC (2013) Optimal design of accelerated degradation test based on gamma process models. Chin J Appl Probab Stat 29:213–224
12. Kallen MJ, van Noortwijk JM (2005) Optimal maintenance decisions under imperfect inspection. Reliab Eng Syst Saf 90:177–185
13. Khanh LS, Mitra F, Anne B (2016) Remaining useful lifetime estimation and noisy gamma deterioration process. Reliab Eng Syst Saf 149:76–87
14. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
15. Li X, Zou T, Fan Y (2015) Bayesian optimal design for step-stress accelerated degradation testing based on gamma process and relative entropy. Lect Notes Mech Eng 19:957–967
16. Lim H (2015) Optimum accelerated degradation tests for the gamma degradation process case under the constraint of total cost. Entropy 17:2556–2572
17. Ling MH, Tsui KL, Balakrishnan N (2015) Accelerated degradation analysis for the quality of a system based on the gamma process. IEEE Trans Reliab 64:463–472
18. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35:161–174
19. Lu DJ, Pandey MD, Xie WC (2013) An efficient method for the estimation of parameters of stochastic gamma process from noisy degradation measurements. J Risk Reliab 227:425–433
20. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. John Wiley & Sons, New York
21. Nelson W (1990) Accelerated testing: statistical models, test plans, and data analysis. John Wiley & Sons, New York
22. Pan Z, Balakrishnan N (2010) Multiple-steps step-stress accelerated degradation modeling based on Wiener and gamma processes. Commun Stat Simul Comput 39:1384–1402
23. Pan Z, Balakrishnan N (2011) Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. Reliab Eng Syst Saf 96:949–957
24. Pan Z, Feng J, Sun Q (2016) Lifetime distribution and associated inference of systems with multiple degradation measurements based on gamma processes. Eksploatacja i Niezawodnosc 18:307–313

25. Pan Z, Sun Q, Feng J (2016) Reliability modeling of systems with two dependent degrading components based on gamma processes. Commun Stat Theory Methods 45:1923–1938
26. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma process. Lifetime Data Anal 11:511–527
27. Park C, Padgett WJ (2006) Stochastic degradation models with several accelerating variables. IEEE Trans Reliab 55:379–390
28. Park SH, Kim JH (2016) Lifetime estimation of LED lamp using gamma process model. Microelectron Reliab 57:71–78
29. Peng CY, Tseng ST (2009) Mis-specification analysis of linear degradation models. IEEE Trans Reliab 58:444–455
30. Peng CY, Tseng ST (2013) Statistical lifetime inference with skew-Wiener linear degradation models. IEEE Trans Reliab 62:338–350
31. Pulcini G (2016) A perturbed gamma process with statistically dependent measurement errors. Reliab Eng Syst Saf 152:296–306
32. Si XS, Wang W, Hu CH, Zhou DH, Pecht MG (2012) Remaining useful life estimation based on a nonlinear diffusion degradation process. IEEE Trans Reliab 61:50–67
33. Singpurwalla ND (1995) Survival in dynamic environments. Stat Sci 10:86–103
34. Tang J, Su TS (2008) Estimating failure time distribution and its parameters based on intermediate data from a Wiener degradation model. Nav Res Logist 55:265–276
35. Tsai CC, Tseng ST, Balakrishnan N (2011) Mis-specification analyses of gamma and Wiener degradation processes. J Stat Plan Inference 141:3725–3735
36. Tsai CC, Tseng ST, Balakrishnan N (2011) Optimal burn-in policy for highly reliable products using gamma degradation process. IEEE Trans Reliab 60:234–245
37. Tsai CC, Tseng ST, Balakrishnan N (2012) Optimal design for degradation tests based on gamma processes with random effects. IEEE Trans Reliab 61:604–613
38. Tsai TR, Sung WY, Lio YL, Chang SI, Lu JC (2016) Optimal two-variable accelerated degradation test plan for gamma degradation processes. IEEE Trans Reliab 65:459–468
39. Tseng ST, Tang J, Ku IH (2003) Determination of burn-in parameters and residual life for highly reliable products. Nav Res Logist 50:1–14
40. Tseng ST, Balakrishnan N, Tsai CC (2009) Optimal step-stress accelerated degradation test plan for gamma degradation processes. IEEE Trans Reliab 58:611–618
41. van Noortwijk JM (2009) A survey of the application of gamma processes in maintenance. Reliab Eng Syst Saf 94:2–21
42. Wang X (2008) A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects. Stat Sin 18:1153–1163
43. White LV (1973) An extension of the general equivalence theorem to nonlinear models. Biometrika 60:345–348
44. White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50: 1–25
45. Whitmore GA, Schenkelberg F (1997) Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. Lifetime Data Anal 3:27–45
46. Yang G (2007) Life cycle reliability engineering. John Wiley & Sons, Hoboken
47. Ye ZS, Shen Y, Xie M (2012) Degradation-based burn-in with preventive maintenance. Eur J Oper Res 62:360–367
48. Ye ZS, Wang Y, Tsui KL, Pecht M (2013) Degradation data analysis using Wiener processes with measurement errors. IEEE Trans Reliab 62:772–780
49. Ye ZS, Xie M, Shen Y, Tang LC (2012) Degradation-based burn-in planning under competing risks. Technometrics 54:159–168
50. Ye ZS, Xie M, Tang LC, Chen N (2014) Semiparametric estimation of gamma processes for deteriorating products. Technometrics 56:504–513
51. Yu HF, Tseng ST (2002) Designing a screening degradation experiment. Nav Res Logist 49:514–526

Chapter 10
Misspecification Analysis of Gamma with Inverse Gaussian Degradation Processes

Sheng-Tsaing Tseng and Yu-Cheng Yao

Abstract Degradation models are widely used to assess the lifetime information of highly reliable products. In this study, motivated by a stress relaxation data, we investigate the misspecification effect on the prediction of product's mean time to failure (MTTF) when the degradation model is wrongly fitted. Assuming the true model comes from gamma degradation process, but wrongly treated as inverse Gaussian degradation process, we first derive an analytic expression for the asymptotic distribution of quasi maximum likelihood estimate (QMLE) of the product's MTTF. Next, the penalty for the model misspecification is addressed comprehensively. The result demonstrates that the effect on the accuracy of the product's MTTF prediction strongly depends on the ratio of critical value to the scale parameter of the gamma degradation process.

Keywords Degradation model • Gamma process • Inverse Gaussian process • Quasi maximum likelihood estimate • Misspecification analysis

10.1 Introduction

Assessing the reliability information (e.g., the mean time to failure (MTTF) or the pth-quantile of lifetime distribution) of products is an essential task in the continual enhancement of a product's quality and reliability. For highly reliable products, however, it is difficult to assess the lifetime using traditional life tests, which only record time-to-failure. Even the technique of adopting an accelerated test with higher levels of stress, including elevated temperatures or voltages, is of little help, since no failures are likely to occur over a reasonable period of time. In this case, if the products have quality characteristics (QCs) whose degradation over time is related to reliability, then collecting degradation data or accelerated degradation data can provide timely lifetime information for highly reliable products.

S.-T. Tseng (✉) • Y.-C. Yao
Institute of Statistics, National Tsing Hua University, 30043, Hsinchu, Taiwan, Republic of China
e-mail: sttseng@stat.nthu.edu.tw; s9924802@m99.nthu.edu.tw

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_10

Bagdonavicius and Nikulin [2], Lawless [5], Meeker and Escobar [7], Nelson [8] and Tseng and Yu [15] have all described approaches of assessing reliability information via degradation data. Stochastic process formulations are well-known for this problem in the literature, and [12] has provided a survey on these stochastic processes for describing failure-mechanisms. Among them, Wiener process is one of the prominent ones and has been studied rather extensively. Doksum and Normand [4] presented Wiener process to describe the relationship between biomarker process values at random time points and a vector of covariates in a HIV study. Whitmore and Schenkelberg [19] proposed a time-transformed Wiener process for modeling the resistance of self-regulating heating cables. Yu and Tseng [22] used a Wiener process for addressing the problem of choosing the optimal settings of variables that are influential in the correct identification of significant factors and the experimental cost. Tseng et al. [16] described the linearized light intensity of LED lamps of contact image scanners by a Wiener process. Tang and Su [13] proposed the first-passage times of the test units over certain predetermined non-failure thresholds during the early stage of Wiener degradation process.

From the definition of a Wiener process, it is known that the degradation path is not a strictly increasing function. Generally speaking, a gamma process or inverse Gaussian process (possessing monotone increasing pattern) is more suitable for describing the degradation path of some specific products that display a monotone behavior such as in the case of crack growth. Bagdonavicius and Nikulin [1] modeled degradation paths by a gamma process, and included possibly time-dependent covariates. Lawless and Crowder [6] constructed a tractable gamma process incorporating a random effect, and also discussed a goodness-of-fit test for testing the validity of the model. Park and Padgett [9] provided several new degradation models that incorporate an accelerated test variable based on stochastic processes such as gamma process. Crowder and Lawless [3] used a gamma process to illustrate their single-inspection policy for the maintenance of automobile brakepads. Wang and Xu [17] studies the maximum likelihood estimation of a class of inverse Gaussian process models for degradation data and investigate the subject-to-subject heterogeneity and covariate information which can be incorporated into the model in a natural way. Zhang et al. [23] describes an inverse Gaussian process-based model formulated in a hierarchical Bayesian framework to characterize the growth of the depth of corrosion defects on underground energy pipelines based on inspection data.

Pascual and Montepiedra [10] and Yu [21] discussed lognormal and Weibull accelerated life test (ALT) plans under the misspecification of lifetime distribution. Both ALT and degradation models are standard approaches for assessing reliability information of highly reliable products. Hence, the effect of model misspecification on the accuracy and precision of the lifetime prediction becomes an important issue. Peng and Tseng [11] proposed a general formulation of linear degradation path based on Wiener process and addressed the effects of model misspecification on the prediction of products MTTF. Clearly, a gamma process or inverse Gaussian is better suited for modeling a degradation behavior that has a strictly increasing pattern. Tsai et al. [14] derived an expression for the asymptotic distribution of quasi maximum

likelihood estimate (QMLE) of the products MTTF when the true model comes from gamma degradation process, but is wrongly assumed to be Wiener degradation process and carry out a Monte Carlo simulation study to examine the effect of the corresponding model misspecification for reverse problem.

However, model misspecification between Weiner and gamma process with degradation path of some specific products that display a monotone behavior is not reasonable. Gamma and inverse Gaussian process are usually treated to each other, which makes the model misspecification problem become interesting and important. In the following, motivated by a stress relaxation data, we use the approach of [18] to address the effect of such a model misspecification on the products pth-quantile prediction between gamma and inverse Gaussian processes.

The rest of this paper is organized as follows. Section 10.2 uses an example to describe the motivation of this work. With a known power-linearization function, Sect. 10.3 discusses the effects of model misspecification on the accuracy and precision of the products pth-quantile prediction when the true degradation model comes from a gamma process, but is wrongly treated as an inverse Gaussian process. Section 10.4 also discusses the misspecification problem between gamma and inverse Gaussian process with an unknown power-linearization parameter. Section 10.5 uses the motivating example presented earlier to examine the effects of model misspecification. Finally, some concluding remarks are made at the end of this paper.

10.2 A Motivating Example

In the following, we adopt a stress relaxation data (from [20]) to illustrate the motivation of this study. Stress relaxation is the loss of stress in a component subjected to a constant strain over time (in 3%). Due to excessive stress relaxation, the contacts of electrical connectors often fail and a connector is considered to be failed when its stress relaxation exceeds a pre-defined failure level ω (say $\omega = 30$). Figure 10.1 shows the plot of stress relaxation over time for 6 tested units for temperature 85°C.

The degradation path showed in Fig. 10.1 seems to be a nonlinear function of time, t. However, based on the relationship of stress relaxation and aging time described in [20], the suggestion of a monotonic transformation is $\Lambda(t;\delta) = t^{\delta}$ satisfied the initial condition $\Lambda(0;\delta) = 0$.

If the degradation path is treated as a gamma degradation model, then the following process can be used to model the degradation path of the stress relaxation data:

$$\mathrm{M_G} : Y(t) \sim \mathrm{Ga}(\alpha\Lambda(t;\delta_1), \beta), \tag{10.1}$$

where $\mathrm{Ga}(\alpha, \beta)$ is gamma distribution with shape parameter α and rate parameter β Similarly, if the degradation path is treated as an inverse Gaussian degradation

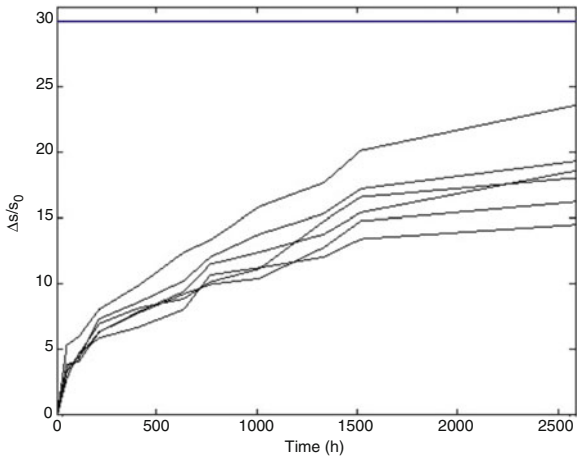

Fig. 10.1 Degradation paths for stress relaxation data

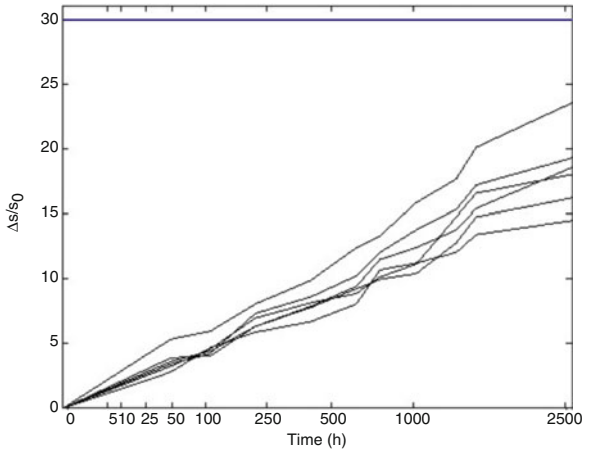

Fig. 10.2 Gamma degradation paths for stress relaxation data with transformed time ($\delta = 0.3889$)

model, then the following process can also be used to model degradation path of the stress relaxation data:

$$\mathrm{M_{IG}} : Y(t) \sim \mathrm{IG}(\mu \Lambda(t; \delta_2), \lambda \Lambda^2(t; \delta_2)), \tag{10.2}$$

where $\mathrm{IG}(\mu, \lambda)$ is inverse Gaussian distribution with location parameter μ and shape parameter λ. By using MLE, we can obtain $\hat{\delta}_1 = 0.3889$ and $\hat{\delta}_2 = 0.3823$. Figures 10.2 and 10.3 show the degradation path with transformed aging time which is plugged MLEs of δ_1 and δ_2 in for gamma and inverse Gaussian degradation models, respectively. The results demonstrate that both two models are appropriate to describe the stress relaxation degradation data.

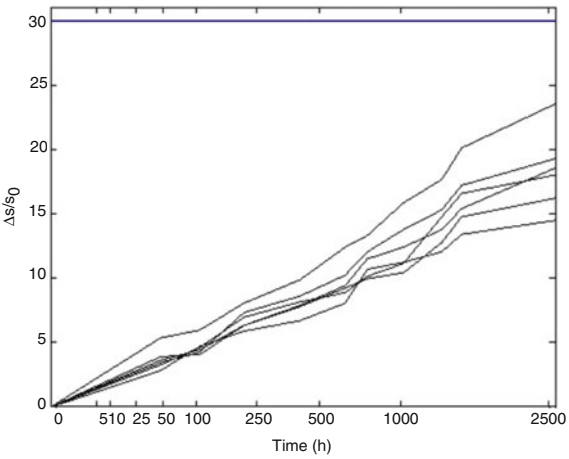

Fig. 10.3 Inverse gaussian degradation paths for stress relaxation data with transformed time ($\delta = 0.3823$)

Table 10.1 Estimations of pth-quantile of time to failure

p	Gamma	IG	RB	RV
0.5%	3432.89	2727.81	20.54%	1.39
1.0%	3828.79	3177.58	17.01%	1.41
5.0%	5082.16	4634.10	−8.82%	1.43
10.0%	5862.46	5555.10	−5.24%	1.42
50.0%	9336.29	9684.28	3.73%	1.33
90.0%	14143.70	15328.62	8.38%	1.25
95.0%	15790.00	17229.19	9.11%	1.23
99.0%	19257.51	21174.82	9.96%	1.20
99.5%	20658.54	22747.29	10.11%	1.19

Table 10.1 also shows the corresponding relative bias (RB) and relative variability of products pth-quantile under various combinations of $q = 0.5\%, 1\%, 5\%, 10\%, 50\%, 90\%, 95\%, 99\%$ and 99.5% if the true degradation model comes from Gamma process.

Most of all products pth-quantile, lower-percentile especially, have significant difference. The pth-quantile of time to failure can be affected by model misspecification, which turns out to be an interesting research topic. We now state the decision problem as follows: How do we address the effects of model misspecification on accuracy and precision of estimating products pth-quantile when M_G is misspecified as M_{IG}.

10.3 Misspecifying Gamma Process as Inverse Gaussian Process When δ Is Known

In some basic engineering applications, the relationship between degradation path of the products QC and aging time calculated by physical property and mechanism is well-known, which can support that monotonic transformation is independent of unknown parameter. Hence, we first assume δ is known for bringing out the effects of model misspecification based on a products pth-quantile of time to failure. For the above mentioned two models: $\mathrm{M_G}$ and $\mathrm{M_{IG}}$ we will address the effects of model misspecification. More specifically, we first apply the result of [18] to derive an expression for the asymptotic distribution of QMLE of the products pth-quantile of time to failure when the true degradation model is $\mathrm{M_G}$, but is fitted wrongly by model $\mathrm{M_{IG}}$. The effect of this model misspecification can then be addressed sequentially.

Assume that n units are tested, and the degradation measurements of each unit are available at time $t_1, t_2, \ldots, t_m$. Set $\boldsymbol{\theta}_{\mathrm{G}|\delta} = (\alpha, \beta)$ and $\boldsymbol{\theta}_{\mathrm{IG}|\delta} = (\mu, \lambda)$ are parameters of $\mathrm{M_G}$ and $\mathrm{M_{IG}}$ respectively. When δ is known, $\Lambda(t;\delta)$ can be simply treated as $\Lambda(t)$. The sample path of the ith unit at time t_j is given by

$$Y_i(t_j) \sim \mathrm{Ga}(\alpha\Lambda(t_j), \beta), 1 \le i \le n, 1 \le j \le m.$$

For fixed $\boldsymbol{\theta}_{\mathrm{G}|\delta}$ let $\boldsymbol{\theta}^*_{\mathrm{G}|\delta}$ be the value of $\boldsymbol{\theta}_{\mathrm{G}|\delta}$ that minimizes the expected negative log-likelihood $\mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}|\delta}))$ with respect to $\mathrm{M_G}$ that is,

$$\boldsymbol{\theta}^*_{\mathrm{IG}|\delta} = \arg\min_{\boldsymbol{\theta}_{\mathrm{IG}|\delta}} \mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}|\delta})).$$

Let

$$Z_{ij} = Y_i(t_j) - Y_i(t_{j-1}),$$

and

$$\Delta(t_j) = \Lambda(t_j) - \Lambda(t_{j-1}).$$

By the independent increment property of the inverse Gaussian process, the log-likelihood function under model $\mathrm{M_{IG}}$ can be expressed as follows:

$$L(\boldsymbol{\theta}_{\mathrm{IG}|\delta}) = \prod_{i=1}^{n}\prod_{j=1}^{m} \sqrt{\frac{\lambda\Delta^2(t_j)}{2\pi z_{ij}^3}} \mathrm{e}^{-\frac{\lambda(z_{ij}-\mu\Delta(t_j))^2}{2\mu^2 z_{ij}}}. \tag{10.3}$$

Then, we have

$$\mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}|\delta})) \propto -\frac{1}{2} nm \ln\lambda + k_0, \tag{10.4}$$

where

$$k_0 = \frac{\lambda n \alpha \Lambda(t_m)}{2\mu^2\beta} - \frac{n\lambda}{2}\left(\frac{2\Lambda(t_m)}{\mu} - \sum_{j=1}^{m}\frac{\beta\Delta^2(t_j)}{\alpha\Delta(t_j)-1}\right).$$

By minimizing $\mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}|\delta}))$ with respect to μ and λ , we obtain the following:

$$\boldsymbol{\theta}^*_{\mathrm{IG}|\delta} = (\mu^*, \lambda^*) = \left(\frac{\alpha}{\beta}, \frac{m}{\beta}\left(\sum_{j=1}^{m}\frac{\Delta^2(t_j)}{\alpha\Delta(t_j-1)} - \frac{\Lambda(t_m)}{\alpha}\right)^{-1}\right). \tag{10.5}$$

Now, let

$$\mathbf{A}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) = \begin{pmatrix} -\frac{n\lambda^*\Lambda(t_m)}{\mu^{*3}} & 0 \\ 0 & -\frac{mn}{2\lambda^{*2}} \end{pmatrix}, \tag{10.6}$$

$$\mathbf{B}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) = \begin{pmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{pmatrix}, \tag{10.7}$$

where

$$B_{11} = \frac{n\lambda^{*2}\alpha\Lambda(t_m)}{\mu^{*6}\beta^2}$$

$$B_{12} = -\frac{n\lambda^*}{2\mu^{*3}}\left(\frac{\Lambda(t_m)}{\alpha} + \sum_{j=1}^{m}\frac{\Delta^2(t_j)}{1-\alpha\Delta(t_j)}\right),$$

$$B_{22} = \frac{n\beta^2}{4}\left(\sum_{j=1}^{m}\frac{\Delta^4(t_j)}{(\alpha\Delta(t_j)-1)^2(\alpha\Delta(t_j)-2)} + k_1\right),$$

$$k_1 = \frac{\Lambda(t_m)}{\alpha^3} + \frac{2}{\alpha^2}\sum_{j=1}^{m}\frac{\Delta^2(t_j)}{1-\alpha\Delta(t_j)},$$

and

$$\mathbf{C}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) = \mathbf{A}^{-1}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) \cdot \mathbf{B}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) \cdot \mathbf{A}^{-1}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}).$$

Let $\hat{t}_{p,\mathrm{IG}|\delta}$ denote the QMLE of the product's pth-quantile of time to failure when the true model is $\mathrm{M_G}$ Then

$$\hat{t}_{p,\mathrm{IG}|\delta} = t_{p,\mathrm{IG}|\delta}(\hat{\boldsymbol{\theta}}_{\mathrm{IG}|\delta}), \tag{10.8}$$

where $\hat{\boldsymbol{\theta}}_{\mathrm{IG}|\delta}$ is QMLE of $\boldsymbol{\theta}_{\mathrm{IG}|\delta}$ obtained by maximizing the likelihood function $L(\boldsymbol{\theta}_{\mathrm{IG}|\delta})$ in (10.3), and $t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}_{\mathrm{IG}|\delta})$ satisfies

$$\Phi\left(\sqrt{\frac{\lambda}{\omega}}\left(z_0-\frac{\omega}{\mu}\right)\right)-\mathrm{e}^{\frac{2\lambda z_0}{\mu}}\,\Phi\left(-\sqrt{\frac{\lambda}{\omega}}\left(z_0+\frac{\omega}{\mu}\right)\right)=p$$

where $z_0 = \Lambda(t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}_{\mathrm{IG}|\delta}))$ Then, by theorem 3.2 of [18] and the use of delta-method, we obtain the following result.

Theorem 10.1

$$\hat{t}_{p,\mathrm{IG}|\delta} \sim \mathrm{N}(t^*_{p,\mathrm{IG}|\delta}, \gamma(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta})), \tag{10.9}$$

where

$$t^*_{p,\mathrm{IG}|\delta} = t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}^*_{\mathrm{IG}|\delta}), \tag{10.10}$$
$$\gamma(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) = (\nabla t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}^*_{\mathrm{IG}|\delta}))^{\mathrm{T}} \cdot \mathbf{C}(\boldsymbol{\theta}_{\mathrm{G}|\delta} : \boldsymbol{\theta}^*_{\mathrm{IG}|\delta}) \cdot \nabla t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}^*_{\mathrm{IG}|\delta}),$$

and

$$\nabla t_{p,\mathrm{IG}|\delta}(\boldsymbol{\theta}_{\mathrm{IG}|\delta}) = \left(\frac{\partial t_{p,\mathrm{IG}|\delta}}{\partial \mu}, \frac{\partial t_{p,\mathrm{IG}|\delta}}{\partial \lambda}\right)^{\mathrm{T}}.$$

Let $\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G})$ denote the approximate bias of using $\hat{t}_{p,\mathrm{IG}|\delta}$ to estimate $t_{p,\mathrm{IG}|\delta}$ when the true model is $\mathrm{M_G}$. From (10.9), we get

$$\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G}) = \mathrm{E}(\hat{t}_{p,\mathrm{IG}|\delta}) - t_{p,\mathrm{G}|\delta} = t^*_{p,\mathrm{IG}|\delta} - t_{p,\mathrm{G}|\delta}, \tag{10.11}$$

where $t_{p,\mathrm{G}|\delta} = t_{p,\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta})$ is the product's pth-quantile of time to failure respect to the true model $\mathrm{M_G}$, satisfying

$$\frac{\Gamma(\alpha\Lambda(t_{p,\mathrm{G}|\delta}), \beta\omega)}{\Gamma(\alpha\Lambda(t_{p,\mathrm{G}|\delta}))} = p, \tag{10.12}$$

$\Gamma(\alpha, z)$ is the incomplete gamma function defined by

$$\Gamma(\alpha, z) = \int_z^\infty x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x,$$

and $\Gamma(\alpha) = \Gamma(\alpha, 0)$.

Now, we define the following criterion to measure the relative bias of model misspecification:

$$\kappa_{\mathrm{IG}|\mathrm{G},\delta} = \frac{\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G})}{t_{p,\mathrm{G}|\delta}}. \tag{10.13}$$

Similarly, let $\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G})$ denote the approximate mean square error of $\hat{t}_{p,\mathrm{IG}|\delta}$ when it is used to estimate the product's pth-quantile of time to failure when the true model is $\mathrm{M_G}$. From (10.9) and (10.11), we obtain

$$\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G}) = \gamma(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) + (t^*_{p,\mathrm{IG}|\delta} - t_{p,\mathrm{G}|\delta})^2. \tag{10.14}$$

The approximate mean square error takes into account two types of errors – the random sampling error and the estimation bias.

To measure the relative variability of model misspecification, we use the following criterion:

$$\varphi_{\mathrm{IG}|\mathrm{G},\delta} = \frac{\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}|\delta} \mid \mathrm{M_G})}{\mathrm{AMSE}(\hat{t}_{p,\mathrm{G}|\delta} \mid \mathrm{M_G})}. \tag{10.15}$$

Note that $\mathrm{AMSE}(\hat{t}_{p,\mathrm{G}|\delta} \mid \mathrm{M_G}) = \mathrm{AVar}(\hat{t}_{p,\mathrm{G}|\delta})$ and

$$\mathrm{AVar}(\hat{t}_{p,\mathrm{G}|\delta}) = (\nabla t_{p,\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta}))^{\mathrm{T}} \cdot \mathbf{I}^{-1}_{\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta}) \cdot \nabla t_{p,\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta}), \tag{10.16}$$

where $\mathbf{I}_{\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta})$ is Fisher information matrix of gamma process, and

$$\nabla t_{p,\mathrm{G}|\delta}(\boldsymbol{\theta}_{\mathrm{G}|\delta}) = \left(\frac{\partial t_{p,\mathrm{G}|\delta}}{\partial \alpha}, \frac{\partial t_{p,\mathrm{G}|\delta}}{\partial \beta} \right)^{\mathrm{T}}.$$

10.4 Misspecifying Gamma Process as Inverse Gaussian Process When δ Is Unknown

The relationship between degradation path of the product's QC and aging time, in general, is unknown, but can be described by a power-function transformation with an unknown parameter δ such as $\Lambda(t;\delta) = t^{\delta}$. Assuming that δ is unknown, the parameters $\boldsymbol{\theta}_{\mathrm{G}}$ and $\boldsymbol{\theta}_{\mathrm{IG}}$ in $\mathrm{M_G}$ and $\mathrm{M_I G}$ can be extended to $\boldsymbol{\theta}_{\mathrm{G}} = (\alpha, \beta, \delta_1)$ and $\boldsymbol{\theta}_{\mathrm{IG}} = (\mu, \lambda, \delta_2)$ respectively. Moreover, $\Delta(t_j;\delta)$ is redefined as $\Delta(t_j;\delta) = \Lambda(t_j;\delta) - \Lambda(t_{j-1};\delta)$. From (10.3), the log-likelihood function under model $\mathrm{M_{IG}}$ can be expressed as follows:

$$\mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}})) \propto -\frac{1}{2} nm \ln \lambda + k_2, \tag{10.17}$$

where

$$k_2 = \frac{\lambda n \alpha \Lambda(t_m;\delta_1)}{2\mu^2\beta} - \frac{n\lambda}{2}\left(\frac{2\Lambda(t_m;\delta_2)}{\mu} - \sum_{j=1}^{m}\frac{\beta\Delta^2(t_j;\delta_1)}{\alpha\Delta(t_j;\delta_1)-1}\right).$$

By minimizing $\mathrm{E}_{\mathrm{M_G}}(-\ln L(\boldsymbol{\theta}_{\mathrm{IG}}))$ with respect to μ and λ, we obtain the following:

$$\boldsymbol{\theta}^*_{\mathrm{IG}} = (\mu^*, \lambda^*, \delta_2^*) = \left(\frac{\alpha\Lambda(t_m;\delta_1)}{\beta\Lambda(t_m;\delta_2)}, g_\lambda(\delta_2^*;\alpha,\beta,\delta_1), \delta_2^*\right), \tag{10.18}$$

where

$$g_\lambda(\delta_2^*;\alpha,\beta,\delta_1) = \frac{m}{\beta}\left(\sum_{j=1}^{m}\frac{\Delta^2(t_j;\delta_2)}{\alpha\Delta(t_j;\delta_1)-1} - \frac{\Lambda^2(t_m;\delta_2)}{\alpha\Lambda(t_m;\delta_1)}\right)^{-1},$$

and δ_2^* satisfies

$$\frac{1}{m}\sum_{j=1}^{m}\frac{\partial}{\partial\delta_2^*}\ln\Delta(t_j;\delta_2^*) = -\frac{1}{2}\frac{\partial \ln g_\lambda(\delta_2^*;\alpha,\beta,\delta_1)}{\partial\delta_2^*}.$$

Now, set

$$\mathbf{A}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) = \mathrm{E}_{\mathrm{M_G}}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta}_{\mathrm{IG}})}{\partial\theta_{\mathrm{IG}r}\partial\theta_{\mathrm{IG}s}}\right)\Bigg|_{\boldsymbol{\theta}_{\mathrm{IG}}=\boldsymbol{\theta}^*_{\mathrm{IG}}}, \tag{10.19}$$

$$\mathbf{B}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) = \mathrm{E}_{\mathrm{M_G}}\left(\frac{\partial \ln L(\boldsymbol{\theta}_{\mathrm{IG}})}{\partial\theta_{\mathrm{IG}r}}\frac{\partial \ln L(\boldsymbol{\theta}_{\mathrm{IG}})}{\partial\theta_{\mathrm{IG}s}}\right)\Bigg|_{\boldsymbol{\theta}_{\mathrm{IG}}=\boldsymbol{\theta}^*_{\mathrm{IG}}}, \tag{10.20}$$

and

$$\mathbf{C}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) = \mathbf{A}^{-1}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) \cdot \mathbf{B}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) \cdot \mathbf{A}^{-1}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}),$$

where $\theta_{\mathrm{IG}r}$ is the rth element of $\boldsymbol{\theta}_{\mathrm{IG}}$.

Let $\hat{t}_{p,\mathrm{IG}}$ denote the QMLE of the product's pth-quantile of time to failure when the true model is $\mathrm{M_G}$ Then

$$\hat{t}_{p,\mathrm{IG}} = t_{p,\mathrm{IG}}(\hat{\boldsymbol{\theta}}_{\mathrm{IG}}), \tag{10.21}$$

where $\hat{\boldsymbol{\theta}}_{\mathrm{IG}} = (\hat{\mu}, \hat{\lambda}, \hat{\delta_2})$ is QMLE of $\boldsymbol{\theta}_{\mathrm{IG}}$ and $t_{p,\mathrm{IG}}(\boldsymbol{\theta}_{\mathrm{IG}})$ satisfies

$$\Phi\left(\sqrt{\frac{\lambda}{\omega}}\left(z_1 - \frac{\omega}{\mu}\right)\right) - \mathrm{e}^{\frac{2\lambda z_1}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{\omega}}\left(z_1 + \frac{\omega}{\mu}\right)\right) = p$$

where $z_1 = \Lambda(t_{p,\mathrm{IG}}(\boldsymbol{\theta}_{\mathrm{IG}}); \delta_2)$ Then, by theorem 3.2 of [18] and the use of delta-method, we obtain the following result.

Theorem 10.2

$$\hat{t}_{p,\mathrm{IG}} \sim \mathrm{N}(t^*_{p,\mathrm{IG}}, \gamma(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}})), \tag{10.22}$$

where

$$t^*_{p,\mathrm{IG}} = t_{p,\mathrm{IG}}(\boldsymbol{\theta}^*_{\mathrm{IG}}), \tag{10.23}$$

$$\gamma(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) = (\nabla t_{p,\mathrm{IG}}(\boldsymbol{\theta}^*_{\mathrm{IG}}))^{\mathrm{T}} \cdot \mathbf{C}(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) \cdot \nabla t_{p,\mathrm{IG}}(\boldsymbol{\theta}^*_{\mathrm{IG}}),$$

and

$$\nabla t_{p,\mathrm{IG}}(\boldsymbol{\theta}_{\mathrm{IG}}) = \left(\frac{\partial t_{p,\mathrm{IG}}}{\partial \mu}, \frac{\partial t_{p,\mathrm{IG}}}{\partial \lambda} \right)^{\mathrm{T}}.$$

Let $\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G})$ denote the approximate bias of using $\hat{t}_{p,\mathrm{IG}}$ to estimate $t_{p,\mathrm{IG}}$ when the true model is $\mathrm{M_G}$. From (10.22), we get

$$\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G}) = \mathrm{E}(\hat{t}_{p,\mathrm{IG}}) - t_{p,\mathrm{G}} = t^*_{p,\mathrm{IG}} - t_{p,\mathrm{G}}, \tag{10.24}$$

where $t_{p,\mathrm{G}} = t_{p,\mathrm{G}}(\boldsymbol{\theta}_{\mathrm{G}})$ is the product's pth-quantile of time to failure respect to the true model $\mathrm{M_G}$, satisfying

$$\frac{\Gamma(\alpha \Lambda(t_{p,\mathrm{G}}; \delta_1), \beta\omega)}{\Gamma(\alpha \Lambda(t_{p,\mathrm{G}}; \delta_1))} = p. \tag{10.25}$$

Now, we define the following criterion to measure the relative bias of model misspecification:

$$\kappa_{\mathrm{IG|G}} = \frac{\mathrm{ABias}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G})}{t_{p,\mathrm{G}}}. \tag{10.26}$$

Similarly, let $\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G})$ denote the approximate mean square error of $\hat{t}_{p,\mathrm{IG}}$ when it is used to estimate the product's pth-quantile of time to failure when the true model is $\mathrm{M_G}$. From (10.22) and (10.24), we obtain

$$\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G}) = \gamma(\boldsymbol{\theta}_{\mathrm{G}} : \boldsymbol{\theta}^*_{\mathrm{IG}}) + (t^*_{p,\mathrm{IG}} - t_{p,\mathrm{G}})^2. \tag{10.27}$$

To measure the relative variability of model misspecification, we use the following criterion:

$$\varphi_{\mathrm{IG|G}} = \frac{\mathrm{AMSE}(\hat{t}_{p,\mathrm{IG}} \mid \mathrm{M_G})}{\mathrm{AMSE}(\hat{t}_{p,\mathrm{G}} \mid \mathrm{M_G})}. \tag{10.28}$$

Note that $\mathrm{AMSE}(\hat{t}_{p,\mathrm{G}} \mid \mathrm{M_G}) = \mathrm{AVar}(\hat{t}_{p,\mathrm{G}})$ and

$$\mathrm{AVar}(\hat{t}_{p,\mathrm{G}}) = (\nabla t_{p,\mathrm{G}}(\boldsymbol{\theta}_{\mathrm{G}}))^{\mathrm{T}} \cdot \mathbf{I}_{\mathrm{G}}^{-1}(\boldsymbol{\theta}_{\mathrm{G}}) \cdot \nabla t_{p,\mathrm{G}}(\boldsymbol{\theta}_{\mathrm{G}}), \tag{10.29}$$

where $\mathbf{I}_{\mathrm{G}}(\boldsymbol{\theta}_{\mathrm{G}})$ is Fisher information matrix of gamma process, and

$$\nabla t_{p,\mathrm{G}}(\boldsymbol{\theta}_{\mathrm{G}}) = \left(\frac{\partial t_{p,\mathrm{G}}}{\partial \alpha}, \frac{\partial t_{p,\mathrm{G}}}{\partial \beta}, \frac{\partial t_{p,\mathrm{G}}}{\partial \delta_1} \right)^{\mathrm{T}}.$$

10.5 Data Analysis

Example 1 (the case when δ known) Given $\delta = 0.33$ assume that there are 20 samples ($n = 20$) with measurement time $(t_1, \ldots, t_8) = (1, 2, 4, 8, 16, 32, 64, 128)$ and the true degradation process follows a gamma process with parameter

$$\boldsymbol{\theta}_{\mathrm{G}|\delta} = (\alpha, \beta) = (15, 450). \tag{10.30}$$

Now, if the model is misspecified as inverse Gaussian process, then we have

$$\boldsymbol{\theta}^*_{\mathrm{IG}|\delta} = (\mu^*, \lambda^*) = (0.0333, 0.4299). \tag{10.31}$$

Given a critical value $\omega = -\log(0.5)$ and $p = 0.01$ substituting (10.30) into (10.12) and (10.31) into (10.10), we have $t_{.01,\mathrm{G}|\delta} = 6515.59$ and $t^*_{.01,\mathrm{IG}|\delta} = 6242.05$ respectively. Now from (10.11) and (10.13), we estimate the relative bias as

$$\kappa_{\mathrm{IG}|\mathrm{G},\delta} = \frac{\mathrm{ABias}(\hat{t}_{.01,\mathrm{IG}|\delta} \mid \mathrm{M_G})}{t_{.01,\mathrm{G}|\delta}} = -4.20\%.$$

Also, substituting the same parameter settings into (10.14) and (10.16), we have $\mathrm{AMSE}(\hat{t}_{.01,\mathrm{IG}|\delta} \mid \mathrm{M_G}) = 398413$ and $\mathrm{AMSE}(\hat{t}_{.01,\mathrm{G}|\delta} \mid \mathrm{M_G}) = 323951$ Hence, the relative variability $\varphi_{\mathrm{IG}|\mathrm{G},\delta}$ is 398413/323951 = 1.23. From the relative bias and the relative variability, the effects of the model misspecification here seems to be intensely serious.

These results demonstrate that the effect of model misspecification on the accuracy and precision of the product's 1% of time to failure prediction is serious under the parameter setting of simulated data, and it is seen that the product's pth-quantile of time to failure affect the relative bias and relative variability. In the following, we will discuss the sensitivity analysis of the relative bias and relative variability for the product's pth-quantile of time to failure.

As the same as previous parameters setting, Fig. 10.4 shows that the sensitivity analysis of relative bias $\kappa_{\mathrm{IG}|\mathrm{G},\delta}$ for the product's pth-quantile of time to failure. From Fig. 10.4, we observe that the relative bias is far away 0 at tail-quantile. Furthermore,

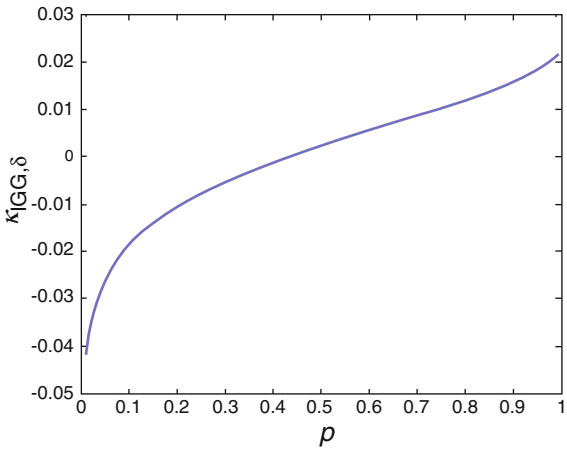

Fig. 10.4 The sensitivity analysis of relative bias, $\kappa_{\mathrm{IG}|\mathrm{G},\delta}$

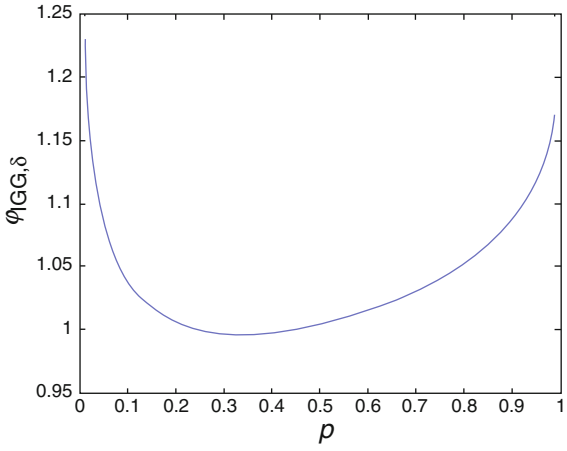

Fig. 10.5 The sensitivity analysis of relative variability, $\varphi_{\mathrm{IG}|\mathrm{G},\delta}$

Fig. 10.5 presents the plot of the sensitivity analysis of relative variability $\varphi_{\mathrm{IG}|\mathrm{G},\delta}$. The result shows that the relative variability become pretty large when p goes to 0 or 1. Hence, the effects on the accuracy and precision of the product's pth-quantile of time to failure prediction become critical when p closes to 0 or 1.

Example 2 (the case when δ unknown) Next, we use the stress relaxation data in motivating example (where the power functions for each unit are different) for evaluating the effects of the model misspecification. In this example, we assume that the true model is $\mathrm{M_G}$. The MLE of $\boldsymbol{\theta}_{\mathrm{G}}$ can be determined to be

$$\hat{\boldsymbol{\theta}}_{\mathrm{G}} = (\hat{\alpha}, \hat{\beta}, \hat{\delta}_1) = (1.6030, 1.8590, 0.3889). \tag{10.32}$$

In the following, we treat the above estimate $\hat{\boldsymbol{\theta}}_{\mathrm{G}}$ as the true parameter $\boldsymbol{\theta}_{\mathrm{G}}$. In addition, based on real data shown in Fig. 10.1 ($n = 6, (t_1, \ldots, t_{10}) = (46, 108, 212, 344, 446, 626, 729, 927, 1005, 1218)$), and substituting $\boldsymbol{\theta}_{\mathrm{G}}$ into (10.18), we can

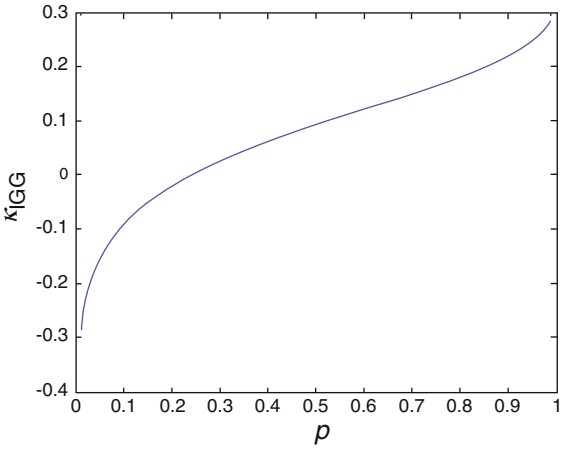

Fig. 10.6 The sensitivity analysis of relative bias, $\kappa_{\mathrm{IG|G}}$

obtain

$$\boldsymbol{\theta}^*_{\mathrm{IG}} = (\mu^*, \lambda^*, \delta_2^*) = (1.0031, 1.0473, 0.3697). \tag{10.33}$$

Given a critical value $\omega = 30$ and $p = 0.01$ substituting (10.32) into (10.25) and (10.33) into (10.23), we have $t_{.01,\mathrm{G}} = 3828.74$ and $t^*_{.01,\mathrm{IG}} = 2537.35$ respectively. Now from (10.24) and (10.26), we estimate the relative bias as

$$\kappa_{\mathrm{IG|G}} = \frac{\mathrm{ABias}(\hat{t}_{.01,\mathrm{IG}} \mid \mathrm{M_G})}{t_{.01,\mathrm{G}}} = -33.73\%.$$

Also, substituting the same parameter settings into (10.27) and (10.29), then we have $\mathrm{AMSE}(\hat{t}_{.01,\mathrm{IG}} \mid \mathrm{M_G}) = 1697277$ and $\mathrm{AMSE}(\hat{t}_{.01,\mathrm{G}} \mid \mathrm{M_G}) = 788210$ Therefore, the relative variability $\varphi_{\mathrm{IG|G}}$ is 1697277/788210 = 2.15. From the relative bias (RB) and the relative variability (RV), the effects of the model misspecification in relaxation data seems to be quite serious. These results demonstrate that the effect of model misspecification on the accuracy and precision of the product's 1% of time to failure prediction is serious under the parameter setting of relaxation data. Furthermore, Figs. 10.6 and 10.7 also show the sensitivity analysis of RB ($\kappa_{\mathrm{IG|G}}$) and RV ($\varphi_{\mathrm{IG|G}}$) for the product's pth-quantile of time to failure, where $0 \leq p \leq 1$ respectively. From the results, it demonstrates that the RB is very serious at the tailed-quantile. The same situation also happened to the RV shown in Fig. 10.7.

10.6 Conclusions

Motivated by a real stress-relaxation data, we discuss the effects of model misspecification when the true degradation model is a gamma process and an inverse Gaussian process with an unknown linearization transformation parameter, but is

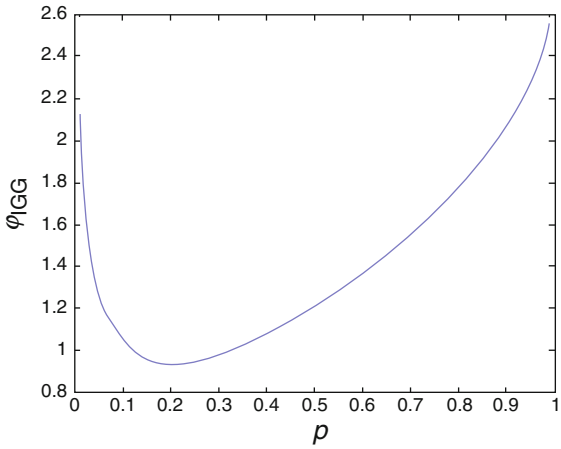

Fig. 10.7 The sensitivity analysis of relative variability, $\varphi_{IG|G}$

misspecified as an inverse Gaussian process. The relative bias and relative variability are used to measure the effects of such a model misspecification. The results demonstrate that the effects on the accuracy and precision of the product's pth-quantile prediction become critical when p is close to 0 or 1. Finally, one remark is addressed. The dual problem of this study what is the misspecification effect when the true degradation model is an inverse Gaussian process; but wrongly misspecified as gamma degradation process shall also be an interesting and challenging topic for future research.

References

1. Bagdonavicius V, Nikulin MS (2001) Estimation in degradation models with explanatory variables. Lifetime Data Anal 7:85–103
2. Bagdonavicius V, Nikulin MS (2002) Accelerated life models: modeling and statistical analysis. Chapman & Hall/CRC, Boca Raton
3. Crowder M, Lawless J (2007) On a scheme for predictive maintenance. Eur J Oper Res 176:1713–1722
4. Doksum KA, Normand SLT (1995) Gaussian models for degradation processes-part I: methods for the analysis of biomarker data. Lifetime Data Anal 1:131–144
5. Lawless JF (2002) Statistical models and methods for lifetime data. John Wiley & Sons, New York
6. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
7. Meeker WQ, Escobar LA (1998) statistical methods for reliability data. John Wiley & Sons, New York
8. Nelson W (1990) Accelerated testing: statistical models, test plans, and data analyses. John Wiley & Sons, New York
9. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma process. Lifetime Data Anal 11:511–527

10. Pascual FG, Montepiedra G (2005) Lognormal and Weibull accelerated life test plans under distribution misspecification. IEEE Trans Reliab 54:43–52
11. Peng CY, Tseng ST (2009) Mis-specification analysis of linear degradation models. IEEE Trans Reliab 58:444–455
12. Singpurwalla ND (1995) Survival in dynamic environments. Stat Sci 10:86–103
13. Tang J, Su TS (2008) Estimating failure time distribution and its parameters based on intermediate data from a Wiener degradation model. Nav Res Logist 55:265–276
14. Tsai CC, Tseng ST, Balakrishnan N (2011) Optimal burn-in policy for highly reliable products using gamma degradation process. IEEE Trans Reliab 60:234–245
15. Tseng ST, Yu HF (1997) A termination rule for degradation experiments. IEEE Trans Reliab 46:130–133
16. Tseng ST, Tang J, Ku IH (2003) Determination of burn-in parameters and residual life for highly reliable products. Nav Res Logist 50:1–14
17. Wang X, Xu D (2010) An inverse Gaussian process model for degradation data. Technometrics 52:188–197
18. White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50: 1–25
19. Whitmore GA, Schenkelberg F (1997) Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. Lifetime Data Anal 3:27–45
20. Yang G (2007) Life Cycle Reliability Engineering. John Wiley & Sons, New York
21. Yu HF (2007) Designing a screening experiment with a reciprocal Weibull degradation rate. Comput Ind Eng 52:175–191
22. Yu HF, Tseng ST (2002) Designing a screening experiment for highly reliable products. Nav Res Logist 49:514–526
23. Zhang S, Zhou W, Qin H (2013) Inverse Gaussian process-based corrosion growth model for energy pipelines considering the sizing error in inspection data. Corros Sci 73:309–320

Part III
Applications

Chapter 11
Practical Applications of a Family of Shock-Degradation Failure Models

Mei-Ling T. Lee and G. A. Whitmore

Abstract Many systems experience gradual degradation while simultaneously being exposed to a stream of random shocks of varying magnitude that eventually cause failure when a shock exceeds the residual strength of the system. This failure mechanism is found in diverse fields of application. Lee and Whitmore Shock-degradation failure processes and their survival distributions. Manuscript submitted for journal publication, 2016) presented a family of system failure models in which shock streams that follow a Fréchet process are superimposed on a degrading system described by a stochastic process with stationary independent increments. They referred to them as *shock-degradation failure models*. In this article, we discuss applications of these models and investigate practical issues and extensions that help to make these models more accessible and useful for studies of system failure. This family has the attractive feature of defining the failure event as a first passage event and the time to failure as a first hitting time (FHT) of a critical threshold by the underlying stochastic process. FHT models have found use in many real-world settings because they describe the failure mechanism in a realistic manner and also naturally accommodate regression structures. This article discusses a variety of data structures for which this model is suitable, as well as the estimation methods associated with them. The data structures include conventional censored survival data, data sets that combine readings on system degradation and failure event times, and data sets that include observations on the timing and magnitudes of shocks. This assortment of data structures is readily handled by *threshold regression* estimation procedures. Predictive inferences and risk assessment methods are also available. This article presents three case applications related to osteoporotic hip fractures in elderly women, divorces for cohorts of Norwegian couples, and deaths of cystic fibrosis patients. This article closes with discussion and concluding remarks.

M.-L.T. Lee (✉)
School of Public Health, University of Maryland, 20742, College Park, MD, USA
e-mail: MLTLEE@UMD.EDU

G.A. Whitmore
McGill University, 1001 rue Sherbrooke Ouest, H3A 1G5, Montreal, QC, Canada
e-mail: george.whitmore@mcgill.ca

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_11

Keywords Applications • Degradation process • Fréchet process • Shock stream • Strength process • Survival distribution • Threshold regression

11.1 Introduction

Many systems experience gradual deterioration while being exposed simultaneously to a stream of random shocks of varying magnitude that eventually cause failure when a shock exceeds the residual strength of the system. This basic situation is found in many fields of application as illustrated by the following examples:

1. An equipment component experiences normal wear and tear during use but is also exposed to mechanical shocks from random external forces of varying intensity. The weakening component finally fails when a sufficiently strong shock arrives to break it.
2. A new business may experience gradual financial erosion or strengthening through time while experiencing shocks or disturbances from economic events that impact its business sector. The business fails if an external shock forces it into bankruptcy.
3. The human skeleton weakens with age and disease but is also exposed to occasional traumatic events such as blunt force injuries, stumbles or falls. The skeleton 'fails' when the trauma produces a skeletal fracture.
4. Many lung diseases such as chronic obstructive pulmonary disease (COPD) or cystic fibrosis involve a progressive deterioration of lung function. The time course of these diseases is punctuated by occasional acute exacerbations, brought on by infections or other assaults, that create moments of life-threatening crisis for the patient. A sufficiently weak patient dies when an exacerbation overwhelms the body's defenses.
5. The cohesion of a marriage may gradually weaken or strengthen through time while simultaneously being subjected to stresses or disturbances of varying severity, such as financial problems, alcoholism or infidelity. The marriage fails if a shock is sufficient to break the marriage bond.

In this article, we consider practical applications of a family of shock-degradation failure models that were initially described in [7]. This family is one in which shock streams are superimposed on a degrading underlying condition for the system and failure occurs when a shock causes the weakened system to fail. The family of models has already found several applications, which will be introduced later. These shock-degradation models have the attractive feature of defining the failure event and failure time as the first hitting time (FHT) of a threshold by the underlying stochastic process. Such FHT models are useful in practical applications because, first, they usually describe the failure mechanism in a realistic manner and, second, they naturally accommodate regression structures that subsequently can be analyzed and interpreted using *threshold regression* methods. The family encompasses a wide class of underlying degradation processes, as demonstrated in [7]. The shock stream itself is assumed to be generated by a Fréchet process that we describe momentarily.

11.2 Shock-Degradation Failure Models

We now describe the two principal components of this family of shock-degradation failure models.

11.2.1 The Shock Process

The kind of shocks that impact a system require careful definition so the potential user of the model will know if it is suitable for a particular application or may require some extension or modification (of which there are many). As illustrated by the examples presented in the introduction, a *shock* refers to any sudden and substantial force applied to the system. The shock may be a sharp physical force, a sudden and significant economic disturbance, a physiological reaction to trauma or assault, or an instance of major emotional distress, depending on the context. We consider only shocks that create a momentary negative excursion or displacement in the condition of the system, with the system returning to its previous state or condition after the shock is absorbed. The shocks are therefore not cumulative in their impact. For example, if a fall doesn't produce a bone fracture then there is no lasting damage. The shock is also concentrated at a 'moment in time'. This moment may be a few minutes, hours or days depending on the time scale of the application. For example, a COPD exacerbation may work itself out over a few days or weeks but this interval is just a 'moment' when considered over a life span of years.

To define a Fréchet shock process, we start with any partition of the open time interval $(0, s]$ into n intervals $(0, t_1], \ldots, (t_{n-1}, t_n]$, where $0 = t_0 < t_1 < \cdots < t_n = s$. The largest interval in the time partition is denoted by $\Delta = \max_j(t_j - t_{j-1})$. Let V_j be a positive value associated with interval $(t_{j-1}, t_j]$. Assume that the sequence $V_1, \ldots, V_n$ is a set of mutually independent draws from the following cumulative distribution function (c.d.f.):

$$P(V_j \leq v) = G(v)^{t_j - t_{j-1}} \quad \text{for each } j = 1, \ldots, n. \tag{11.1}$$

Here $G(v)$ is the cumulative distribution function (c.d.f.) of a Fréchet distribution, which has the following mathematical form [3]:

$$G(v) = \exp\left[-(\alpha/v)^{\beta}\right] \quad \text{for } v > 0, \alpha > 0, \beta > 0. \tag{11.2}$$

The Fréchet process $\{V(t), 0 < t \leq s\}$ associated with the generating distribution in (11.2) is the limiting sequence of $V_1, \ldots, V_n$ as the norm of the time partition Δ tends to 0 (and n increases accordingly). The process can be extended analytically

to the whole positive real line by allowing s to increase without limit. At its essence, the process $\{V(t)\}$ describes a steam of local maxima in the sense that the maximum of $V(t)$ for all t in any open interval $(a, b]$ has the c.d.f. $G(v)^{b-a}$. Thus, the process is defined by:

$$P\left[\max_{\tau\in(0,t]} V(\tau) \le v\right] = G(v)^t \quad \text{for all } v > 0, t > 0. \tag{11.3}$$

Lee and Whitmore [7] pointed out a number of properties of the Fréchet shock process. Importantly, they noted that scale parameter α is the $\exp(-1)$th fractile or 37th percentile of the maximum shock encountered in one unit of time, irrespective of the value of shape parameter β. They also show that larger values of β tighten the distribution of shock magnitudes about this fixed percentile.

11.2.2 The Degradation Process

The shock-degradation model assumes that the system of interest has an initial strength $y_0 > 0$ and that this strength degrades over time. The system strength process is denoted by $\{Y(t), t \ge 0\}$ and the degradation process is denoted by $\{W(t), t \ge 0\}$. The model assumes $\{W(t)\}$ is a stochastic process with stationary independent increments and initial value $W(0) = 0$. It also assumes that the degradation process has a cumulant generating function (c.g.f.) defined on an open set $\mathscr{Z}$, denoted by $\kappa(u) = \ln E\{\exp[uW(1)]\}$. The system strength process at time t is $Y(t) = y_0 \exp[W(t)]$. This type of degradation process is one in which the system strength changes in increments that are *proportional* to the residual strength at any moment. As a result, the system strength $Y(t)$ never reaches a point of zero strength but may approach it asymptotically. Therefore, in this model setting, the system cannot fail through degradation alone. The failure occurs when a shock exceeds the residual strength.

Lee and Whitmore [7] considered the following two families of stochastic processes that possess stationary independent increments as important illustrations of the class:

Wiener diffusion process. A Wiener diffusion process $\{W(t), t \ge 0\}$, with $W(0) = 0$, having mean parameter μ, variance parameter $\sigma^2 \ge 0$ and c.g.f.:

$$\kappa(u) = \ln E\{\exp[uW(1)]\} = u(\mu + u\sigma^2/2). \tag{11.4}$$

Gamma process. A gamma process $\{X(t), t \ge 0\}$, with $X(0) = 0$, having shape parameter $\zeta > 0$ and scale parameter $\eta > 0$. Define $W(t)$ as the negative of $X(t)$, that is, $W(t) = -X(t)$ in order to assure that the sample paths of $\{W(t)\}$

are monotonically decreasing. The c.g.f. of $W(t)$, as a negative-valued gamma process, has the form:

$$\kappa(u) = \ln E\{\exp[uW(1)]\} = \zeta \ln\left(\frac{\eta}{\eta + u}\right) \quad \text{for } \eta + u > 0. \tag{11.5}$$

Lee and Whitmore [7] also considered the following family of models in which degradation has a smooth deterministic trajectory:

Deterministic exponential process. System strength follows a deterministic exponential time path of the form $Y(t) = y_0 \exp(\lambda t)$ where λ denotes the exponential rate parameter.

This family is simple but an adequate practical representation for some important applications.

11.2.3 The Shock-Degradation Survival Distribution

Lee and Whitmore [7] showed that their shock-degradation model has the following survival distribution:

$$\overline{F}(s) = E_{\mathscr{C}}\left[P(S > s|\mathscr{C})\right] = E_{\mathscr{C}}\left\{\exp\left[-cQ(s)\right]\right\}, \tag{11.6}$$

where S denotes the survival time,

$$c = \left(\frac{\alpha}{y_0}\right)^{\beta} \quad \text{and} \quad Q(s) = \int_0^s e^{-\beta W(t)} dt. \tag{11.7}$$

The notation $E_{\mathscr{C}}$ denotes an expectation over the set of degradation sample paths $\mathscr{C} = \{W(t) : 0 \le t \le s, W(0) = 0\}$. The random quantity $Q(s)$ is seen to be a stochastic integral of the strength process, raised to power $-\beta$, over time interval $(0, s]$. The authors went on to show how the survival function can be expanded in a series involving expected moments of $Q(s)$ and derive exact formulas for these moments. Both a Taylor series expansion and an Euler product-limit series expansion are considered. The authors pointed out that conditions for the Fubini-Tonelli Theorem, governing the interchange of integration and expectation operators for the infinite series implicit in (11.6), fails to hold. Thus, they proposed the use of the following finite expansions as approximations for the survival function $\overline{F}(s)$:

Taylor series kth-order expansion

$$\overline{F}_k(s) = \exp(-cm_1)\left[\sum_{\ell=0}^{k}(-1)^{\ell}\frac{c^{\ell}}{\ell!}m_{\ell}^{*}\right]. \tag{11.8}$$

Euler product-limit kth-order expansion

$$\overline{F}_k(s) = \exp(-cm_1)\left[\sum_{\ell=0}^{k}(-1)^{\ell}\frac{k!}{\ell!(k-\ell)!}\left(\frac{c}{k}\right)^{\ell} m_{\ell}^{*}\right]. \tag{11.9}$$

Here $m_1 = E_{\mathscr{C}}[Q(s)]$ is the expected first moment of $Q(s)$ and $m_{\ell}^{*} = E_{\mathscr{C}}[Q_{*}(s)^{\ell}], \ell = 0, 1, 2, \ldots,$ are the expected central moments of $Q(s)$, with $Q_{*}(s) = Q(s) - m_1$. By definition, $m_0^{*} = 1$ and $m_1^{*} = 0$.

Lee and Whitmore [7] also showed that the following log-survival function is a lower bound for the logarithm of the exact survival function (11.6):

$$\ln \overline{F}_L(s) = \exp(-cm_1) = \exp\{-cE_{\mathscr{C}}[Q(s)]\} = -c\left[\frac{e^{\kappa s}-1}{\kappa}\right] \tag{11.10}$$

where κ is short-hand for $\kappa(-\beta)$, the c.g.f. evaluated at $-\beta$. This lower bound is quite tight for degradation processes with modest variability, especially for survival time below the median. Note that this lower-bound survival function is the leading expression $\exp(-cm_1)$ found in approximations (11.8) and (11.9).

As pointed out by Lee and Whitmore [7], both the exact survival function (11.6) and its lower bound in (11.10) are over-parameterized when they are estimated from censored survival data alone. Such data cannot be informative about every feature of the degradation and shock processes. Readings on the underlying degradation process $\{W(t)\}$ or shock process $\{V(t)\}$ individually are needed to estimate and validate the full model. This insight is completely reasonable because survival outcomes alone have limited information content when it comes to separating the influences of degradation and shocks in the failure mechanism.

A number of simple extensions for the shock-degradation model make it quite adaptable to practical application. The model can be modified to include a power transformed time scale, an analytical time scale or an exogenous cure rate, as may be required by a particular application. For example, a geometric Wiener degradation model can have an endogenous cure rate if the underlying process is one of growing strength rather than degrading strength (as in the social context of marriage dissolution, for example, or the survival of a business organization). An exogenous cure rate, on the other hand, recognizes that a subgroup of individuals under study may not be susceptible to failure. Our case applications demonstrate a power transformed time scale as well as cure rates of both the exogenous and endogenous variety.

Next, we present a few properties of the lower bound for the survival function given in (11.10) that are useful in understanding the behavior of the shock-degradation model where the degradation process is not highly variable.

Hazard function. The hazard function corresponding to the lower-bound survival function (11.10) is:

$$h_L(s) = c\exp(\kappa s). \tag{11.11}$$

Thus, the hazard is increasing, constant or decreasing exponentially according to whether κ is positive, 0 or negative, respectively. As $h_L(0) = c = (\alpha/y_0)^\beta$, we see quite reasonably that the initial risk of failure depends on α/y_0, the ratio of the 37th percentile shock to the initial strength of the system, modulated by the shock shape parameter β.

Probability density function. The probability density function (p.d.f.) of the lower-bound survival function (11.10) is:

$$\ln f_L(s) = \ln h_L(s) + \ln \overline{F}_L(s) = \ln(c) + \kappa s - c\left[\frac{e^{\kappa s} - 1}{\kappa}\right]. \tag{11.12}$$

The density function has a positive density of c at the origin and steadily declines as s increases if $\kappa \leq c$. If $\kappa > c$, the density function has a single mode at $\ln(\kappa/c)/\kappa$.

Cure rate. The lower-bound survival function in (11.10) has a positive probability of never failing if parameter $\kappa < 0$; specifically, $\ln P(S = \infty) = c/\kappa$.

11.3 Data Structures

Practical applications of shock-degradation models present a wide variety of data structures. We now elaborate on some of these data structures and their associated measurement and observation challenges.

11.3.1 Direct Readings on the Degradation Process

We first consider the topic of obtaining direct readings on the degradation process. As pointed out by Lee and Whitmore [7], readings on their shock-degradation process are unlikely to be disturbed significantly by the shock process when the readings are gathered over brief moments of time. Put differently, large shocks occur so infrequently in the time continuum that they are rarely encountered when making one or a few observations on the underlying degradation process in a limited observation period. As a practical illustration, a patient suffering from COPD may only have an acute exacerbation a few times each year so the patient's lung condition on any random day of the year is likely to be close to its chronic stable level. Moreover, in actual clinical practice, a patient visit for spirometric testing is usually only scheduled when the patient's condition is known to be stable.

To see the point mathematically, equation (11.1) gives the following probability for the maximum shock V observed in a time interval of length $\Delta t > 0$:

$$P(V \leq v) = G(v)^{\Delta t} = \exp\left[-\Delta t\,(\alpha/v)^\beta\right]. \tag{11.13}$$

For fixed values of α, β and $v > 0$, this formula gives a probability that approaches 1 as Δt approaches zero. Thus, in this limiting sense, the probability of a material shock is vanishingly small if the process is observed at any arbitrary moment of time and, moreover, the probability of any material shock remains vanishingly small if shocks are observed for any finite number of such moments. Relaxing this theoretical statement to accept that Δt may not be vanishingly small, the mathematics implies that no shock of practical significance will be present in any finite number of observation intervals if each interval is sufficiently short. To illustrate the point numerically, consider a shock level v equal to α, the 37th percentile, and set β to 1. If the unit time is, say, one calendar year and $\Delta t = 1/365$ (representing one day) then $P(V \leq v) = 0.997$. Thus, on only 3 occasions in 1000 will the maximum shock on that day exceed the 37th percentile maximum shock for a year. The practical lesson of this result is that significant shocks in a shock-degradation process occur sparingly and are rarely discovered at random moments of observation.

11.3.2 Observations on Failure Times, System Strength and Shocks

The preceding section has focussed on the survival function, which is an essential mathematical element for analyzing censored survival data for systems. Yet, actual applications frequently offer opportunities to observe and measure the underlying strength or degradation of the system at time points prior to failure, at failure, or at withdrawal from service or end of study. Observations on the strength process refer to measurements of the intrinsic strength of the system at moments of relative stability which are little influenced by shocks, as we noted earlier. Thus, in general, applications may involve censored survival data that are complemented by periodic readings on the underlying condition of the system.

To expand on the types of observation on degradation that may be available, we mention that degradation readings might be feasible for all systems under study. Alternatively, they may be available only for survivors because failure may destroy the system and make any reading on residual strength of the failing system impossible. In contrast, readings may be available only for failing items because measurement of residual strength may require a destructive disassembly of a system, which will not be feasible for systems still in service.

Our model assumes that strength and, therefore, degradation follows a one-dimensional process with failure being triggered when system strength is forced to zero by a shock. In reality, however, strength is usually a complex multidimensional process. Moreover, readings on the underlying strength may be unavailable. Rather, investigators may only have access to readings on a marker or surrogate process so that only indirect measurements of degradation are at hand. The marker process might be highly correlated with the actual degradation process but a high correlation

is not assured. For example, electromigration of metallic material in the presence of an electrical field is a spatially complex degradation process that can cause failure of electronic components such as integrated circuits. This degradation process might be monitored by a single performance measure (for example, transported mass) that only captures the phenomenon in a crude manner. An alternative scenario is one in which we estimate a surrogate strength process that is a composite index (a linear regression function, say) constructed from readings on one or more observable marker processes that can be monitored through time. The practical realities of the context determine what can be done to produce an adequate model for the degradation process in the actual application.

Opportunities occasionally arise to observe and measure shocks in a variety of ways. For example, in COPD, various markers describe the magnitude of an acute exacerbation including severity of symptoms, recovery time, and the like. As a practical matter, however, shocks are often not observed if they are minor. In this situation, it is a left-truncated shock distribution that is observed and the truncation point may need to be estimated from the data. Again, considering COPD exacerbations as an illustration, these may not be measured unless they are above a threshold of severity; for instance, severe enough to require prescription medication or hospitalization.

The most general data situation encountered in practical application is one in which survival times and readings on degradation are jointly observed for individual systems. The data record for each individual system is a longitudinal one that consists of a sequence of readings (none, one, two, or more) on the strength or degradation process, gathered at irregularly spaced time points, together with a survival outcome (either a censoring time or a failure time). The readings on the strength or degradation process may be observations on the actual process or on one or more marker processes. A regression structure can be introduced if the setting provides data on relevant covariates for system parameters. In this situation, maximum likelihood estimation can be used. The evolving state or condition of a system often possesses the Markov property, which allows a tidy handling of longitudinal data under the threshold regression approach. Published methods for threshold regression with longitudinal data found in [5] have immediate application in this situation, as one of our later case applications will demonstrate. Applications that involve joint observation of survival and degradation require mathematical extensions of previous results [6, 8]. In the next section, we summarize pertinent results from [7].

11.4 Joint Observation of Survival and Degradation

We now consider the data elements found in a longitudinal record consisting of periodic readings on system degradation, ending with either a failure event or a survival event (that is, a right-censored failure event). Our line of development

follows the method of Markov decomposition proposed in [5]. We limit our study now to the lower-bound survival function in (11.10). The likelihood function for a longitudinal record will be a product of conditionally independent events of two types. The first type is a failure event. In this event, the system has an initial strength, say y_0, and then fails at time s later. The likelihood of this event is given by a p.d.f. like that in (11.12). The second type is a survival event; more precisely, an event in which the system has an initial strength, say y_0, survives beyond a censoring time s (that is, $S > s$) and has a recorded strength at time s of, say, $y > 0$. Note that $y = Y(s) = y_0 \exp[W(s)] = y_0 \exp(w)$, where w denotes the amount of degradation corresponding to strength y. Thus, this second event involves the following set of strength sample paths: $\mathscr{C}^* = \{Y(t) : 0 \le t \le s, Y(0) = y_0, Y(s) = y = y_0 \exp(w)\}$; in other words, the sample paths are pinned down at two end points but are otherwise free to vary. The likelihood of this second type of event is $P\{S > s, W(s) \in [w, w + dw]\}$. This probability can be factored into the product of a conditional survival probability and the p.d.f. for degradation level w at time s, as follows:

$$P\{S > s, W(s) \in [w, w + dw]\} = P(S > s|w)g(w)dw = \overline{F}(s|w)g(w)dw. \quad (11.14)$$

The p.d.f. $g(w)$ is known from the specified form of the degradation process. The conditional survival function $\overline{F}(s|w)$ is less straightforward and has yet to be derived in a general form for the shock-degradation model. Lee and Whitmore [7] presented a more limited but very useful mathematical result which we now employ. For a degradation process with modest variability, they noted that a lower bound on the survival function is quite tight. For a pinned degradation process, they showed:

$$\ln \overline{F}(s|w) = \ln E_{\mathscr{C}^*}\{\exp[-cQ(s|w)]\} \ge -cE_{\mathscr{C}^*}[Q(s|w)] = \ln \overline{F}_L(s|w), \quad (11.15)$$

where c and $Q(s)$ are defined as in (11.7). Expectation $E_{\mathscr{C}^*}[Q(s|w)]$ does not have a general closed form for the family of degradation processes with stationary independent increments. However, Lee and Whitmore [7] derived a closed form for $\ln \overline{F}_L(s|w)$ in the important case where the degradation process is a Wiener process. The derivation builds on properties of a Brownian bridge process. Their formula for the lower bound of the survival function for a pinned Wiener process is as follows:

$$\ln \overline{F}_L(s|w) = -c\sqrt{\frac{2\pi s}{\beta^2\sigma^2}} \exp(z_1^2/2)\left[\Phi(z_2) - \Phi(z_1)\right], \quad (11.16)$$

where

$$z_1 = \frac{w - \beta(\sigma^2/2)s}{\sqrt{\sigma^2 s}}, \quad z_2 = \frac{w + \beta(\sigma^2/2)s}{\sqrt{\sigma^2 s}},$$

and $\Phi(\cdot)$ denotes the standard normal c.d.f.. The p.d.f. of the degradation level $W(s)$ at time s in the Wiener case is given by

$$g(w) = \frac{1}{\sqrt{2\pi\sigma^2 s}} \exp[-(w-\mu s)^2/2\sigma^2 s]. \tag{11.17}$$

Thus, (11.16) and (11.17) are the two components required for evaluating the joint probability in (11.14) for lower-bound survival of a system beyond a censoring time $S > s$ and having the degradation level w at time s.

11.5 Case Applications

We next present three case applications of shock-degradation models. The first application concerns osteoporotic hip fractures and has been published elsewhere. We summarize it briefly to demonstrate the wide range of potential applications. The second application looks at Norwegian divorces. It has not been published previously so we present more details on its development and findings. The third application considers survival times for cystic fibrosis patients. A deterministic version of the shock-degradation model was previously published for this case. We present an extension of the published model that incorporates explicitly the stochastic time course of patient lung function.

11.5.1 Osteoporotic Hip Fractures

Osteoporotic hip fractures in elderly women were studied by He et al. [4] using this shock-degradation model in conjunction with threshold regression. They studied times to first and second fractures. The underlying strength process in their model represented skeletal health. The shock process represented external traumas, such as falls and stumbles, which taken together with chronic osteoporosis, might trigger a fracture event. Threshold regression was used to associate time to fracture with baseline covariates of study participants.

The system strength model used by He et al. [4] is the deterministic exponential process that we described in Sect. 11.2.2, for which the log-survival function is:

$$\ln \overline{F}(s) = \left(\frac{\alpha}{y_0}\right)^{\beta} \left[\frac{e^{-\beta\lambda s} - 1}{\beta\lambda}\right]. \tag{11.18}$$

The authors equated parameters $\ln(y_0)$ and λ to linear combinations of covariates while fixing the remaining parameters. The data set consists of censored times to first and second fractures. The models were fitted by maximum likelihood methods

using (11.18). Technical details and discussion of the study findings can be found in the original publication.

11.5.2 Norwegian Divorces

Extensive data on the durations of marriages contracted in Norway are found in [9]. The data for the 1960, 1970 and 1980 marriage cohorts were presented and carefully analyzed by Aalen et al. [1]. They used the data to demonstrate various statistical concepts, models and techniques related to time-to-event data analysis. We use the same cohort data here to demonstrate the shock-degradation model. In this example, only censored survival data for the marriages are available. The data set has no longitudinal measurements on the strength of the marital unions themselves. Recall our explanation in the introduction that marriage might be viewed as a shock-degradation process in which some marriages actually tend to strengthen through time. In our model, a marriage is subjected to a stream of minor and major shocks and, if the marriage fails, the failure will occur at the moment of a shock that exceeds the residual strength of the marital bond. The data appear in Table 5.2 of [1]. The marriage duration numbers represent completed years of marriage. The midpoints of the yearly intervals are taken as the durations for divorces that occurred in the year. The marriage cohorts ignore cases lost to follow-up because of death or emigration. Marriages lasting more than 35, 25 and 15 years for the 1960, 1970 and 1980 cohorts are censored. The respective sample sizes are 23,651, 29,370 and 22,230 for these three cohorts.

We use the maximum likelihood method to estimate our model. We assume at the outset that the degradation process is a Wiener diffusion process for which $\kappa = \kappa(-\beta) = -\beta\mu + \beta^2\sigma^2/2$. We use the lower-bound log-survival function (11.10) and its corresponding log-hazard function (11.11) in the sample likelihood calculation. We expected this bound to be reasonably tight to the actual survival function in this application and our sensitivity analysis discussed shortly justifies the approach. We incorporate regression functions for model parameters. Parameters are made to depend on linear combinations of covariates with logarithmic link functions being used for positive-valued parameters. As only censored survival data are available, some parameters of the process model cannot be estimated. Referring to the survival function in (11.10) for the Wiener case, we see immediately that μ and σ for the process cannot be estimated separately. Hence, we can only estimate the parameter $\xi = \mu - \beta\sigma^2/2$. Further, the survival function depends on α and y_0 only through the ratio α/y_0. In terms of logarithms, the ratio becomes $\ln(\alpha) - \ln(y_0)$. Thus, these two logarithmic regression functions cannot be distinguished mathematically and are not estimable by maximum likelihood unless the regression functions depend on different sets of covariates. This mathematical result points out the practical reality that censored survival data cannot separate the effect of a covariate on the degradation process from its effect on the scale parameter α of the shock process when the covariates act multiplicatively, as implied by the use of log-linear

regression functions. We therefore do not estimate α but rather set its value to 1. As α is the 37th percentile of the shock distribution irrespective of β, setting α to 1 makes this percentile the *unit measure* for the latent degradation process. Even with α set to 1, parameters β, y_0 and ξ cannot be independently estimated from censored survival data if all three have intercept terms. We therefore fix one more parameter and choose to set $\xi = 0.001$. A positive value is chosen in anticipation that some marriages are not prone to fail. The magnitude 0.001 is chosen to produce a convenient scaling of the parameter values. In the end, therefore, only parameters β and y_0 will be estimated for the principal model. In addition, however, we also wish to extend the model using a power transformation of the time scale, that is, a transformation of form $s = t^{\gamma}$ with $\gamma > 0$. We estimate power exponent γ using a logarithmic link function. The power transformation implies that marriage breakdown occurs on a time scale that is accelerating or decelerating relative to calendar time depending on the value of exponent γ.

Given the preceding specifications for the model, the sample log-likelihood function in this case application has the form:

$$\ln L(\boldsymbol{\theta}) = \sum_{i \in N} (\alpha / y_0)^{\beta} \left[\exp(-\beta \xi s_i) - 1\right] / (\beta \xi) + \sum_{i \in N_1} \beta \ln(\alpha / y_0) - \beta \xi s_i + \ln(\gamma) + (\gamma - 1) \ln(t_i). \quad (11.19)$$

Here $\boldsymbol{\theta}$ denotes the vector of regression coefficients. Index set N includes all couples in the sample data set and index set N_1 is the subset of couples for whom divorce durations are observed. The observed or censored marriage duration for couple i is denoted by t_i and $s_i = t_i^{\gamma}$ is the power-transformed duration. The terms in the first sum of (11.19) are log-survival probabilities. Those in the second sum are log-hazard values, taken together with the values $\ln(\gamma) + (\gamma - 1) \ln(t_i)$ which represent the log-Jacobian terms for the power transformation. The log-survival probabilities and log-hazard values are based on formulas for the lower-bound found in (11.10) and (11.11), respectively.

Table 11.1 presents the threshold regression results for the Norwegian marriage duration data. The parameters β and γ are made to depend on indicator variables for the marriage cohort through log-linear regression functions. The assumption is that socio-economic trends in Norway may have changed marriage stability over these three decades. The estimate for the initial marital strength y_0 is 3.871. This value is in units of the 37th percentile annual maximum shock. In other words, the initial marital strength is just under four modest shocks away from divorce. Estimates of β, the shape parameter for shocks, vary slightly (but significantly) across the cohorts, being 4.967 in 1960, 4.655 in 1970 and 4.744 in 1980. The magnitudes of this parameter suggest that modest shocks are numerous but extreme shocks that threaten a marriage are rare. For example, given the estimate for β in 1960 (4.967), the initial marital strength of 3.871 represents the 99.9th percentile of the maximum annual shock. Estimates of the power exponent γ for the time scale transformation

Table 11.1 Threshold regression estimates for the shock-degradation model applied to duration data for Norwegian marriage cohorts from 1960, 1970 and 1980. The 1960 cohort serves as the reference

Parameter and covariate	Estimated reg. coef.	Std. err.	P-value
$\ln(y_0)$			
intercept	1.354	0.0588	0.000
$\ln(\beta)$			
1970 cohort	−0.065	0.0042	0.000
1980 cohort	−0.046	0.0039	0.000
intercept	1.603	0.0375	0.000
$\ln(\gamma)$			
1970 cohort	0.031	0.0096	0.001
1980 cohort	0.218	0.0091	0.000
intercept	0.527	0.0098	0.000

also vary moderately (but significantly) across the cohorts, being 1.693 in 1960, 1.746 in 1970 and 2.106 in 1980. The estimates indicate an accelerating analytical time. The implication is that shock patterns will occur on a more compressed time scale as the marriage lengthens, for those marriages that eventually end in divorce. Interestingly, the power exponent is increasing with the decade, suggesting that the shock patterns are accelerating across cohorts (so extreme shocks of given size occur more frequently in calendar time). Finally, the proportions of marriages that would eventually end in divorce if the time horizon were extended indefinitely (and spousal deaths are ignored) are 0.215 for 1960, 0.326 for 1970 and 0.290 for 1980. We looked at asymptotic correlation coefficients for parameter estimates for suggestions of multicollinearity. The output shows one large correlation coefficient, namely, −0.998 for the intercepts of $\ln(\beta)$ and $\ln(y_0)$. This large value indicates that the magnitudes of errors in these two estimates are almost perfectly offsetting; in other words, the pattern of shocks and initial marital strength are mathematically difficult to distinguish when only censored survival data are available.

Figure 11.1 compares the Kaplan-Meier plots and the fitted shock-degradation survival functions for marriage duration in the Norwegian marriage cohorts. The fits are quite good, considering that only covariates for the marriage cohort are taken into account. The slight lack of fit that does appear for the shock-degradation model may be produced in part by administrative artifacts that cause the timing of a divorce decree to deviate from our theoretical survival model. Under Norwegian law, a divorce can be granted under several conditions. Spouses can divorce after legal separation for one year or without legal separation if they have lived apart for two years. Divorce can also be granted in cases of abuse or if spouses are closely related. These laws have the effect of lengthening the recorded marriage duration beyond the first hitting time for the threshold. Likewise, administrative and judicial delays may further delay divorce decrees and, hence, lengthen duration. We do not try to model these administrative influences.

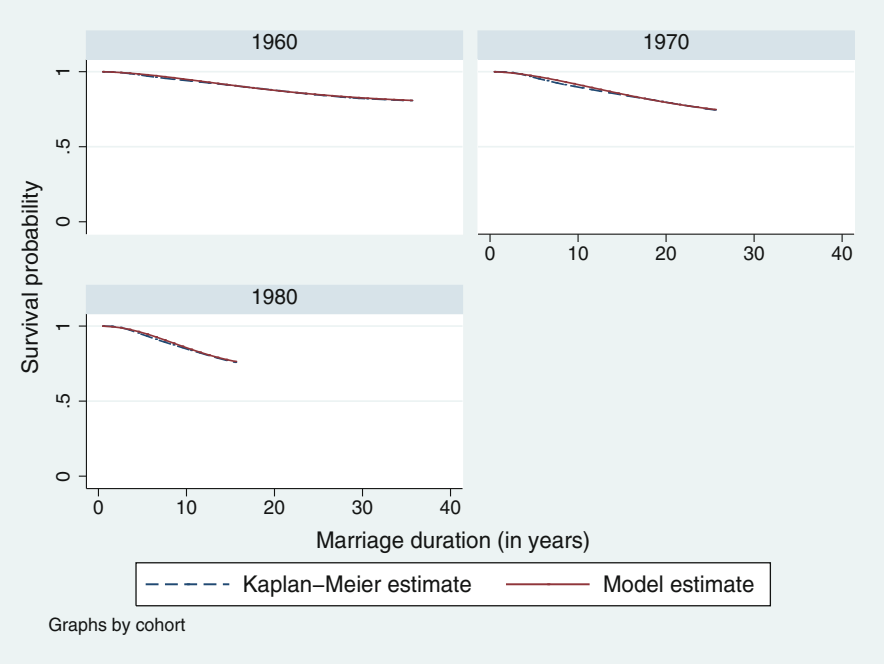

Fig. 11.1 Comparison of Kaplan-Meier plots and fitted shock-degradation survival functions for marriage duration in three Norwegian marriage cohorts

We conducted a sensitivity analysis on the model to see if the variance parameter σ^2 of the Wiener process might differ from zero and could be estimated. As the model is already parameter rich for a case application that uses only censored survival data, we take the fit of the lower-bound survival model as a starting point. We then fix the regression coefficients for the β parameter at their lower-bound fitted values. Next, we set aside the fixed value of 0.001 for ξ and allow μ and σ to vary in the model (recall that $\xi = \mu - \beta\sigma^2/2$). A 4th-order Euler product-limit approximation for the true survival function is used. Starting with $\mu = 0.001$ and $\sigma = 0$, the sample log-likelihood begins to decline slowly as σ increases away from zero, which indicates that the lower-bound fitted model cannot be improved. Specifically, the sample log-likelihood for the fitted lower-bound model in Table 11.1 is $-94,783.6$. This log-likelihood declines to $-94,787.2$ as σ increases from 0 to 0.009. The estimate of μ simultaneously increases from 0.001 to 0.00117. Estimates of the other parameters, $\ln(y_0)$ and the regression coefficients for $\ln(\gamma)$, change little as μ and σ change. As anticipated by the theoretical analysis in Lee and Whitmore [7], the numerical analysis reaches a breakdown point as the variability of the degradation process increases. In this case, the breakdown occurs just beyond $\sigma = 0.009$, which is a modest level of variability in this case setting.

11.5.3 Survival Times for Cystic Fibrosis Patients

Aaron et al. [2] used the shock-degradation model, together with threshold regression, to assess one-year risk of death for cystic fibrosis (CF) patients. Their data were drawn from the Canadian registry of cystic fibrosis patients. They used the version of the model having a deterministic exponential degradation trajectory. The authors developed a CF health index comprised of risk factors for CF chronic health, including clinical measurements of lung function. This CF health index was the strength process in their model. The shock process represented acute exacerbations, both small and large, to which CF patients are exposed. Death is triggered when an exacerbation overwhelms the residual health of a CF patient. The authors' modeling focussed on short-term risk of death and produced a risk scoring formula that could be used for clinical decision making.

We extend the model in Aaron et al. [2] and consider the estimation of a stochastic degradation process for cystic fibrosis. For this task, we have created a synthetic data set that imitates data from a CF registry sufficiently well to demonstrate the technicalities of our modeling approach. Briefly, our synthetic data set has 2939 cystic fibrosis patients who have lung function measured during regular clinic visits that are usually about one year apart. Patients are assumed to be stable during these scheduled visits, that is, they are not experiencing an acute exacerbation at the time of the visit. Our synthetic data set has 40,846 scheduled clinic visits for all patients combined and includes 532 patient deaths. As shown by Aaron et al. [2], the health status of a cystic fibrosis patient is affected by many risk factors but the most important by far is the lung measurement *forced expiratory volume in one second* (FEV1), expressed as a fraction or percentage of the same measurement for a healthy person of the same height, age and sex. We refer to the fractional value as FEV1pred, which stands for *FEV1 predicted*. We take the patient's FEV1pred as our strength measure and assume that its logarithm follows a Wiener diffusion process. Again, following Aaron et al. [2], we note that each patient has a longitudinal record of regular lung function measurements. We use the Markov decomposition approach to decompose each longitudinal record into a series of individual records as described in [5]. At each scheduled clinic visit, the patient either survives until the next scheduled visit or dies before the next visit. We therefore can estimate our Wiener shock-degradation model using maximum likelihood methods incorporating both censored survival times and degradation readings. We use the lower-bound results found in (11.16) and (11.17) to implement the maximum likelihood estimation, imitating the procedure in [2]. The parameters to be estimated by threshold regression methods for this shock-degradation model include α and β for the shock process and μ and σ for the Wiener degradation process. Note that y_0 and w are known because strength process $Y(t)$ is equated with the patient's evolving FEV1pred levels.

Table 11.2 Threshold regression estimates for a shock-degradation model applied to survival times of cystic fibrosis patients for a synthetic data set. The regression functions have no covariates

Parameter	Estimate	Std. err.
$\ln(\alpha)$	−2.042	0.0154
$\ln(\beta)$	1.313	0.0188
μ	−0.0448	0.0012
$\ln(\sigma)$	−1.445	0.0035

The log-likelihood function to be maximized in this case application is the following:

$$\ln L(\boldsymbol{\theta}) = \sum_{i \in N_0} \left[\ln \overline{F}_L(s_i|w_i) + \ln g(w_i) \right] + \sum_{i \in N_1} \ln f_L(s_i). \tag{11.20}$$

Here $\boldsymbol{\theta} = (\alpha, \beta, \mu, \sigma)$ denotes the vector of parameters to be estimated. Index set N_0 includes all clinic visits i in which the patient survives until the next scheduled clinic visit at time s_i and has a log-FEV1pred reading of w_i on that next visit. Index set N_1 is the set of clinic visits i in which the patient dies at time s_i after the visit (and before the next scheduled visit). Quantities $\ln \overline{F}_L(s_i|w_i)$ and $\ln g(w_i)$ are calculated from formulas (11.16) and (11.17), respectively. Quantity $\ln f_L(s_i)$ is calculated from formula (11.12).

Table 11.2 presents the parameter estimates. The sample size is $n = 40,846$ clinic visits. The table shows logarithmic estimates for α, β and σ. The maximum likelihood estimates for α and β are 0.130 and 3.72, respectively. As α represents the 37th percentile of maximum annual shocks for a Fréchet process, our estimate for α suggests that cystic fibrosis patients have acute exacerbations that are more severe than a decline of 13% points in FEV1pred in about 2 of every 3 years on average. Given the estimates for both parameters, one can calculate from (11.2) that the 90th percentile of the most severe annual exacerbation of a cystic fibrosis patient is about 24% points of FEV1pred. The estimate of μ for the Wiener degradation process suggests that FEV1pred declines about 4.48% annually in cystic fibrosis patients. The estimate for σ is 0.236. This value indicates that log-FEV1pred varies annually by this amount, which is equivalent to average annual variation of about 27% in FEV1pred. Finally, our estimation routine also generates an estimate of the asymptotic correlation matrix of the parameter estimates (not shown). The matrix is helpful in checking for multicollinearity. The only correlation coefficient that is moderately large is that for the estimates of $\ln(\alpha)$ and $\ln(\beta)$, which happens to be 0.65. The number shows that estimation errors for these parameters tend to lie in the same direction, that is, tending to be positive or negative together.

11.6 Discussion and Concluding Remarks

As we have illustrated, many systems in diverse fields of application degrade with time, asymptotically approaching a point of zero strength. These systems are often simultaneously exposed to a stochastic stream of shocks that momentarily weaken the system as they arrive. These systems ultimately fail when an incident shock exceeds the residual strength of the system. Thus, degradation is a contributing factor to failure but it is a shock that delivers the *coup de grâce*. Our illustrations also show that in certain circumstances a system that is strengthening rather than deteriorating with time may also fail because a sufficiently severe shock overwhelms its strength.

The survival function in (11.6) is based on a system strength process whose logarithm has stationary independent increments. This model formulation happens to be mathematically conjugate with the Fréchet process and, hence, yields a tidy mathematical form. If the geometric aspect is dropped in favor of a strength process that itself has stationary independent increments then some tractable forms of the survival function can be derived. For example, a deterministic linear strength process is a simple example. We do not develop or explore this alternative class of models here.

The shock stream assumed in our stochastic model does not produce cumulative damage, by assumption. In reality, however, there are shock systems that generate cumulative damage, such as fluctuating stresses that produce crack propagation in materials. A practical approximation to a cumulative damage process can be obtained within our modeling approach by assuming that the degradation process itself in our model is a monotone stochastic process, like the gamma process for example. Again, however, it would be a shock, superimposed on cumulative damage, that causes failure.

Acknowledgements Mei-Ling T. Lee is supported in part by NIH Grant R01 AI121259. We thank research colleagues in the fields of osteoporosis, cystic fibrosis and chronic obstructive pulmonary disease for making us aware of the important role that physical traumas and acute exacerbations play in initiating critical medical events in their fields. Given our previous awareness of the importance of shock processes in causing failure in engineering systems, it was not difficult for us to see the shock-degradation process as a general failure mechanism and to anticipate its natural extension to social and economic systems as well.

References

1. Aalen OO, Borgan O, Gjessing HK (2008) Survival and event history analysis: a process point of view. Statistics for biology and health. Springer, New York
2. Aaron SD, Stephenson AL, Cameron DW, Whitmore GA (2015) A statistical model to predict one-year risk of death in patients with cystic fibrosis. J Clin Epidemiol 68(11):1336–1345. doi:10.1016/j.jclinepi.2014.12.010

3. Gumbel EJ (1953) Probability table for the analysis of extreme-value data: introduction. Applied mathematics series, vol 22. National Bureau of Standards, United States Department of Commerce, Washington, DC, pp 1–15
4. He X, Whitmore GA, Loo GY, Hochberg MC, Lee M-LT (2015) A model for time to fracture with a shock stream superimposed on progressive degradation: the Study of Osteoporotic Fractures. Stat Med 34(4):652–663. doi:10.1002/sim.6356
5. Lee M-LT, Whitmore GA (2006) Threshold regression for survival analysis: modeling event times by a stochastic process reaching a boundary. Stat Sci 21:501–513
6. Lee M-LT, Whitmore GA (2010) Proportional hazards and threshold regression: their theoretical and practical connections. Lifetime Data Anal 16:196–214
7. Lee M-LT, Whitmore GA (2016) Shock-degradation failure processes and their survival distributions, manuscript under review
8. Lee M-LT, Whitmore GA, Rosner BA (2010) Threshold regression for survival data with time-varying covariates. Stat Med 29:896–905
9. Mamelund S-E, Brunborg H, Noack T. Divorce in Norway 1886–1995 by calendar year and marriage cohort. Technical report 97/19. Statistics Norway

Chapter 12
Statistical Methods for Thermal Index Estimation Based on Accelerated Destructive Degradation Test Data

Yimeng Xie, Zhongnan Jin, Yili Hong, and Jennifer H. Van Mullekom

Abstract Accelerated destructive degradation test (ADDT) is a technique that is commonly used by industries to access material's long-term properties. In many applications, the accelerating variable is temperature. In such cases, a thermal index (TI) is used to indicate the strength of the material. For example, a TI of 200 °C may be interpreted as the material can be expected to maintain a specific property at a temperature of 200 °C for 100,000 h. A material with a higher TI possesses a stronger resistance to thermal damage. In literature, there are three methods available to estimate the TI based on ADDT data, which are the traditional method based on the least-squares approach, the parametric method, and the semiparametric method. In this chapter, we provide a comprehensive review of the three methods and illustrate how the TI can be estimated based on different models. We also conduct comprehensive simulation studies to show the properties of different methods. We provide thorough discussions on the pros and cons of each method. The comparisons and discussion in this chapter can be useful for practitioners and future industrial standards.

Keywords Degradation path • Long-term properties • Material durability • Material reliability • Monotonic splines • Semiparametric methods

12.1 Introduction

12.1.1 Background

Polymeric materials are common in various industrial applications. In current industrial practice, a thermal index (TI) is often used to rate the long-term performance of polymeric materials. As specified in industrial standard UL 746B [1], the TI of a

Y. Xie • Z. Jin • Y. Hong (✉) • J.H. Van Mullekom
Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
e-mail: xym@vt.edu; jinxx354@vt.edu; yilihong@vt.edu; vanmuljh@vt.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_12

polymeric material can be considered as a measure of the material's ability to retain a specific property (e.g., physical or electrical properties) under exposure to elevated temperatures over a prolonged period of time (e.g., 100,000 h). The interpretation of the TI is as follows. A material with a TI value of 200 °C is expected to maintain the specific property for exposure to a temperature of 200 °C for 100,000 h. Thus, a material with a higher TI rating is expected to demonstrate a stronger resistance to thermal exposure as compared to those with a lower TI rating. The TI can also be used to determine suitability for a particular application, and for comparing multiple materials. When a material is introduced to a field, its TI can be compared to a list of similar materials with known TI values, which can give insights for the long term performance of the new material. Therefore, estimating the TI for a material is an important task in evaluating material performance.

To estimate the TI, data that track the material property over time need to be collected. Such data are referred to as degradation data. However, the degradation of the material performance is often gradual and can take years to observe deterioration. To collect information in a timely manner, accelerated degradation test (ADTs) are often used. In the setting of TI estimation, temperature is the accelerating variable. When measuring the material performance, such as the tensile strength, the sample will be stretched until it breaks. Because the sample is destroyed in the testing procedure, only one measurement can be collected from one sample. Such type of ADT is called an accelerated destructive degradation test (ADDT). Due to the nature of the testing, ADDT is a commonly used technique for evaluating long-term performance of polymeric materials. Examples of ADDT data include the Adhesive Bond B data in [2], the Polymer data in [3], the Seal Strength data in [4], and the Formulation K data in [5].

To use the ADDT data for the TI estimation, a statistical method is needed. In the literature, there are three methods available to estimate the TI based on ADDT data, which are the traditional approach based on the least-squares method, the parametric approach based on maximum likelihood (ML) method, and the semiparametric approach based on splines method. The traditional procedure is the one that is currently specified in the industrial standards UL 746B [1], which is commonly used to evaluate material in applications. The traditional approach is a two-step approach using polynomial fittings and least-squares methods. In the statistical literature, the parametric method is also commonly used to model the ADDT data, and the ML method is used for parameter estimation. Recently, a semiparametric method was proposed to analyze ADDT data in Xie et al. [5]. The basic idea of the semiparametric method is to use monotonic splines to model the baseline degradation path and use a parametric method to model the effect of accelerating variable.

The objective of this chapter is to provide a comprehensive review of the three methods and illustrate how the TI can be estimated based on different methods. We also conduct comprehensive simulation studies to show the properties of different methods. Then, we provide thorough discussions on the pros and cons of each method. The comparisons and discussions in this chapter can be useful for practitioners and future industrial standards.

12.1.2 Related Literature

Degradation data were used to access products and material reliability in early work such as Nelson [6, Chapter 11], and Lu and Meeker [7]. There are two types of degradation data: repeated measures degradation test (RMDT) data and ADDT data. For RMDT data, multiple measurements can be taken from the same unit. For ADDT data, only one measurement can be taken from the same unit, due to the destructive nature of the measuring procedure. Different types of methods are used to analyze RMDT and ADDT data. The majority of the degradation literature is on RMDT data analysis, features two major classes of models: the general path model (e.g., [8] and [9]) and stochastic process models (e.g., [10, 11], and [12]). A review of statistical degradation models and methods are available in Meeker et al. [13], and Ye and Xie [14].

This chapter focuses on the analysis of ADDT data and their corresponding TI estimation procedures. Regarding ADDT analysis, the traditional approach for TI estimation using the least-squares method is described in UL 746B [1]. Parametric models are quite common in ADDT analysis, for example, in Escobar et al. [2], Tsai et al. [3], and Li and Doganaksoy [4]. King et al. [15] applied both the traditional and parametric approaches to ADDT data analysis and TI estimations. King et al. [15] also did comprehensive comparisons for the two approaches in TI estimations. Xie et al. [5] developed a semiparametric approach for ADDT data analysis, in which the monotonic splines are used to model the baseline degradation path and the Arrhenius relationship is used to describe the temperature effect. However, the TI estimation procedure was not developed in [5].

In this chapter, we develop the TI estimation based on the semiparametric method after providing a review of the existing methods in TI estimations. We also conduct comprehensive simulations to compare the three methods. In terms of software implementation, Hong et al. [16] implements the three methods and their corresponding TI estimation procedures into an R package “ADDT”. Details and illustrations of the R package ADDT is available in Jin et al. [17].

12.1.3 Overview

The rest of this chapter is organized as follows. Section 12.2 introduces the concept of ADDT, examples of ADDT data, and the concept of TI. Section 12.3 presents the three different methods that can be used to model ADDT data and their corresponding procedures for TI estimation. The three different methods are the traditional method, the parametric method, and the semiparametric method. Section 12.5 conducts extensive simulations to compare the performance of the estimation procedures. Section 12.6 provides a comprehensive discussion on the pros and cons of each method and suggestions for practitioners.

12.2 Accelerated Tests and Thermal Index

In this section, we give a more detailed introduction to ADDT and TI.

12.2.1 Test Plans

The test plan of an ADDT consists of the temperature levels, the measuring time points, and the number of samples allocated to each combination of the temperature levels and measuring time points. Table 12.1 illustrates a test plan for an ADDT. Four elevated temperature levels are considered in the test, which are 250 °C, 260 °C, 270 °C, and 280 °C. There are five measuring time points considered in this plan, which are 552, 1008, 2016, 3528, and 5040 h. At the initial time (time zero), there are ten sample units tested under the normal temperature level to serve as the baseline. For each combination of temperature level and time points, there are five sample units tested to obtain the measurements for the material property. To measure some properties like tensile strength, the unit will be destroyed after the measurement. Note that equal sample allocation is used in Table 12.1. However, unequal sample size allocation is also seen in practice. See King et al. [15] for more detailed discussion on the test plans.

12.2.2 Data and Notation

The ADDT data record the material property (e.g., the tensile strength of the material) for each unit. Here, we use the Adhesive Bond B example in Escobar et al. [2] to illustrate the ADDT data. Figure 12.1 shows a scatter plot of the Adhesive Bond B data. In general, we observe that there is a decreasing trend over time, while for higher temperature level, the rate of decreasing is faster than those under lower temperature levels.

Here we introduce some notations to the ADDT data that will be necessary for the development of the statistical methods. Let n be the number of temperature levels and n_i be the number of measuring time points for temperature level i. The value of

Table 12.1 Illustration of sample size allocation for an ADDT

Temperature	Measuring points (hours)					
(°C)	0	552	1008	2016	3528	5040
–	10					
250		5	5	5	5	5
260		5	5	5	5	5
270		5	5	5	5	5
280		5	5	5	5	5

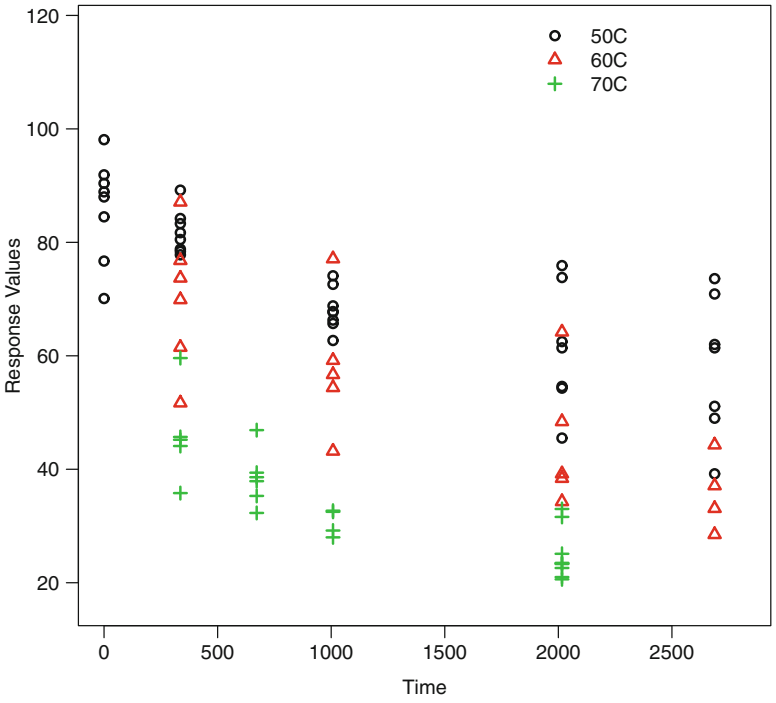

Fig. 12.1 Scatter plot of the Adhesive Bond B data. The x-axis is time in hours and the y-axis is strength in Newtons

the ith temperature level is denoted by A_i. The corresponding time points are denoted by $t_{ij}, j = 1, \cdots, n_i$. Note that it is possible that the measure time points are different from different temperature levels. Let n_{ij} be the number samples tested at time t_{ij} for temperature level i. Note that the number of samples tested at each time point t_{ij} can also vary. We denote the degradation measurement by y_{ijk} for the kth sample at level i of the temperature level i and measuring time t_{ij}, $i = 1, \cdots, n, j = 1, \cdots, n_i$, and $k = 1, \cdots, n_{ij}$. The total number of measured samples are $N = \sum_{i=1}^{n} \sum_{j=1}^{n_i} n_{ij}$.

12.2.3 Thermal Index

In this section, we introduce the general concept of the thermal index (TI). In the following, we will use the tensile strength as the interested material property. In a common framework of degradation modeling, the failure time is defined as the first time when the degradation level passes the failure threshold. For example, a failure is said to have occurred when the tensile strength of a sample reaches a certain percentage (e.g., 50%) of the original tensile strength.

For degradation processes that are accelerated by temperature, the Arrhenius relationship is widely used to model the relationship between the degradation and temperature. In particular, the Arrhenius model uses the following transformed temperature,

$$h(A) = \frac{-11605}{A + 273.16}, \tag{12.1}$$

where A is the temperature value in degrees Celsius, the constant 11605 is the reciprocal of the Boltzmann's constant (in units of eV). Note that the constant 273.16 is for converting the Celsius temperature scale to the Kelvin temperature scale. For the convenience of modeling, we define,

$$x = \frac{1}{A + 273.16}, \quad \text{and} \quad x_i = \frac{1}{A_i + 273.16}.$$

Through the modeling of the degradation data, which will be detailed in Sect. 12.3, the mean time to failure at x can be described by a relationship $m(x)$. For targeted time to failure t_d (e.g., $t_d = 100{,}000$ h), the corresponding temperature level R can be obtained by solving x_d from $m(x_d) = t_d$. Because

$$x_d = m^{-1}(t_d) = \frac{1}{R + 273.16},$$

we obtain the corresponding temperature value R as

$$R = \frac{1}{m^{-1}(t_d)} - 273.16. \tag{12.2}$$

The temperature level R in (12.2) is defined as the TI for the material. Figure 12.2 illustrates the temperature-time relationship based on the Arrhenius relationship and the corresponding TI.

Note that the targeted time to failure is not required to be fixed at 100,000 h. For example, if there is an existing material with a known TI (e.g., 220 °C), its targeted time to failure t_d^{old} can be obtained. For a new material, its TI can be obtained by using t_d^{old} as the targeted time. In this case, the TI for the new material is called the relative TI because it compares to an existing material; see King et al. [15] for more details.

12.3 Statistical Methods for Thermal Index Estimations

This section covers the statistical methods for the TI estimation. We first review the traditional and the parametric methods as described in King et al. [15]. Then, we derive the TI estimation based on the semiparametric model in Xie et al. [5].

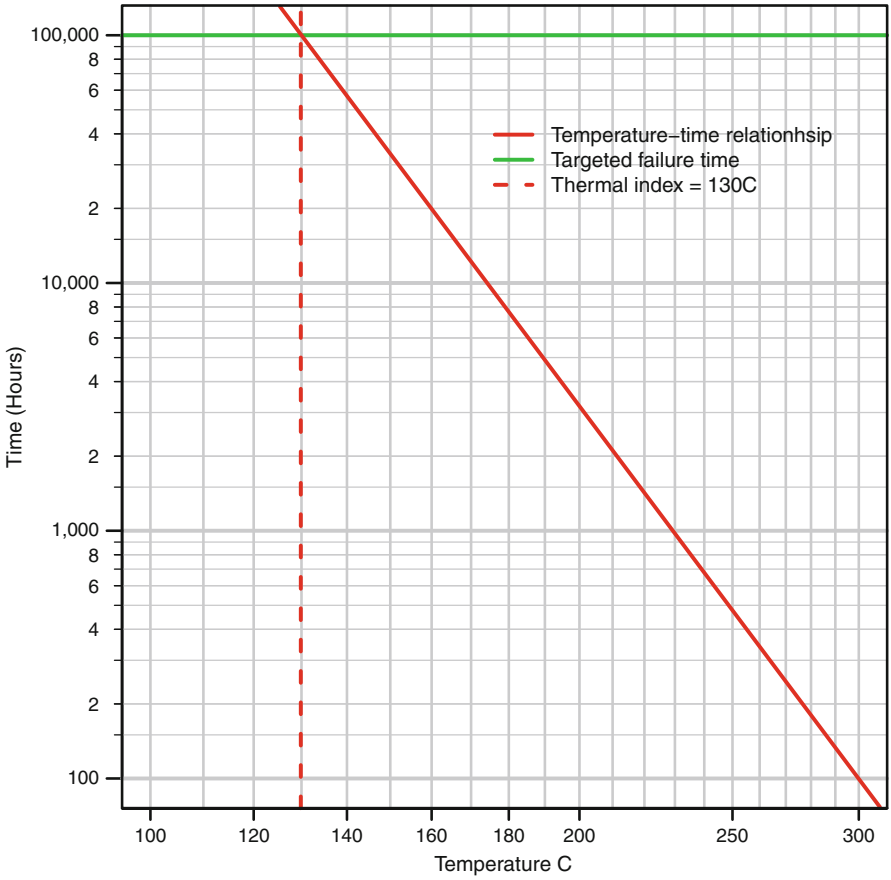

Fig. 12.2 Illustration of temperature-time relationship and TI. The x-axis is temperature A on the scale of $1/(A + 273.16)$, and the y-axis is time in hours on base 10 logarithm scale

12.3.1 The Traditional Method

The traditional method is the methodology that is described in UL 746B [1], which is the currently accepted standard for ADDT data analysis in industry. The procedure essentially is a two-step approach. The basic idea is to find an appropriate model to link the time to failure to the level of degradation, and then estimate the parameters by applying the least-squares technique. The estimated failure time is done by interpolating the fitted curves. If there is no material-specific knowledge on the degradation relationship, the UL standards recommend using a third-order polynomial fitting.

Specifically, for temperature level i, one first computes the points $\{t_{ij}, \bar{y}_{ij\cdot}\}$, $j = 1, \cdots, n_i$, where

$$\bar{y}_{ij\cdot} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$$

is the average of the batch of observations at time t_{ij} for temperature level i. A third order polynomial $a_{0i} + a_{1i}t + a_{2i}t^2 + a_{3i}t^3$ is used to fit the data points $\{t_{ij}, \bar{y}_{ij\cdot}\}$, $j = 1, \cdots, n_i$, separately for each temperature level. Here, $(a_{0i}, a_{1i}, a_{2i}, a_{3i})'$ are the polynomial coefficients to be estimated by the least-squares method.

After obtaining the estimates of $(a_{0i}, a_{1i}, a_{2i}, a_{3i})'$, the mean failure time m_i for temperature level i can be obtained through interpolation. In particular, one needs to solve,

$$a_{0i} + a_{1i}m_i + a_{2i}m_i^2 + a_{3i}m_i^3 = y_f,$$

where y_f is the failure threshold. The failure threshold is usually set at 50% of the initial strength, though different values may be set according to the specification of different applications.

Through the polynomial interpolation, a set of data points $\{x_i, m_i\}, i = 1, \cdots, n$ are obtained where x_i is the transformed temperature as defined in (12.1). The least-squares method is used again to fit a straight line to data points $\{x_i, \log_{10}(m_i)\}, i = 1, \cdots, n$. That is, to fit the following model,

$$\log_{10}(m_i) = \beta_0 + \beta_1 x_i, i = 1, \cdots, n,$$

to obtain the estimates of β_0 and β_1. Note that the base 10 logarithm is used here because it is more popular in engineering literature. In the traditional method the temperature-time relationship is represented as

$$\log_{10}[m(x)] = \beta_0 + \beta_1 x. \tag{12.3}$$

With the fitted temperature-time relationship in (12.3), the TI based on the traditional method is obtained as

$$R = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16\,.$$

where t_d is the target time and $t_d = 100{,}000$ is often used.

The traditional method is fairly intuitive and straightforward to compute, which is an advantage. Here we provide some other considerations for the traditional method. The interpolation based method requires the degradation level to reach the failure threshold so that m_i can be obtained for level i. Otherwise, all data collected at level i can not be used for analysis. The number of temperature levels for the ADDT is usually small (i.e., around 4). Thus only a few number of observations are available to fit the model in (12.3). Furthermore, the two-step approach detailed in the current

standard does not specify a method to quantify the statistical uncertainty associated with the TI estimation. For ADDT data, one would expect higher temperature levels to yield shorter lifetimes. Due to randomness in the data and the flexibility of polynomials, the traditional method can produce estimated failure times that are not monotonically increasing with temperature, which would usually be unrealistic. With parametric models, most of the concerns are avoided.

12.3.2 The Parametric Method

In statistical literature, parametric methods are prevalent in ADDT data analysis such as in [2, 3], and [4]. In the parametric method, the primary method for estimation and inference is based on a parametric model and maximum likelihood theory. Here, we give a brief description for the parametric method summarized in King et al. [15].

In this setting, the parametric model for the degradation measurement is represented as

$$y_{ijk} = \mu(t_{ij}; x_i) + \epsilon_{ijk}, \tag{12.4}$$

where $\mu(t_{ij}; x_i)$ is the underlying degradation path and ϵ_{ijk} is an error term. Since the tensile strength is decreasing over time, the function $\mu(t; x_i)$ is specified as a decreasing function of t. Consequently, a higher temperature usually lead to a higher rate of degradation. The function $\mu(t; x_i)$ is also a decreasing function of the temperature.

For a specific x, the mean time to failure $m(x)$ can be solved from $\mu[m(x); x] = y_f$, leading to the temperature-time relationship as $m(x) = \mu^{-1}(y_f; x)$. The TI can be solved from $m(x_d) = t_d$, which is equivalent to solving for x_d in $\mu(t_d; x_d) = y_f$. The TI can be computed from the solution x_d. That is,

$$R = \frac{1}{x_d} - 273.16.$$

To proceed with the modeling, one needs to be specific about the form of $\mu(t, x)$. For polymer materials, the parametric form in [18] is often used. In particular,

$$\mu(t; x) = \frac{\alpha}{1 + \left[\dfrac{t}{\eta(x)}\right]^{\gamma}}, \tag{12.5}$$

where α is the initial degradation level, $\eta(x) = \exp(\nu_0 + \nu_1 x)$ is the scale factor based on the Arrhenius model, and γ is the shape parameter determining the steepness of the degradation path.

Let $p = y_f/\alpha$ be the proportion of decreasing for the failure threshold from the initial degradation level. Based on the model in (12.5), the mean time to failure at

x, $m(x)$, is obtained by solving $\mu[m(x); x] = p\alpha$. Specifically, the temperature-time relationship is

$$\log_{10}[m(x)] = \beta_0 + \beta_1 x,$$

where

$$\beta_0 = \frac{\nu_0}{\log(10)} + \frac{1}{\gamma \log(10)} \log\left[\frac{1-p}{p}\right], \quad \text{and} \quad \beta_1 = \frac{\nu_1}{\log(10)}.$$

When $p = 1/2$, β_0 reduces to $\nu_0/\log(10)$. The TI at t_d can be computed as

$$R = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16. \tag{12.6}$$

The model in (12.4) is estimated by the ML method. The error term is modeled as

$$\varepsilon_{ijk} \sim \mathrm{N}(0, \sigma^2), \quad \text{and} \quad \mathrm{Corr}(\varepsilon_{ijk}, \varepsilon_{ijk'}) = \rho, \ k \neq k'. \tag{12.7}$$

The parameter ρ represents the within-batch correlation. The unknown parameters is denoted by $\boldsymbol{\theta} = (\nu_0, \nu_1, \alpha, \gamma, \sigma, \rho)'$. The likelihood is

$$L(\boldsymbol{\theta}) = \prod_{i,j} (2\pi)^{-\frac{n_{ij}}{2}} |\Sigma_{ij}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[\boldsymbol{y}_{ij} - \boldsymbol{\mu}(t_{ij}, x_i)]' \Sigma_{ij}^{-1} [\boldsymbol{y}_{ij} - \boldsymbol{\mu}(t_{ij}, x_i)]\right\}, \tag{12.8}$$

where $\boldsymbol{y}_{ij} = (y_{ij1}, \cdots, y_{ijn_{ij}})'$ is the corresponding vector of degradation measurements which follows the multivariate normal distribution with mean vector $\boldsymbol{\mu}(t_{ij}; x_i)$, an $n_{ij} \times 1$ vector of $\boldsymbol{\mu}(t_{ij}; x_i)$'s, and covariance matrix Σ_{ij}, an $n_{ij} \times n_{ij}$ matrix with σ^2 on the diagonal entries and $\rho\sigma^2$ on the off-diagonal entries. The parameter estimates $\widehat{\boldsymbol{\theta}}$ are obtained by maximizing (12.8). The estimate of R is obtained by evaluating (12.6) at the estimate $\widehat{\boldsymbol{\theta}}$.

The parametric model can overcome the shortcoming of the traditional method, and allows for statistical inference. However, for the parametric method, one needs to find an appropriate form for $\mu(t_{ij}; x_i)$.

12.3.3 The Semiparametric Method

Xie et al. [5] proposed the following semi-parametric functional form for $\mu(t_{ij}; x_i)$ for the model in (12.4). That is,

$$\mu(t_{ij}; x_i) = g\left[\eta_i(t_{ij}; \beta); \boldsymbol{\gamma}\right], \tag{12.9}$$

$$\eta_i(t;\beta) = \frac{t}{\exp(\beta s_i)}, \quad s_i = x_i - x_{\max}.$$

Here, $g(\cdot)$ is a monotonic decreasing function with parameter vector $\boldsymbol{\gamma}$, and β is the parameter for the temperature effect. The quantity $x_{\max} = 1/(\max_i\{A_i\} + 273.16)$ is the transformed value of the highest level of temperature. At the highest temperature level, $s_{\max} = x_{\max} - x_{\max} = 0$, then $\mu(t; x_{\max}) = g(t;\boldsymbol{\gamma})$. Thus, the function $g(\cdot)$ is interpreted as the baseline degradation path. The advantage of using the maximum temperature level as the baseline is that its degradation level will reach the failure threshold in most ADDTs. The $g(\cdot)$ is constructed nonparametrically by monotonic splines, which is the nonparametric component of the model. The use of the monotonic splines retains the physical meaning of the degradation mechanism (i.e., monotonicity), and it is also flexible because one does not need to find a parametric form for the degradation paths. The Arrhenius model is used for describing the acceleration effect, which is the parametric component of the model. Thus, the model in (12.9) is called a semiparametric model.

The distribution of the error terms ε_{ijk} is specified in (12.7). Let $\boldsymbol{\theta} = (\boldsymbol{\gamma}', \beta, \sigma, \rho)'$ be the vector containing all of the unknown parameters. The estimation of $\boldsymbol{\theta}$ is through an iterative procedure that maximizes the loglikelihood function. The details of monotonic spline construction and parameter estimation are given in Xie et al. [5].

Here we derive the TI estimation based on the semiparametric model in (12.9). Let $g_0 = g(0)$ be the initial degradation level and p be the proportion reducing from the initial degradation (i.e., $p = y_f/g_0$). The mean time to failure for the temperature level x is denoted by $m(x)$, which can be solved from

$$g\left[\frac{m(x)}{\exp[\beta(x - x_{\max})]}\right] = pg_0.$$

We obtain the temperature-time relationship as, $m(x) = g^{-1}(pg_0)\exp[\beta(x - x_{\max})]$, which is equivalent to

$$\log_{10}[m(x)] = \beta_0 + \beta_1 x.$$

Here,

$$\beta_0 = \log_{10}[g^{-1}(pg_0)] - \frac{\beta x_{\max}}{\log(10)}, \quad \text{and} \quad \beta_1 = \frac{\beta}{\log(10)}.$$

The TI is computed as,

$$R = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16. \tag{12.10}$$

The estimates of the TI R can be obtained by substituting the estimate of $\boldsymbol{\theta}$ into (12.10).

12.4 An Illustration of Thermal Index Estimation

In this section, we provide an illustration of TI estimation using the Adhesive Bond B data in Escobar et al. [2]. The computing was done by using the R package ADDT by Hong et al. [9].

12.4.1 Degradation Path Modeling

We apply the traditional method, the parametric method, and the semiparametric method to the Adhesive Bond B data. For the traditional method, Fig. 12.3 shows the polynomial interpolation for the Adhesive Bond B data, when the failure threshold is set to $p = 50\%$. For the temperature level 50 °C, the degradation level has not reached the failure threshold yet. The estimated time to failure m_{50} is not available. Thus, data from this level is discarded from the analysis. In contrast, all data can be used in the parametric and semiparametric methods.

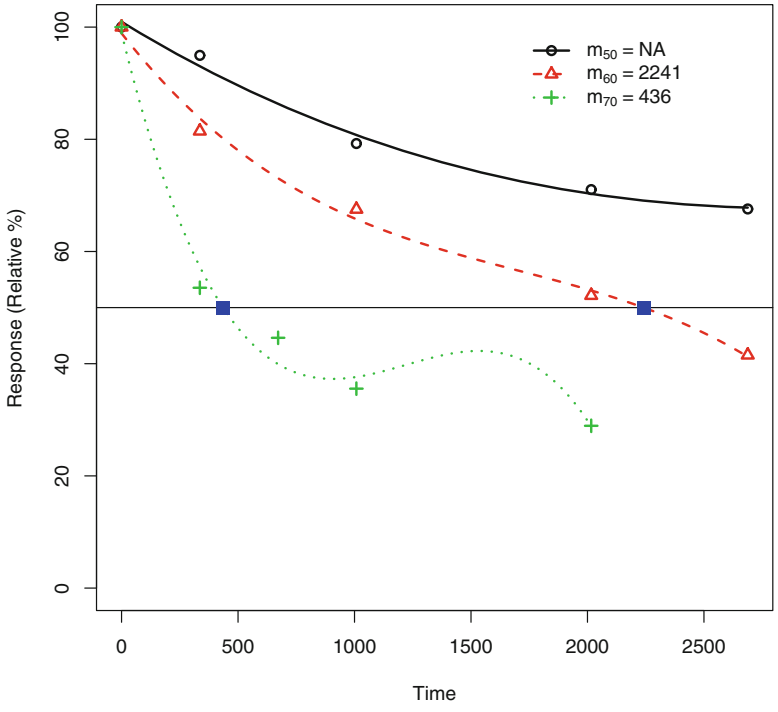

Fig. 12.3 Polynomial interpolation for the traditional method for the Adhesive Bond B data. The failure threshold is $p = 50\%$

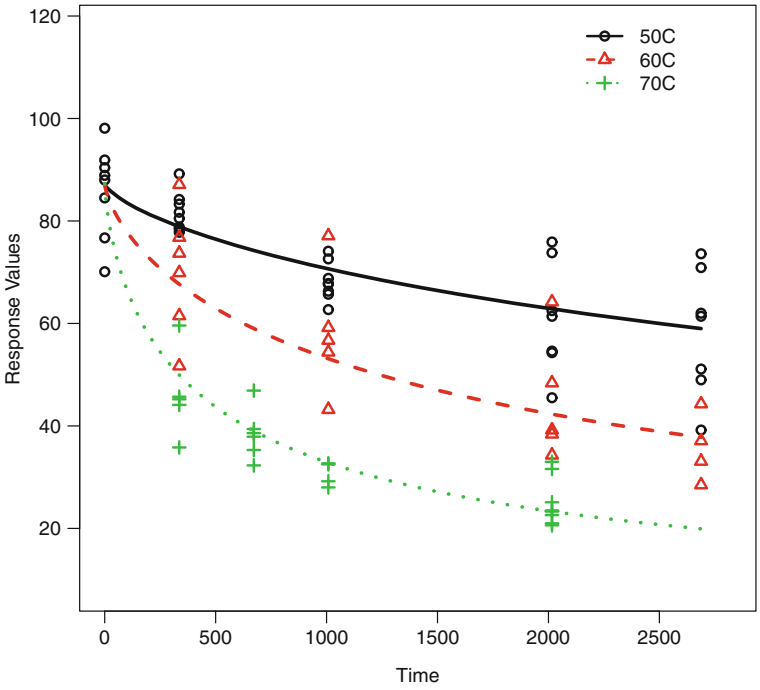

Fig. 12.4 Fitted degradation paths using the parametric method for the Adhesive Bond B data. The x-axis is time in hours and the y-axis is strength in Newtons

Figure 12.4 shows the fitted degradation paths using the parametric method for the Adhesive Bond B data, while Fig. 12.5 shows similar results based on the semiparametric method. Both methods provide good fits to the data. The results in Xie et al. [5] show that the semiparametric method tends to have a better fit to the degradation data.

12.4.2 TI Estimation

For illustrations, we compute the TI based on the three methods presented previously. Table 12.2 shows the estimated parameters for the temperature-time relationship, and the corresponding TI for the Adhesive Bond B data. In the computing, we use $t_d = 100{,}000$ and $p = 50\%$. Figure 12.6 shows the fitted temperature-time relationship lines using the three methods and the corresponding estimated TI for the Adhesive Bond B data. The results based on the parametric method and semiparametric method are quite close to each other, while the results from traditional method is different from these two methods. Section 12.5 will conduct a simulation study to evaluate the estimation performance.

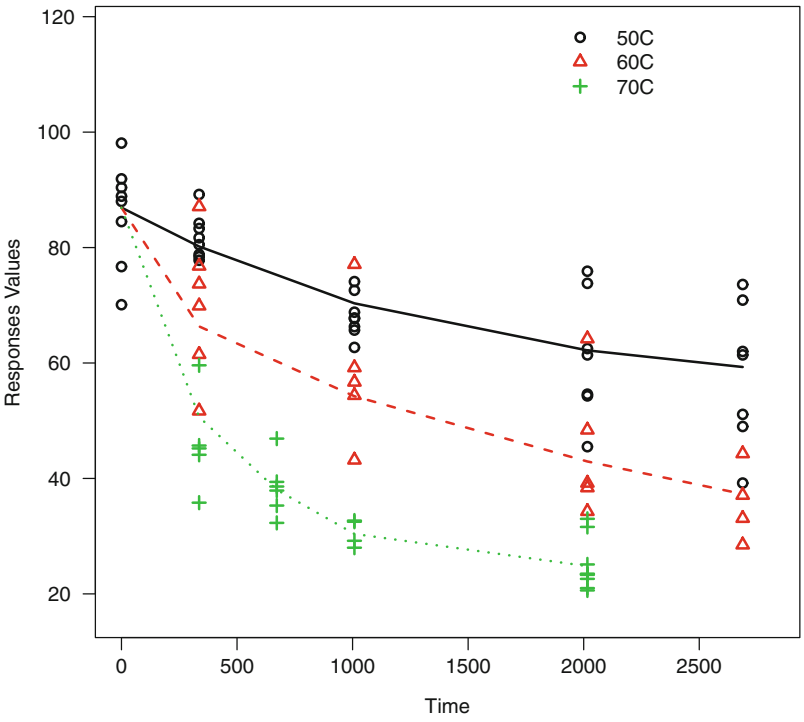

Fig. 12.5 Fitted degradation paths using the semiparametric method for the Adhesive Bond B data. The x-axis is time in hours and the y-axis is strength in Newtons

Table 12.2 Estimated parameters for the temperature-time relationship and TI based on the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for the Adhesive Bond B data, when $t_d = 100{,}000$ and $p = 50\%$

Methods	β_0	β_1	TI
TM	−21.05	8128.4	39
PM	−16.18	6480.4	33
SPM	−16.81	6697.1	34

12.5 Simulation Studies

In this section, simulations are carried out to compare the performance of the traditional method, the parametric method, and the semiparametric method in terms of estimating the TI. We will consider two settings, under which the parametric model is correctly specified, or the parametric model is incorrectly specified.

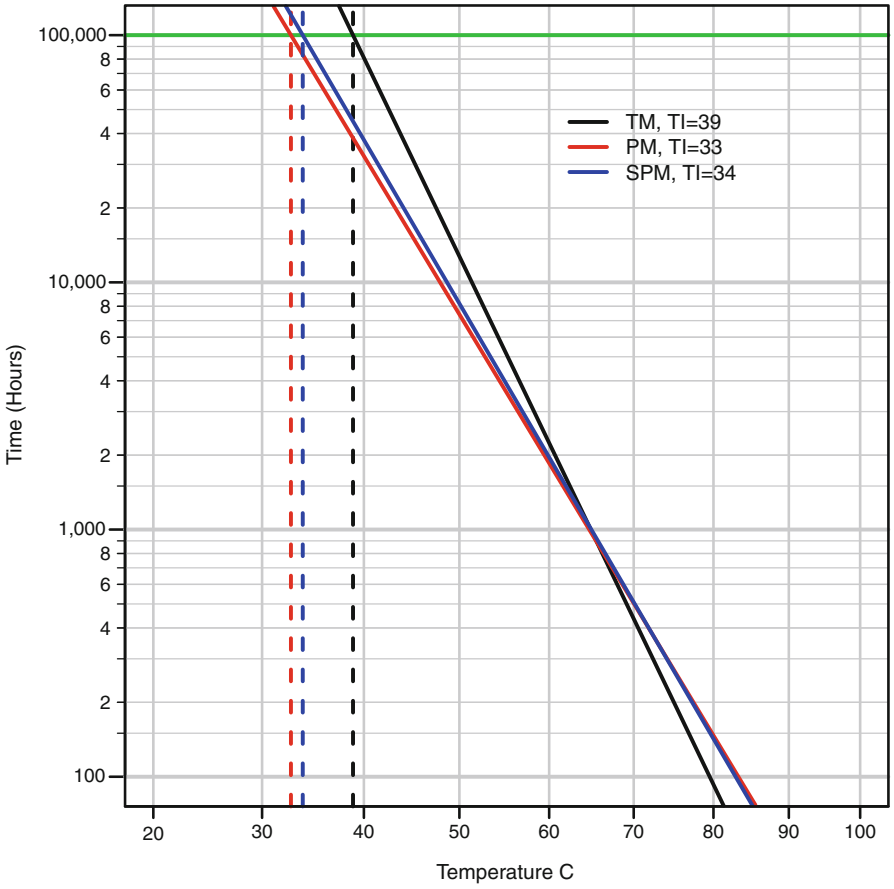

Fig. 12.6 Fitted temperature-time relationship lines using the traditional method (TM), the parametric method (PM) and the semiparametric method (SPM), and the corresponding estimated TI for the Adhesive Bond B data

12.5.1 Simulation Settings

For the first setting (Setting I), we generate degradation data from the parametric model in (12.5), and for the parametric method, we still use the same model in (12.5) to fit the data, which is corresponding to the case that the model is correctly specified. For model (12.5), the parameter values used in the simulation are $\alpha = 9000, \nu_0 = -16, \nu_1 = 12500, \gamma = 2, \sigma = 1000$, and $\rho = 0$. The failure threshold was set to be $p = 50\%$ of the initial degradation level and we used $t_d = 100,000$ h. Under this configuration, the true TI is $R = 181\,^{\circ}\text{C}$. The correlation in Setting I is $\rho = 0$, to speed up the simulations.

Table 12.3 The temperature levels and measuring time points for the eight simulation scenarios

Scenarios	Temperature levels (°1)					Time points (hours)				
1: Temp. 3, Time 4		250	260	270		552	1008	2016	3528	
2: Temp. 4, Time 4		250	260	270	280	552	1008	2016	3528	
3: Temp. 4, Time 4	240	250	260	270		552	1008	2016	3528	
4: Temp. 5, Time 4	240	250	260	270	280	552	1008	2016	3528	
5: Temp. 3, Time 5		250	260	270		552	1008	2016	3528	5040
6: Temp. 4, Time 5		250	260	270	280	552	1008	2016	3528	5040
7: Temp. 4, Time 5	240	250	260	270		552	1008	2016	3528	5040
8: Temp. 5, Time 5	240	250	260	270	280	552	1008	2016	3528	5040

For Setting II, we examine model misspecification by generating degradation data from a parametric model that is different from (12.5), but we fit the model in (12.5) for the parametric method. In particular, the following model was used to generate data for Setting II,

$$\mu(t;x) = \alpha \exp\left\{-\left[\frac{t}{\eta(x)}\right]\right\}, \tag{12.11}$$

which was used in Li and Doganaksoy [4] to describe the degradation of polymer strength. Here, $\eta(x) = \exp(\nu_0 + \nu_1 x)$. For model (12.11), the parameter values were set to $\alpha = 9000$, $\nu_0 = -15.6$, $\nu_1 = 12471$, $\sigma = 1000$, and $\rho = 0$. Those values were chosen to match the mean time to failure under 270 °C and the true TI $R = 181$ °C to Setting I so that the results from both settings are comparable.

For each setting, eight scenarios were considered. For each scenario, we vary the number of time points and the temperature levels. Table 12.3 lists the configuration for each scenario. We considered both four time points and five time points to check the sensitivity to time constraints. The number of temperature levels is from three to five to check the sensitivity to temperature factors. We also considered the range of temperature levels with either higher or lower temperature levels to check the effect of temperature level in terms of distance from use levels. A similar simulation study design was used in King et al. [15].

12.5.2 Results Under the Correct Model

Table 12.4 shows the estimated mean, bias, standard deviation (SD), and root of mean squared error (RMSE) of the TI estimators for the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for Setting I: the parametric model is correctly specified. Figure 12.7 visualizes the results in Table 12.4. We observe that those scenarios with more time points and temperature levels tend to have better precision in estimating TI for all methods. Testing at higher

Table 12.4 Estimated mean, bias, SD, and RMSE of the TI estimators for the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for Setting I: the parametric model is correctly specified

Scenarios	True TI	Mean			Bias			SD			RMSE		
		TM	PM	SPM	TM	PM	SPM	TM	PM	SPM	TM	PM	SPM
1: Temp. 3, Time 4	181	170	179	179	11	2	2	14	9	9	18	9	9
2: Temp. 4, Time 4	181	178	180	181	3	1	1	8	6	6	8	6	6
3: Temp. 4, Time 4	181	171	181	181	11	0	0	13	5	6	17	5	6
4: Temp. 5, Time 4	181	178	181	181	4	0	0	8	4	4	9	4	4
5: Temp. 3, Time 5	181	179	179	179	2	2	2	9	9	9	10	9	9
6: Temp. 4, Time 5	181	182	180	181	1	1	0	5	5	6	5	5	6
7: Temp. 4, Time 5	181	177	180	181	4	1	1	6	5	5	7	5	5
8: Temp. 5, Time 5	181	180	181	182	1	1	0	4	4	4	4	4	4

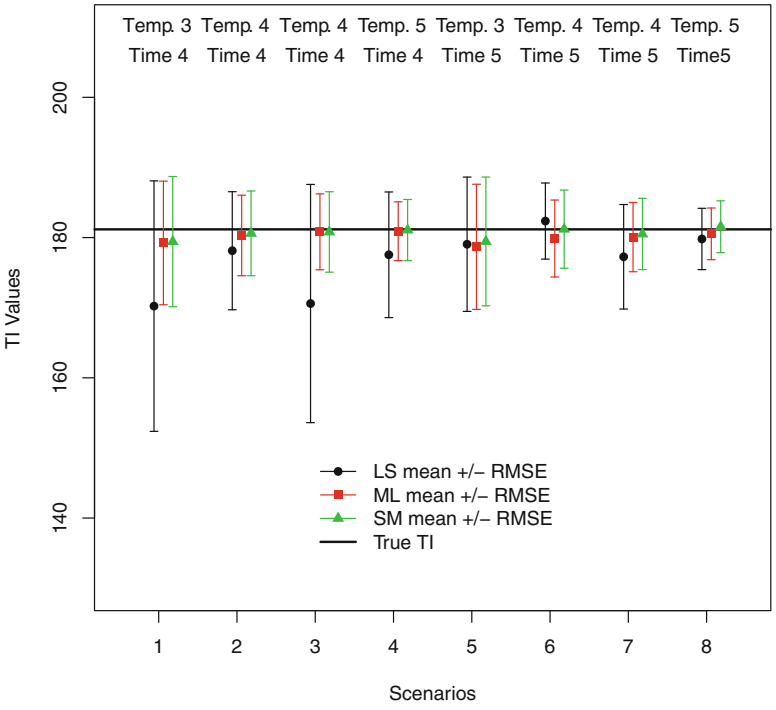

Fig. 12.7 Plot of the estimated mean, bias, SD, and RMSE of the TI estimators for the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for Setting I: the parametric model is correctly specified

Table 12.5 Estimated mean, bias, SD, and RMSE of the TI estimators for the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for Setting II: the parametric model is incorrectly specified

Scenarios	True TI	Mean			Bias			SD			RMSE		
		TM	PM	SPM	TM	PM	SPM	TM	PM	SPM	TM	PM	SPM
1: Temp. 3, Time 4	181	178	180	179	3	1	2	16	11	12	17	11	12
2: Temp. 4, Time 4	181	179	180	180	1	1	1	10	7	8	10	8	8
3: Temp. 4, Time 4	181	176	180	179	4	0	1	17	7	8	17	7	8
4: Temp. 5, Time 4	181	178	180	180	2	0	1	10	5	6	10	5	6
5: Temp. 3, Time 5	181	179	178	178	2	2	3	12	10	11	12	10	12
6: Temp. 4, Time 5	181	178	179	180	3	2	1	8	7	7	9	7	7
7: Temp. 4, Time 5	181	179	181	180	2	0	1	8	6	6	8	6	6
8: Temp. 5, Time 5	181	178	180	180	3	0	0	6	4	5	6	4	5

temperature levels tends to provide better precision for all the methods. Among the three methods, the traditional method tends to perform worse than the other two methods. This observation for the traditional method is consistent with the findings in King et al. [15]. The performance of the newly added semiparametric is comparable to the parametric method.

12.5.3 Results Under a Misspecified Model

Table 12.5 shows the estimated mean, bias, SD, and RMSE of the TI estimators for the traditional method, the parametric method, and the semiparametric method for Setting II: the parametric model is incorrectly specified. Figure 12.8 visualizes the results in Table 12.5. We observe similar patterns to Setting I. That is, those scenarios with more time points and temperature levels tend to have better precision in estimating TI for all methods, and the traditional method tends to perform worse than the other two methods, which is also consistent with the findings in King et al. [15]. Surprisingly the parametric method performs well even under model misspecification. Similarly, the performance of the newly added semiparametric method is comparable to the parametric method.

12.6 Discussions

In literature, there are three methods available to estimate the TI based on ADDT data, which are the traditional method, the parametric method, and the semiparametric method. In this chapter, we provide a comprehensive review of the three methods and illustrate how the TI can be estimated based on different models. We also conduct a simulation study to show the properties of different methods. The

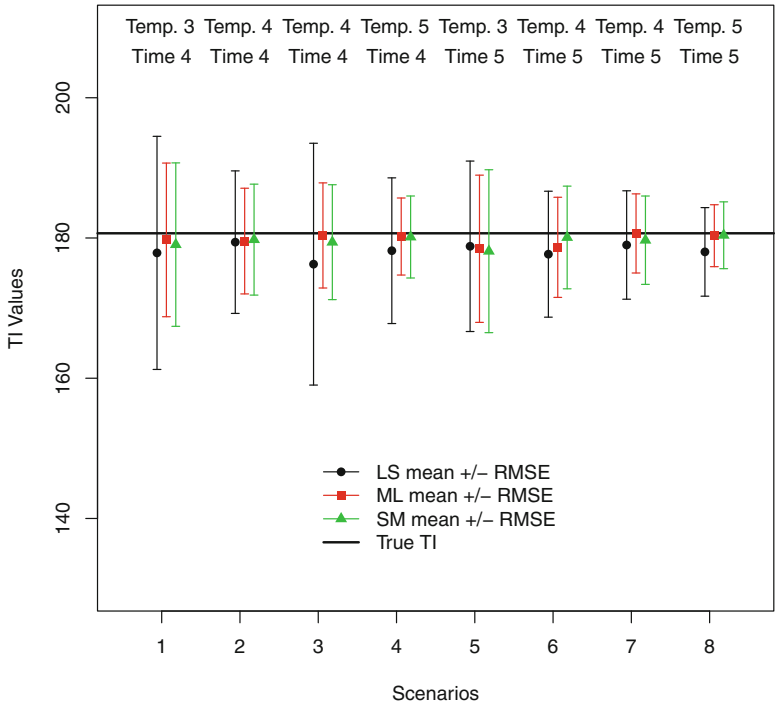

Fig. 12.8 Plot of the estimated mean, bias, SD, and RMSE of the TI estimators for the traditional method (TM), the parametric method (PM), and the semiparametric method (SPM) for Setting II: the parametric model is incorrectly specified

comparisons and discussion in this chapter can be useful for practitioners and future industrial standards.

Here, we provide a summary on the pros and cons of each method.

- Regarding estimation performance, if there are fewer temperature levels or number of time points, the traditional method tends to not performance well. When there are five temperature levels and five time points, the traditional method works well. Both the parametric and semiparametric methods perform better than the traditional methods and their performance are comparable to each other.
- Regarding model assumptions, the traditional method does not require specific forms for the underlying degradation path because it uses polynomial interpolation. The semiparametric method does not require a specific form but assumes that the underlying path is monotone and smooth. The parametric method assumes a specific form, which requires the strongest assumption. However, the simulation study shows that the parametric model used here is flexible to some extent under model misspecification.

- Regarding data use, both the parametric and semiparametric methods use all of the data for analyses, including those have not yet reached the failure threshold. The traditional method will discard the data from the temperature which has not reached the failure threshold yet.
- Both the parametric and semiparametric methods can quantify the uncertainties in the estimation (see King et al. [15], and Xie et al. [5] for details). Because the traditional method requires two steps to estimate the TI, it is challenging to quantify the statistical uncertainties.
- The semiparametric method is the most computationally intensive method, and the parametric method is in the middle in term of computational time. All the three methods is implemented in the R package ADDT. The chapter in [17] gives a detailed illustration for the use of the package.

In summary, it is of advantages to use the parametric and semiparametric methods in the ADDT analysis and TI estimation. In practice, one can compare the model fitting of both the parametric and semiparametric methods (e.g., AIC values) to determine which models can provide a better description to the ADDT data. The practitioner should also weigh the pros and cons discussed in this section in conjunction with the minimum AIC model for final model selection. Details of model comparisons can be found in Xie et al. [5].

Acknowledgements The authors thank William Q. Meeker for his helpful comments on earlier version of the paper. The authors acknowledge Advanced Research Computing at Virginia Tech for providing computational resources. The work by Hong was partially supported by the National Science Foundation under Grant CMMI-1634867 to Virginia Tech.

References

1. UL746B (2013) Polymeric materials – long term property evaluations, UL 746B. Underwriters Laboratories, Incorporated
2. Escobar LA, Meeker WQ, Kugler DL, Kramer LL (2003) Accelerated destructive degradation tests: data, models, and analysis. In: Lindqvist BH, Doksum KA (eds) Mathematical and statistical methods in reliability, chap. 21. World Scientific Publishing Company, River Edge
3. Tsai C-C, Tseng S-T, Balakrishnan N, Lin C-T (2013) Optimal design for accelerated destructive degradation tests. Qual Technol Quant Manag 10:263–276
4. Li M, Doganaksoy N (2014) Batch variability in accelerated-degradation testing. J Qual Technol 46:171–180
5. Xie Y, King CB, Hong Y, Yang Q (2017, in press) Semi-parametric models for accelerated destructive degradation test data analysis. Technometrics, doi:10.1080/00401706.2017.1321584
6. Nelson WB (1990) Accelerated testing: statistical models, test plans, and data analysis. Hoboken, Wiley
7. Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 34:161–174
8. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. John Wiley & Sons, Hoboken
9. Hong Y, Duan Y, Meeker WQ, Stanley DL, Gu X (2015) Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. Technometrics 57:180–193

10. Whitmore GA (1995) Estimation degradation by a Wiener diffusion process subject to measurement error. Lifetime Data Anal 1:307–319
11. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Anal 11:511–527
12. Wang X, Xu D (2010) An inverse Gaussian process model for degradation data. Technometrics 52:188–197
13. Meeker WQ, Hong Y, Escobar LA (2011) Degradation models and data analyses. In: Encyclopedia of statistical sciences. Wiley, Hoboken
14. Ye Z, Xie M (2015) Stochastic modelling and analysis of degradation for highly reliable products. Appl Stoch Model Bus Ind 31:16–32
15. King CB, Xie Y, Hong Y, Van Mullekom JH, DeHart SP, DeFeo PA (2017, in press) A comparison of traditional and maximum likelihood approaches to estimating thermal indices for polymeric materials. J Qual Technol
16. Hong Y, Xie Y, Jin Z, King C (2016) ADDT: a package for analysis of accelerated destructive degradation test data. R package version 1.1. [Online]. Available: http://CRAN.R-project.org/package=ADDT
17. Jin Z, Xie Y, Hong Y, Van Mullekom JH (2017) ADDT: an R package for analysis of accelerated destructive degradation test data. In: Chen DG, Lio YL, Ng HKT, Tsai TR (eds) Statistical modeling for degradation data, chap. 14. Springer, New York
18. Vaca-Trigo I, Meeker WQ (2009) A statistical model for linking field and laboratory exposure results for a model coating. In: Martin J, Ryntz RA, Chin J, Dickie RA (eds) Service life prediction of polymeric materials, chap. 2. Springer, New York

Chapter 13
Inference on Remaining Useful Life Under Gamma Degradation Models with Random Effects

Man Ho Ling, Hon Keung Tony Ng, and Kwok-Leung Tsui

Abstract Prognostics and system health management becomes an important topic in modern reliability study. In prognostics and system health management, remaining useful life is one of the vital indexes to yield an advance warning of impending failure in a system, thereby helping in executing preventive actions prior to failure occurrence and helping in making maintenance decisions. To obtain precise statistical inference on the remaining useful life, we consider degradation models that incorporate unit-specific random effects that model heterogeneity in the degradation of distinct systems, and propose a parametric bootstrap confidence interval for the remaining useful life of each system. A Monte Carlo simulation study is carried out to evaluate the performance of the proposed methodology. To illustrate the suggested model and inferential methods, a real data set of light intensity of light emitting diodes is analyzed.

Keywords Bootstrap • Degradation models • Gamma process • Random effects • Remaining useful life • Maximum likelihood estimation

M.H. Ling
Department of Mathematics and Information Technology, The Education University of Hong Kong, Tai Po, Hong Kong SAR, China
e-mail: amhling@eduhk.hk

H.K.T. Ng (✉)
Department of Statistical Sciences, Southern Methodist University, Dallas, TX, USA
e-mail: ngh@mail.smu.edu

K.L. Tsui
Department of Systems Engineering and Engineering Management, The City University of Hong Kong, Kowloon Tong, Hong Kong, China
e-mail: kltsui@cityu.edu.hk

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_13

13.1 Introduction

Nowadays, as many powerful sensors and signal processing techniques appear, prognostics and system health management (PHM) becomes an important topic in reliability study. PHM is a systematic approach for failure prevention by monitoring the health/status of products and systems, predicting failure progression, yielding an advance warning of impending failure in a system, and mitigating operating risks through repair or replacement [14], thereby helping in making maintenance decisions and executing preventive actions prior to failure occurrence to extend system life. Therefore, many applications of PHM can be found in a variety of fields including electronics, smart grid, nuclear plant, power industry, aerospace and military application, fleet industrial maintenance, and public health management [14].

Remaining useful life (RUL), defined as the length from current time to the end of the useful life, is one of the vital indexes in PHM. RUL has been commonly used in reliability studies with practical applications. In reliability studies, multiple degradation features can be extracted for degradation detection and quantification. While the application of health/status monitoring is established, degradation data that describe quality characteristics over time are measured, one can estimate the RUL for making a timely maintenance decision for failure avoidance. As RUL estimation is one of the critical assessments in PHM and condition-based maintenance, it is of interest in developing efficient statistical inferential procedures for the RUL. Si et al. [15] provided a comprehensive review on RUL estimation based on different statistical models. Gebraeel and Pan [9] discussed a Bayesian approach for computing and updating the estimate of RUL of bearings. Recently, Fan et al. [8] developed the physics-of-failure based damage models to predict the RUL of light emitting diodes (LEDs). However, unit-to-unit variation has not been taken into account in these existing studies. Therefore, we aim to propose a model that incorporates unit-specific random effects in this paper.

Since time-to-failure can be viewed as the first-passage time of passing a specified threshold in a degradation process, there exists a comprehensive literature on statistical inference based on degradation models and their applications. For instance, Singpurwalla [16] used degradation models with covariates for modeling the product lifetime in a dynamic environment. Ebrahimi [6] proposed degradation models for fatigue cracks in a system. Wang [17, 18] considered the gamma degradation models for degradation data of bridge beams and carbon-film resistors. Ling et al. [11] considered accelerated degradation models for LEDs. The first-passage time of a Wiener process is known to follow an inverse Gaussian distribution; see Chhikara and Folks [4]. On the other hand, the first-passage time of a gamma process can be approximated by using a two-parameter Birnbaum-Saunders distribution; see Park and Padgett [13]. The estimation of reliability function and mean time-to-failure (MTTF) have been discussed in the literature [1]. However, these lifetime characteristics may not take the information from the current degradation data into account. In contrast, the RUL uses the current degradation data for making a timely maintenance decision for failure avoidance, which can provide an advance warning of impending failure in a system.

For degradation data analysis, three of the popular degradation models, Wiener, inverse Gaussian, and gamma processes, are often used in engineering literature. Due to some attractive properties of the normal distribution, Wiener process has been commonly used to model the degradation process of a product. Wiener process can be used to represent the performance of a product that increases and decreases over time. Readers may refer to articles [5, 12, 19, 20] for details on modeling degradation data using a Wiener process. However, Wiener process is useful for some specific datasets, but it may not be appropriate for modeling item with monotonic degradation. In such cases, it is more reasonable to consider the inverse Gaussian process or the gamma process for modeling the degradation data. Ye and Chen [21] pointed out that the inverse Gaussian process is flexible in incorporating random effects and covariates that account for heterogeneities in degradation data. In this paper, we aim to show that the gamma process degradation model is also flexible in incorporating random effects for the statistical inference of RUL.

This article is organized as follows. In Sect. 13.2, we describe the gamma degradation model with random effects. In Sect. 13.3, we derive the unit-specific RUL based on the proposed gamma degradation model and the corresponding expected value and variance. Then, in Sect. 13.4, we discuss the statistical inference of the RUL and propose a bootstrap procedure to construct confidence interval for unit-specific RUL. A Monte Carlo simulation study is used to evaluate the performance of the proposed methodology in Sect. 13.5. In Sect. 13.6, an illustrative example based on a real degradation data on LEDs is used to demonstrate the statistical inferential procedures studied in this paper. Finally, some concluding remarks are provided in Sect. 13.7.

13.2 Gamma Degradation Model with Random Effects

Gamma degradation models with random effects are motivated by a degradation data of light intensity of LEDs [3]. Twelve LEDs were run at electrical current of 35 mA and the outputs of the light intensity of the LEDs were measured at 5 inspection times. Ling et al. [11] presented a gamma degradation model with time-scale transformation to analyze the degradation data. In degradation analysis, there can be a substantial heterogeneity among degradations of different items. It is therefore of great interest to consider a degradation model with unit-specific random effects that model heterogeneity in degradation of distinct systems. The data obtained from [11] are presented in Fig. 13.1.

Suppose that n items are placed under a multiple levels of constant stress experiment and these items are inspected at times $0 = t_0 < t_1 < t_2 < \cdots < t_m$. Let $y_{i,j} = y_i(t_j)$ represent the measured degradation of the i-th item and inspected at time t_j. Since the quality of a system is quantified in terms of the percentage to the initial value, and so $y_{i,0} \equiv 1$ for $i = 1, 2, \ldots, n$. Moreover, due to the fact that the degradation for each item is bounded between 0 and 1, we can consider the logarithmic (log-) transformation of the degradation measures. We further assume that an increment in the logarithm of the degradation measure,

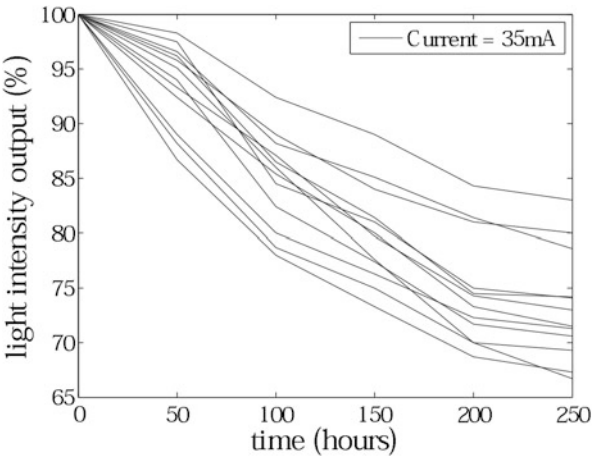

Fig. 13.1 Degradations of light intensity of 12 LEDs at electrical current of 35 mA

$g_{i,j} = \log(y_{i,j-1}) - \log(y_{i,j})$ has a gamma distribution with shape parameter $\alpha_j = \lambda(t_j - t_{j-1}) > 0$, and scale parameter $z > 0$. To incorporate the unit-specific random effects, under the setting presented in [10], we assume that z has an inverse gamma distribution with shape parameter $\delta > 2$ and scale parameter $\gamma > 0$.

Based on this model, we have the corresponding conditional density function of g given z and the marginal density function of z as

$$f_{g|z}(g|z) = \frac{g^{\alpha_j - 1}}{\Gamma(\alpha_j) z^{\alpha_j}} \exp\left(-\frac{g}{z}\right), \qquad g > 0, \tag{13.1}$$

and

$$f_z(z) = \frac{z^{-\delta-1}\gamma^{\delta}}{\Gamma(\delta)} \exp\left(-\frac{\gamma}{z}\right), \qquad z > 0, \tag{13.2}$$

respectively. Lawless and Crowder [10] showed that the marginal density function $f(g)$ is then given by

$$f_g(g) = \int_0^{\infty} f_{g|z}(g|z) f_z(z) dz = \frac{\gamma^{\delta}}{B(\alpha_j, \delta)} \frac{g^{\alpha_j - 1}}{(g+\gamma)^{\alpha_j + \delta}}, \qquad g > 0, \tag{13.3}$$

where $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ is the beta function. They also found that $w_1 = g\delta/(\alpha_j/\gamma)$ has a F distribution with degrees of freedom $d_1 = 2\alpha_j$ and $d_2 = 2\delta$ and $w_2 = g/(g+\gamma)$ has a beta distribution with parameters α_j and δ.

As λ, δ and γ are positive, we can define $\lambda = \exp(a), \delta = \exp(b)$ and $\gamma = \exp(c)$. Under this setting, the log-likelihood function based on the degradation data, $y_{i,j}, i = 1, 2, \ldots, n, j = 0, 1, \ldots, m$, can be expressed as

$$\ell(\boldsymbol{\theta}) = \sum \delta \log(\gamma) - \log(B(\alpha_j, \delta)) + (\alpha_j - 1)\log(g_{i,j}) - (\alpha_j + \delta)\log(g_{i,j} + \gamma), \tag{13.4}$$

where $\sum = \sum_{i=1}^{n}\sum_{j=0}^{m}$ and $\boldsymbol{\theta} = (a, b, c)$. The maximum likelihood estimates (MLEs) of the parameter vector $\boldsymbol{\theta} = (a, b, c)$ can be obtained by maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. Iterative numerical methods, such as the Newton-Raphson method, are required to obtain the MLEs of the model parameters. In our study, the nonlinear programming solver *fminsearch* in MATLAB is used to obtain the MLEs of the model parameters. The MATLAB code for obtaining the MLEs of $\boldsymbol{\theta}$ based on the degradation data can be obtained from the authors upon request.

13.3 Remaining Useful Life

Suppose D_i is the time-to-failure of the i-th system at which the degradation crosses a pre-specified threshold, τ. By using the property that the sum of gamma random variables with the same scale parameter is again gamma distributed, we readily obtain $G(d|z_i) = -\log(y(d|z_i)) \sim Ga(\lambda d, z_i)$. Park and Padgett [13] derived the exact distribution of D_i based on the gamma process, and also provided an approximation to $E[D_i]$ by using a two-parameter Birnbaum-Saunders distribution; see Birnbaum and Saunders [2]. As RUL is defined as the length from the current time, t_*, to the time-to-failure based on the current degradation $y(t_*)$, $\Delta G = \log(y(t_*)) - \log(y(D))$ is of interest in this study. It follows that, conditional on z_i, ΔG has a gamma distribution with shape parameter $\lambda x > 0$ and scale parameter $z_i > 0$, where $x = d - t_*$. Furthermore, as $y(d) = \tau$, we have the conditional probability that the degradation crosses the threshold before time x, given the current degradation $y_i(t_*)$, as follows:

$$\begin{aligned}\Pr(X_i < x|g_*, z_i) &= \Pr(\Delta G > g_*|x, z_i)\\ &\cong 1 - \Phi\left(\frac{g_* - \lambda x z_i}{\sqrt{\lambda x z_i}}\right)\\ &= \Phi\left[\frac{1}{p_i}\left(\sqrt{\frac{x}{q_i}} - \sqrt{\frac{q_i}{x}}\right)\right],\end{aligned} \tag{13.5}$$

where $g_* = \log(y(t_*)) - \log(\tau)$, $p_i = \sqrt{z_i/g_*}$ and $q_i = g_*/(z_i\lambda)$. This result enables us to develop statistical inference on the RUL for an individual item. In addition, conditional on z_i, the remaining useful life $X_i = D_i - t_*$ follows a Birnbaum-Saunders distribution, $BS(p_i, q_i)$, approximately. It follows that, given g_*, the RUL can be estimated by its expected value as

$$\phi_i = E[X_i] = \frac{g_*\delta}{\lambda\gamma} + \frac{1}{2\lambda}. \tag{13.6}$$

In addition, conditional on z_i, the remaining useful life $X_i = D_i - t_*$ has a Birnbaum-Saunders distribution, $BS(p_i, q_i)$, where $p_i = \sqrt{z_i/g_*}, q_i = g_*/(z_i\lambda)$

and $g_* = \log(y(t_*)) - \log(\tau)$. The variance of the remaining useful life is then given by

$$V[X_i] = \frac{g_*^2 \delta}{\lambda^2 \gamma^2} + \frac{g_* \delta}{\lambda^2 \gamma} + \frac{5}{4\lambda^2}. \tag{13.7}$$

The derivations of the above expressions of the expected value and variance are presented in the Appendix.

13.4 Statistical Inference on Remaining Useful Life

Given $\theta = (a, b, c)$, the $100(1-\beta)\%$ asymptotic confidence interval for the remaining useful life for the i-th system, ϕ_i, is then given by

$$\left(\hat{\phi}_i - z_{\beta/2}\widehat{se}(\phi_i), \hat{\phi}_+ z_{\beta/2}\widehat{se}(\phi_i)\right), \tag{13.8}$$

where $z_{\beta/2}$ is the $1-\beta/2$ quantile of the standard normal distribution and $\widehat{se}(\phi_i) = \sqrt{\hat{V}[X_i]}$ is the standard error of ϕ_i.

As θ is unknown, θ can be estimated based on the degradation data. However, the sampling distribution of the estimate of ϕ_i is intractable. Here, we propose using a parametric bootstrap procedure [7] to approximate the sampling distribution of the estimate of ϕ_i and to obtain percentile bootstrap confidence interval of the RUL. An approximate $100(1-\beta)\%$ bootstrap confidence interval for the RUL of the i-th system, ϕ_i, can be constructed as follows:

1. Find the MLE of the vector of model parameters, $\hat{\theta}$, from the original degradation data.
2. Obtain the r-th bootstrap sample $\{\tilde{g}_{i,j}^{(r)}, i = 1, 2, \ldots, n, j = 1, 2, \ldots, m\}$ based on $\hat{\theta}$ and the inspection times $\{t_j, j = 1, 2, \ldots, m\}$.
3. Find the r-th bootstrap estimate of θ based on the r-th bootstrap sample, denoted by $\tilde{\theta}^{(r)}$.
4. For $k = 1, 2, \ldots, 99$, compute the r-th sequence of 99 bootstrap estimates of ϕ_i, denoted by $\phi_{i,k}^{(r)} = \hat{\phi}_i - z_{k/100}\widehat{se}(\phi_i)$, based on $\tilde{\theta}^{(r)}$.
5. Repeat steps 2 to 4 B times to approximate the distribution of the estimate of ϕ_i.
6. Discard all the bootstrap estimates of $\phi_{i,k}^{(r)} \leq 0$.
7. Sort the remaining bootstrap estimates $\phi_{i,k}^{(r)}$ in ascending order, denoted by $\phi_i^{[b]}, b = 1, 2, \ldots, M$, where $M \leq 99B$.

Consequently, a $100(1-\beta)\%$ parametric bootstrap confidence interval for the RUL of the i-th system, ϕ_i is given by

$$\left(\phi_i^{[M\beta/2]}, \phi_i^{[M(1-\beta/2)]}\right). \tag{13.9}$$

13.5 Monto Carlo Simulation Study

To evaluate the performance of the proposed methodology, a Monte Carlo simulation study is carried out to estimate the biases and root mean square errors (RMSE) of the point estimates of the model parameters and the RUL as well as the coverage probabilities (CP) and average widths (AW) of 95% confidence intervals (CI) of the model parameters and the RUL based on $n = 10$ (small sample size), $n = 30$ (moderate sample size), and $n = 100$ (large sample size). Suppose the experiment was terminated at time $t_m = 250$, and the degradation of each specimen was measured at inspection times $t_j = t_m(j/m)$, where $j = 0, 1, \ldots, m$. Consider the inspection frequency $m = 5$ and 10, and let $a = -3$ and $(b, c) = (3, 0)$ (low variability of random effect), $(b, c) = (1.5, 2)$ (moderate variability of random effect), and $(b, c) = (0.8, -3)$ (high variability of random effect). Given the threshold $\tau = 0.5$, the procedure to generate the degradation data of the i-th system, $y_{i,j}$, at the inspection time t_j and the failure time, d_i, are described as follows:

1. Given $\theta = (a, b, c)$, compute $\lambda = \exp(a), \delta = \exp(b)$ and $\gamma = \exp(c)$.
2. Generate z_i from a gamma distribution with scale δ and shape γ.
3. Generate $g_{i,s}$ from a gamma distribution with scale z_i and shape λ, for $s = 1, 2, \ldots, 5000$.
4. Compute $y_{i,j} = y_i(t_j) = \exp\left(-\sum_{s=1}^{t_j} g_{i,s}\right)$.
5. Obtain $d_i = \underset{s=\{1,2,\ldots,5000\}}{\text{argmin}} \; y_i(s) < \tau$.

The results based on 1,000 Monte Carlo simulations are summarized in Tables 13.1, 13.2, 13.3, 13.4, 13.5, and 13.6.

The simulation results show that the performance of the percentile parametric bootstrap confidence interval is satisfactory for estimation of the remaining useful when the sample size is sufficiently large ($n \geq 30$), as the coverage probabilities

Table 13.1 Biases and root mean square errors (RMSE) of the MLEs of the model parameters and the remaining useful life with high variability of random effect ($V[Z] = 7.30 \times 10^{-3}$) for different sample sizes n and different numbers of inspections m

	$a = -3.0, b = 0.8, c = -3.0$					
	Bias			RMSE		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.878	0.059	0.012	3.578	0.717	0.180
b	1.802	0.239	0.045	4.854	1.058	0.231
c	0.974	0.210	0.031	6.386	1.477	0.447
ϕ	−85.250	−99.490	−100.622	298.463	304.979	306.597
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.021	−0.001	−0.002	0.260	0.144	0.075
b	1.358	0.184	0.049	4.033	0.501	0.211
c	1.385	0.213	0.060	4.212	0.695	0.320
ϕ	−88.411	−99.510	−101.018	297.670	301.941	302.632

Table 13.2 Biases and root mean square errors (RMSE) of the MLEs of the model parameters and the remaining useful life with moderate variability of random effect ($V[Z] = 6.08 \times 10^{-4}$) for different sample sizes n and different numbers of inspections m

	$a = -3.0, b = 1.5, c = -2.0$					
	Bias			RMSE		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.414	0.028	−0.001	2.283	0.263	0.130
b	3.246	0.635	0.101	6.898	2.478	0.341
c	2.851	0.612	0.108	7.504	2.633	0.486
ϕ	−40.118	−45.447	−46.323	145.656	145.260	143.997
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.028	−0.001	0.000	0.211	0.113	0.065
b	3.146	0.526	0.071	6.608	2.165	0.308
c	3.135	0.539	0.075	6.749	2.243	0.400
ϕ	−41.369	−45.648	−47.370	141.578	143.704	144.733

Table 13.3 Biases and root mean square errors (RMSE) of the MLEs of the model parameters and the remaining useful life with low variability of random effect ($V[Z] = 1.51 \times 10^{-4}$) for different sample sizes n and different inspection numbers of inspections m

	$a = -3.0, b = 3, c = 0.0$					
	Bias			RMSE		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.247	0.040	0.014	1.435	0.181	0.092
b	3.884	2.700	1.017	7.708	6.173	3.685
c	3.609	2.654	1.001	7.959	6.249	3.730
ϕ	−4.600	−6.839	−9.277	38.196	38.343	38.832
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.038	0.016	0.006	0.159	0.088	0.050
b	3.216	2.237	0.936	6.920	5.668	3.382
c	3.160	2.218	0.928	6.962	5.708	3.413
ϕ	−4.769	−7.039	−7.461	38.333	38.558	38.504

are close to the nominal level of 95%, regardless of variability of random effect. However, the proposed bootstrap method does not provide satisfactory confidence intervals for the model parameters, especially in the case of low variability of random effect ($b = 3.0, c = 0.0$), in terms of the coverage probability. When the variability of random effect is low, the log-likelihood function of c is often a flat curve. Figure 13.2 shows that the MLE of c is far away from the true value of the parameter $c = 0$. This explains why the biases, root mean square errors and coverage probabilities of the parameters b and c are not satisfactory when the variability of random effect is low (Tables 13.3 and 13.6). However, from Eq. (6), $\delta/\gamma = \exp(b - c)$, the poor performance of the estimation of parameters b and c does not influence the performance of the proposed interval estimation method for the RUL.

Table 13.4 Coverage probabilities (CP) and average widths (AW) of the 95% confidence intervals for the model parameters and the remaining useful life with high variability of random effect ($V[Z] = 7.30 \times 10^{-3}$) for different sample sizes n and different numbers of inspections m

	$a = -3.0, b = 0.8, c = -3.0$					
	CP			AW		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.902	0.954	0.948	10.750	3.210	0.728
b	0.776	0.865	0.919	15.290	4.648	0.917
c	0.862	0.903	0.936	25.942	8.091	1.789
ϕ	0.915	0.929	0.932	952.069	860.381	839.207
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.951	0.935	0.950	1.504	0.543	0.297
b	0.759	0.860	0.932	14.658	3.501	0.865
c	0.835	0.875	0.935	16.320	4.284	1.316
ϕ	0.914	0.927	0.933	928.946	844.421	820.014

Table 13.5 Coverage probabilities (CP) and average widths (AW) of the 95% confidence intervals for the model parameters and the remaining useful life with moderate variability of random effect ($V[Z] = 6.08 \times 10^{-4}$) for different sample sizes n and different numbers of inspections m

	$a = -3.0, b = 1.5, c = -2.0$					
	CP			AW		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.939	0.961	0.953	6.960	1.487	0.529
b	0.798	0.880	0.908	13.856	10.492	1.772
c	0.787	0.926	0.930	20.534	12.060	2.390
ϕ	0.923	0.931	0.935	476.330	449.425	430.299
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.941	0.950	0.940	0.893	0.447	0.248
b	0.791	0.879	0.904	13.492	9.305	1.423
c	0.794	0.895	0.914	14.290	9.868	1.766
ϕ	0.916	0.929	0.933	470.197	440.612	432.184

13.6 Illustrative Example: LED Degradation Data

In order to illustrate the proposed model and methodology in this paper, we perform a degradation analysis on the data of light intensity of 12 LEDs. The data are presented in Fig. 13.1 under the proposed heterogeneity model. A failure is defined as the light intensity degrades below a threshold of 50% of the initial value. Table 13.7 presents the estimates of the model parameters and the unit-specific remaining useful life estimates for each LED, along with the corresponding 95% bootstrap confidence intervals. Figure 13.3 displays the 95% confidence intervals for the remaining useful life of each LED. The variance of the random effect is

Table 13.6 Coverage probabilities (CP) and average widths (AW) of the 95% confidence intervals for the model parameters and the remaining useful life with low variability of random effect ($V[Z] = 1.51 \times 10^{-4}$) for different sample sizes n and different numbers of inspections m

	$a = -3.0, b = 3.0, c = 0.0$					
	CP			AW		
$m = 5$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.894	0.904	0.915	3.384	0.646	0.300
b	0.666	0.766	0.887	5.865	4.920	3.776
c	0.300	0.296	0.329	8.178	4.975	3.759
ϕ	0.954	0.954	0.951	173.784	164.338	161.861
$m = 10$	$n = 10$	$n = 30$	$n = 100$	$n = 10$	$n = 30$	$n = 100$
a	0.922	0.934	0.918	0.575	0.317	0.173
b	0.724	0.799	0.884	3.794	3.394	2.878
c	0.188	0.208	0.263	3.502	3.231	2.803
ϕ	0.954	0.954	0.953	172.322	164.144	161.970

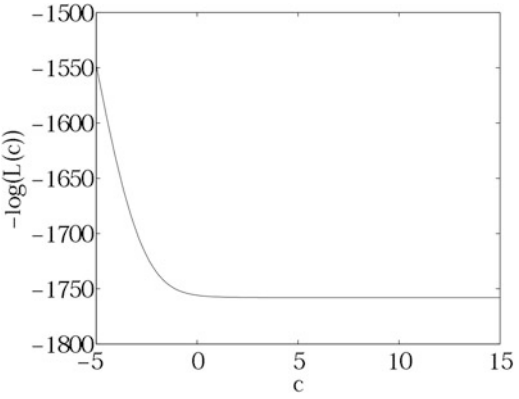

Fig. 13.2 Log-likelihood function with respect to the variable c

$$V[Z] = \frac{\gamma^2}{(\delta-1)^2(\delta-2)} = \frac{\exp(2c)}{(\exp(b)-1)^2(\exp(b)-2)} \approx 1.57 \times 10^{-5},$$

which indicates that the random effect could be neglected and the degradation model with random effect seems to perform slightly better in the remaining useful life estimation.

13.7 Concluding Remarks

In this paper, we proposed a heterogeneity gamma degradation model that incorporates unit-specific random effects in the gamma process. A percentile parametric bootstrap method is suggested for interval estimation for the unit-specific remaining

Table 13.7 Estimates of the model parameters and the remaining useful life for 12 LEDs, along with the corresponding 95% parametric bootstrap confidence interval

Model parameters		
a	b	c (1×10^{-5})
−2.97	3.72	−4.12
(−3.24, −2.57)	(3.39, 4.18)	(−4.52, −3.98)
Remaining useful life		
LED #1	LED #2	LED #3
245.20	318.92	423.77
(93.08, 407.44)	(139.52, 511.66)	(206.78, 658.14)
LED #4	LED #5	LED #6
379.28	393.7	291.61
(177.98, 596.02)	(187.35, 616.15)	(122.02, 473.32)
LED #7	LED #8	LED #9
276.44	252.52	299.67
(112.37, 451.76)	(97.57, 417.88)	(127.21, 484.59)
LED #10	LED #11	LED #12
332.23	331.13	301.96
(147.90, 530.37)	(147.20, 528.83)	(128.66, 487.79)

useful life. The proposed model and methodology are illustrated by the analysis of the degradation data of 12 LEDs. In the illustrative example, we showed that the random effect could be neglected and the degradation model with random effect seems to perform slightly better in the remaining useful life estimation. Moreover, the results from a Monte Carlo simulation study show that the percentile parametric bootstrap technique successfully yields accurate estimates of the remaining useful life. For future research direction, it will be interesting to develop models and estimation methods for individual effect based on degradation data. We are currently working on this problem and hope to report the findings in a future paper.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Project No. 11471275), Research Grants Council Theme-Based Research Scheme (Grant T32-101/15-R) and the Research Grants Council Early Career Scheme Fund (Ref. 28300114). The work of H.K.T. Ng was supported by a grant from the Simons Foundation (#280601 to Tony Ng). Thanks are also due to the reviewer for the valuable comments and suggestions on an earlier version of this manuscript which led to this improved version. Our sincere thanks also go to the editors of this volume for an invitation to us which provided an impetus for preparing this article.

Appendix

Suppose that, conditional on z, X has a Birnbaum-Saunders distribution, $BS(p, q)$, where $p = \sqrt{z/g_*}, q = g_*/(z\lambda)$. It is well-known that the conditional mean and variance of X, given z, are

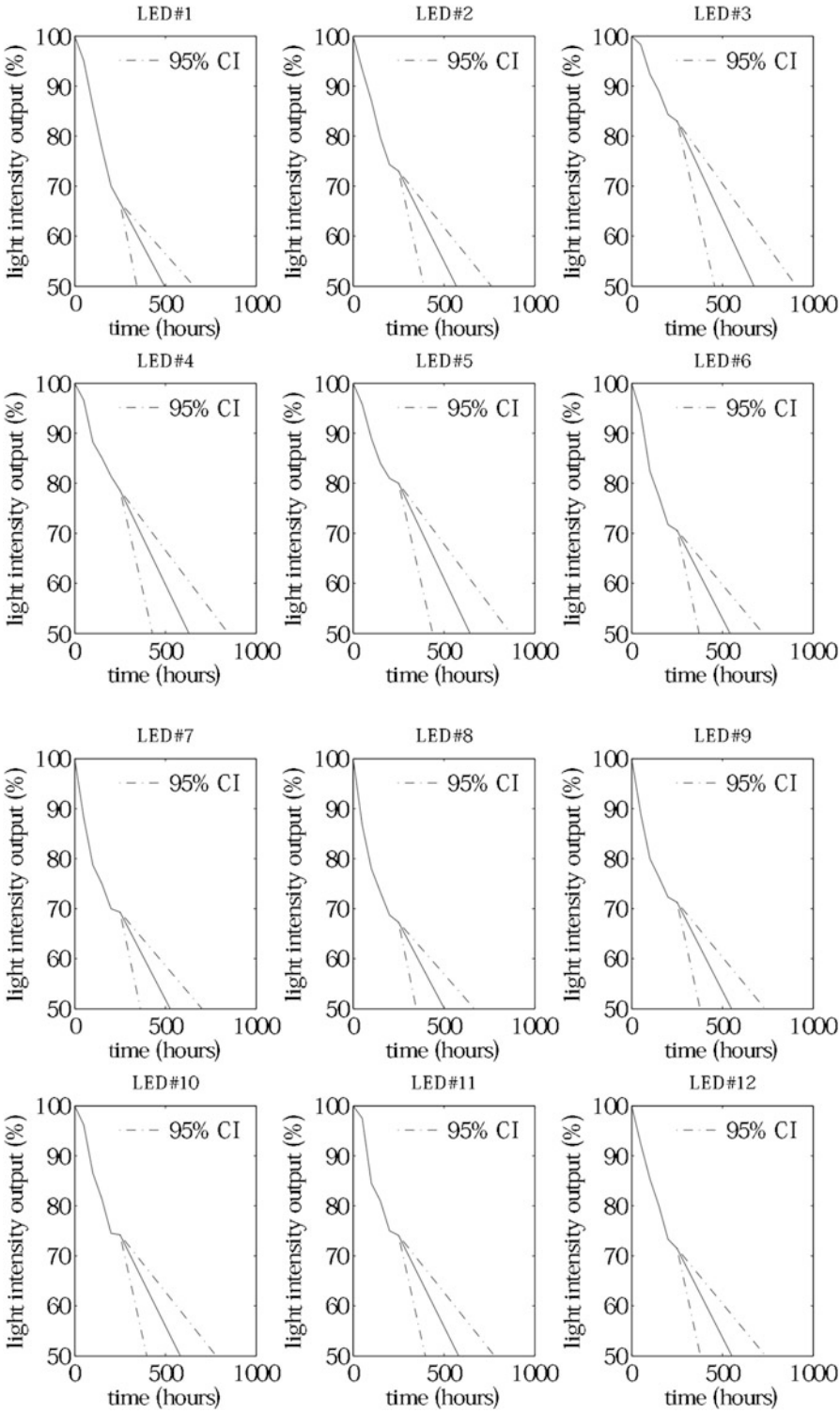

Fig. 13.3 Inference on remaining useful life for the 12 LEDs

$$E[X|z] = q\left(1 + \frac{p^2}{2}\right) \tag{13.10}$$

and

$$V[X|z] = (pq)^2\left(1 + \frac{5p^2}{4}\right), \tag{13.11}$$

respectively. Thus, we have

$$E[X|z] = q\left(1 + \frac{p^2}{2}\right) = \frac{g_*}{z\lambda} + \frac{1}{2\lambda}, \tag{13.12}$$

and

$$E[X^2|z] = V[X|z] + E[X|z]^2 = q^2 + 2p^2q^2 + \frac{3p^4q^2}{2} = \frac{g_*^2}{z^2\lambda^2} + \frac{2g_*}{z\lambda^2} + \frac{3}{2\lambda^2}. \tag{13.13}$$

As z follows an inverse gamma distribution with shape δ and scale γ, it implies that z^{-1} follows a gamma distribution with shape δ and scale γ^{-1}. Also, we have $E[z^{-1}] = \delta/\gamma$ and $E[z^{-2}] = \delta(\delta+1)/\gamma^2$. Therefore, we obtain

$$E[X] = E_z[E[X|z]] = \int_0^\infty \left(\frac{g_*}{z\lambda} + \frac{1}{2\lambda}\right) f_z(z)dz = \frac{g_*\delta}{\lambda\gamma} + \frac{1}{2\lambda}, \tag{13.14}$$

and

$$\begin{aligned} E[X^2] = E_z\left[E\left[X|z\right]\right] &= \int_0^\infty \left(\frac{g_*^2}{z^2\lambda^2} + \frac{2g_*}{z\lambda^2} + \frac{3}{2\lambda^2}\right) f_z(z)dz \\ &= \frac{g_*^2\delta(\delta+1)}{\lambda^2\gamma^2} + \frac{2g_*\delta}{\lambda^2\gamma} + \frac{3}{2\lambda^2}. \end{aligned} \tag{13.15}$$

Consequently, the variance of X is given by

$$V[X] = E[X^2] - E[X]^2 = \frac{g_*^2\delta}{\lambda^2\gamma^2} + \frac{g_*\delta}{\lambda^2\gamma} + \frac{5}{4\lambda^2}. \tag{13.16}$$

References

1. Banjevic D, Jardine AKS (2006) Calculation of reliability function and remaining useful life for a Markov failure time process. IMA J Manag Math 17:115–130
2. Birnbaum ZW, Saunders SC (1969) A new family of life distributions. J Appl Probab 6:319–327

3. Chaluvadi VNH (2008) Accelerated life testing of electronic revenue meters. Ph.D. dissertation, Clemson University, Clemson
4. Chhikara RS, Folks L (1989) The inverse gaussian distribution: theory, methodology and applications. Marcel Dekker, New York
5. Doksum KA, Hoyland A (1992) Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. Technometrics 34:74–82
6. Ebrahimi N (2004) System reliability based on diffusion models for fatigue crack growth. Nav Res Logist 52:46–57
7. Efron B (1982) The Jackknife, the bootstrap and other resampling plans. SIAM, Philadelphia
8. Fan JJ, Yung KC, Pecht M (2011) Physics-of-failure-based prognostics and health management for high-power white light-emitting diode lighting. IEEE Trans Device Mater Reliab 11:407–416
9. Gebraeel N, Pan J (2008) Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. IEEE Trans Reliab 57:539–550
10. Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
11. Ling MH, Tsui KL, Balakrishnan N (2015) Accelerated degradation analysis for the quality of a system based on the gamma process. IEEE Trans Reliab 64:463–472
12. Padgett WJ, Tomlinson MA (2004) Inference from accelerated degradation and failure data based on Gaussian processmodels. Lifetime Data Anal 10:191–206
13. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Anal 11:511–527
14. Si XS, Wang WB, Chen MY, Hu CH, Zhou DH (2013) A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution. Eur J Oper Res 226:54–66
15. Si XS, Wang WN, Hu CH, Zhou DH (2011) Remaining useful life estimation – a review on the statistical data driven approaches. Eur J Oper Res 213:1–14
16. Singpurwalla ND (1995) Survival in dynamic environments. Stat Sci 1:86–103
17. Wang X (2008) A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects. Stat Sin 18:1153–1163
18. Wang X (2009) Nonparametric estimation of the shape function in a gamma process for degradation data. Can J Stat 37:101–118
19. Whitmore GA (1995) Estimating degradation by a Wiener diffusion process subject to measurement error. Lifetime Data Anal 1:307–319
20. Whitmore GA, Schenkelberg F (1997) Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. Lifetime Data Anal 3:27–45
21. Ye ZS, Chen N (2014) The inverse Gaussian process as a degradation model. Technometrics 56:302–311

Chapter 14
ADDT: An R Package for Analysis of Accelerated Destructive Degradation Test Data

Zhongnan Jin, Yimeng Xie, Yili Hong, and Jennifer H. Van Mullekom

Abstract Accelerated destructive degradation tests (ADDTs) are often used to collect necessary data for assessing the long-term properties of polymeric materials. Based on the data, a thermal index (TI) is estimated. The TI can be useful for material rating and comparisons. The R package ADDT provides the functionalities of performing the traditional method based on the least-squares method, the parametric method based on maximum likelihood estimation, and the semiparametric method based on spline methods for analyzing ADDT data, and then estimating the TI for polymeric materials. In this chapter, we provide a detailed introduction to the ADDT package. We provide a step-by-step illustration for the use of functions in the package. Publicly available datasets are used for illustrations.

Keywords Degradation path • Long-term properties • Material durability • Material reliability • Monotonic splines • Semiparametric methods

14.1 Introduction

Accelerated destructive degradation tests (ADDTs) are commonly used to collect data to access the long-term properties of polymeric materials (e.g., [1]). Based on the collected ADDT data, a thermal index (TI) is estimated using a statistical model. In practice, the TI can be useful for material rating and comparisons. In literature, there are three methods available for ADDT data modeling and analysis: the traditional method based on the least-squares approach, the parametric method based on maximum likelihood estimation, and the semiparametric method based on spline models. The chapter in Xie et al. [2] provides a comprehensive review for the three methods for ADDT data analysis and compares the corresponding TI estimation procedures via simulations.

The R package ADDT in Hong et al. [3] provides the functionalities of performing the three methods and their corresponding TI estimation procedures. In this chapter,

Y. Xie • Z. Jin • Y. Hong (✉) • J.H. Van Mullekom
Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
e-mail: jinxx354@vt.edu; xym@vt.edu; yilihong@vt.edu; vanmuljh@vt.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_14

we provide a detailed introduction to the ADDT package. We provide a step-by-step illustration for the use of functions in the package. We also use publicly available datasets for illustrations.

The rest of the chapter is organized as follows. Section 14.2 introduces the three methods, the corresponding TI procedures, and the implementations in the R package. The Adhesive Bond B data ([4]) is used to do a step-by-step illustration. Section 14.3 provides a full analysis of the Seal Strength data ([5]) so that users can see a typical ADDT modeling and analysis process. Section 14.4 contains some concluding remarks.

14.2 The Statistical Methods

14.2.1 Data

In most applications, an ADDT dataset typically includes degradation measurements under different measuring time points, and accelerating variables such as temperature and voltage. In the ADDT package, there are four publicly available datasets ready for users to do analysis, which are the Adhesive Bond B data in [4], the Seal Strength data in [5], the Polymer Y data in [6], and the Adhesive Formulation K data in [7]. Users can load those datasets by downloading, installing the package ADDT and appropriately calling the *data* function. The following gives some example R codes.

```
>install.packages("ADDT")
>library(ADDT)
>data(AdhesiveBondB)
>data(SealStrength)
>data(PolymerY)
>data(AdhesiveFormulationK)
>AdhesiveBondB
>SealStength
```

Table 14.1 shows the Adhesive Bond B dataset. The first column is the acceleration variable, temperature in Celsius. Time points that used to measure the degradation and the degradation values are listed in columns 2 and 3 correspondingly. We illustrate the Adhesive Bond B data in Fig. 14.1. To use the R ADDT package, users need to format the data in the same form as the dataset shown in Table 14.1.

Another dataset that has been frequently used is the Seal Strength data where the strength from ten different seals were measured at five different time points under four different temperature levels. Seal Strength data is shown in Table 14.2. We will use the Adhesive Bond B data and Seal Strength data to illustrate the use of the ADDT package.

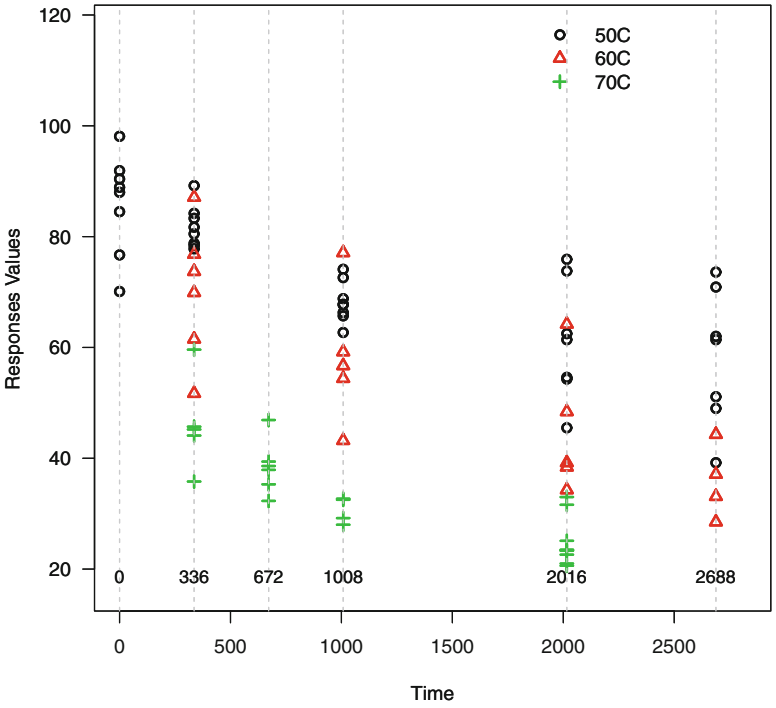

Fig. 14.1 Graphical representation of the Adhesive Bond B dataset. The x-axis stands for the time in hour while y-axis represents the degradation values

14.2.2 The Traditional Method

The traditional method using the least-squares approach is widely accepted and used in various industrial applications. The traditional method is a two-step approach that uses polynomial fittings and the least-squares method to obtain the temperature-time relationship. The TI can be obtained by using the fitted temperature-time relationship. In particular, for each temperature level, indexed by i, we find the mean time to failure m_i satisfies the following equation.

$$a_{0i} + a_{1i}m_i + a_{2i}m_i^2 + a_{3i}m_i^3 = y_f, i = 1, \cdots, n,$$

where y_f is the failure threshold and $(a_{0i}, a_{1i}, a_{2i}, a_{3i})'$ are coefficients. Here n is the number of temperature levels. The temperature-time relationship is expressed as

$$\log_{10}(m_i) = \beta_0 + \beta_1 x_i, i = 1, \cdots, n, \tag{14.1}$$

Table 14.1 The Adhesive Bond B data from Escobar et al. [4], which contains the testing of results of an ADDT for the strength of Adhesive Bond B

TempC	TimeH	Response	TempC	TimeH	Response	TempC	TimeH	Response
50	0	70.1	50	2016	62.5	60	2688	37.1
50	0	76.7	50	2016	73.8	60	2688	44.3
50	0	84.5	50	2016	75.9	70	336	35.8
50	0	88.0	50	2688	39.2	70	336	44.1
50	0	88.9	50	2688	49.0	70	336	45.2
50	0	90.4	50	2688	51.1	70	336	45.7
50	0	91.9	50	2688	61.4	70	336	59.6
50	0	98.1	50	2688	62.0	70	672	32.3
50	336	77.8	50	2688	70.9	70	672	35.3
50	336	78.4	50	2688	73.6	70	672	37.9
50	336	78.8	60	336	51.7	70	672	38.6
50	336	80.5	60	336	61.5	70	672	39.4
50	336	81.7	60	336	69.9	70	672	46.9
50	336	83.3	60	336	73.7	70	1008	28.0
50	336	84.2	60	336	76.8	70	1008	29.2
50	336	89.2	60	336	87.1	70	1008	32.5
50	1008	62.7	60	1008	43.2	70	1008	32.7
50	1008	65.7	60	1008	54.4	70	2016	20.6
50	1008	66.3	60	1008	56.7	70	2016	21.0
50	1008	67.7	60	1008	59.2	70	2016	22.6
50	1008	67.8	60	1008	77.1	70	2016	23.3
50	1008	68.8	60	2016	34.3	70	2016	23.4
50	1008	72.6	60	2016	38.4	70	2016	23.5
50	1008	74.1	60	2016	39.2	70	2016	25.1
50	2016	45.5	60	2016	48.4	70	2016	31.6
50	2016	54.3	60	2016	64.2	70	2016	33.0
50	2016	54.6	60	2688	28.5			
50	2016	61.4	60	2688	33.1			

which is based on the Arrhenius relationship to extrapolate to the normal use condition. With the parameterizations in this temperature-time relationship, the TI, denoted by R, can be estimated as:

$$R = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16\,.$$

where β_0 and β_1 are the same with the coefficients from equation (14.1), and t_d is the target time (usually $t_d = 100{,}000$ is used).

Table 14.2 The Seal Strength data in Li and Daganaksoy [5]. The table shows the strength of seal samples that were measured at five different time points under four different temperature levels

TempC	TimeH	Response	TempC	TimeH	Response	TempC	TimeH	Response	TempC	TimeH	Response
100	0	28.74	200	1680	42.21	250	2520	17.08	300	3360	3.08
100	0	25.59	200	1680	32.64	250	2520	11.52	350	3360	1.24
100	0	22.72	200	1680	32.10	250	2520	13.03	350	3360	1.57
100	0	22.44	200	1680	32.37	250	2520	18.37	350	3360	2.06
100	0	29.48	200	1680	33.59	300	2520	3.86	350	3360	1.56
100	0	23.85	200	1680	26.46	300	2520	4.76	350	3360	1.94
100	0	20.24	200	1680	33.69	300	2520	5.32	350	3360	1.39
100	0	22.33	250	1680	14.29	300	2520	3.74	350	3360	1.91
100	0	21.70	250	1680	20.16	300	2520	4.58	350	3360	1.44
100	0	27.97	250	1680	22.35	300	2520	3.62	350	3360	1.61
200	840	52.52	250	1680	21.96	300	2520	3.58	350	3360	1.50
200	840	30.23	250	1680	13.67	300	2520	3.47	200	4200	14.53
200	840	31.90	250	1680	14.40	300	2520	3.29	200	4200	17.95
200	840	33.15	250	1680	22.37	300	2520	3.63	200	4200	11.90
200	840	34.26	250	1680	13.08	350	2520	1.34	200	4200	17.00
200	840	31.82	250	1680	17.81	350	2520	0.92	200	4200	15.56
200	840	27.10	250	1680	17.82	350	2520	1.31	200	4200	18.07
200	840	30.00	300	1680	10.34	350	2520	1.76	200	4200	13.96
200	840	26.96	300	1680	13.24	350	2520	1.30	200	4200	13.57
200	840	42.73	300	1680	8.57	350	2520	1.47	200	4200	16.35

(continued)

Table 14.2 (continued)

TempC	TimeH	Response	TempC	TimeH	Response	TempC	TimeH	Response	TempC	TimeH	Response
250	840	28.97	300	1680	11.93	350	2520	1.11	200	4200	18.76
250	840	35.01	300	1680	13.76	350	2520	1.25	250	4200	14.75
250	840	27.39	300	1680	16.44	350	2520	1.02	250	4200	11.54
250	840	36.66	300	1680	14.81	350	2520	1.30	250	4200	11.57
250	840	27.91	300	1680	11.50	200	3360	26.72	250	4200	10.83
250	840	31.03	300	1680	11.92	200	3360	21.24	250	4200	12.78
250	840	32.65	300	1680	10.30	200	3360	22.76	250	4200	10.14
250	840	35.08	350	1680	5.78	200	3360	24.39	250	4200	11.45
250	840	28.05	350	1680	5.90	200	3360	15.93	250	4200	12.91
250	840	33.54	350	1680	6.99	200	3360	23.90	250	4200	13.06
300	840	10.63	350	1680	7.94	200	3360	22.09	250	4200	6.76
300	840	8.28	350	1680	7.06	200	3360	23.69	300	4200	1.95
300	840	13.46	350	1680	5.13	200	3360	23.67	300	4200	1.55
300	840	13.47	350	1680	5.80	200	3360	20.94	300	4200	2.19
300	840	9.44	350	1680	6.20	250	3360	14.23	300	4200	2.00
300	840	7.66	350	1680	5.30	250	3360	12.83	300	4200	2.00
300	840	11.16	350	1680	6.34	250	3360	13.02	300	4200	2.33
300	840	8.70	200	2520	9.47	250	3360	16.74	300	4200	1.80
300	840	9.44	200	2520	13.61	250	3360	12.11	300	4200	2.34
300	840	12.23	200	2520	8.95	250	3360	12.24	300	4200	1.88

350	840	13.79	200	2520	8.61	250	3360	18.97	300	4200	2.66
350	840	15.10	200	2520	10.16	250	3360	15.29	350	4200	0.27
350	840	20.58	200	2520	8.82	250	3360	14.38	350	4200	0.20
350	840	18.20	200	2520	8.84	250	3360	14.80	350	4200	0.26
350	840	16.64	200	2520	10.73	300	3360	2.89	350	4200	0.26
350	840	10.93	200	2520	10.63	300	3360	3.31	350	4200	0.27
350	840	12.28	200	2520	7.70	300	3360	1.81	350	4200	0.18
350	840	18.65	250	2520	9.59	300	3360	1.61	350	4200	0.13
350	840	20.80	250	2520	14.37	300	3360	2.65	350	4200	0.20
350	840	15.04	250	2520	12.08	300	3360	2.83	350	4200	0.13
200	1680	31.37	250	2520	11.79	300	3360	2.70	350	4200	0.21
200	1680	37.91	250	2520	17.69	300	3360	2.79			
200	1680	38.03	250	2520	14.05	300	3360	1.83			

In the R package ADDT, we implement the traditional method by using:

```
>addt.fit.lsa<-addt.fit(Response~TimeH+TempC,data=Adh
esiveBondB, proc="LS", failure.threshold=70)
```

The *addt.fit* function in ADDT package fits the traditional model automatically when users specify *proc* = *"LS"* argument. In function *addt.fit*, other arguments include:

- *formula*: We use *Response* ∼ *TimeH+TempC* to represent the model formula. The *Response*, *TimeH*, and *TempC* specify the response, time, and temperature columns in the dataset, respectively. Note that the order of *TimeH* and *TempC* can not be exchanged in the formula.
- *data*: The name of the dataset for analysis. The dataset should have the same layout as the Adhesive Bond B in Table 14.1. Specifically, the order of the three columns should be the same as Adhesive Bond B, which is TempC, TimeH, and Response.
- *initial.value*: We need response measurements at time point 0 to compute the initial degradation level in the model. If the data do not contain that information, the user must supply the *initial.value*. Otherwise, the function will give an error message.
- *failure.threshold*: This argument sets the failure threshold. The default value of the soft failure threshold is 70% of the initial value in the ADDT package examples. Note that in industrial standards such as UL 746B [1], the failure threshold is usually 50%.
- *time.rti*: The *addt.fit* function allows users to specify the expected time associated with the TI. The default value for *time.rti* is t_d =100,000 h.
- *method*: This argument specifies the method that is used in the optimization process. Details can be found in *optim* function in R. The default value is *"Nelder-Mead"*.
- *subset*: This argument allows the users to specify a subset of the dataset for modeling.

The above arguments are the basic model inputs to run *addt.fit*, when *proc*=*"LS"*. Other methods, *proc* =*"ML"* (the parametric method) and *proc* =*"SemiPara"* (the semiparametric method) also require the same arguments. However, there are additional arguments for the other two methods and we will introduce them in Sects. 14.2.3 and 14.2.4.

We store the model fitting results in the *addt.fit.lsa* in this example. Users can print the model summary table and plots upon appropriate call. Examples are listed below:

```
> summary(addt.fit.lsa)

Least Squares Approach:
    beta0     beta1
 -13.7805 5535.0907
est.TI: 22
```

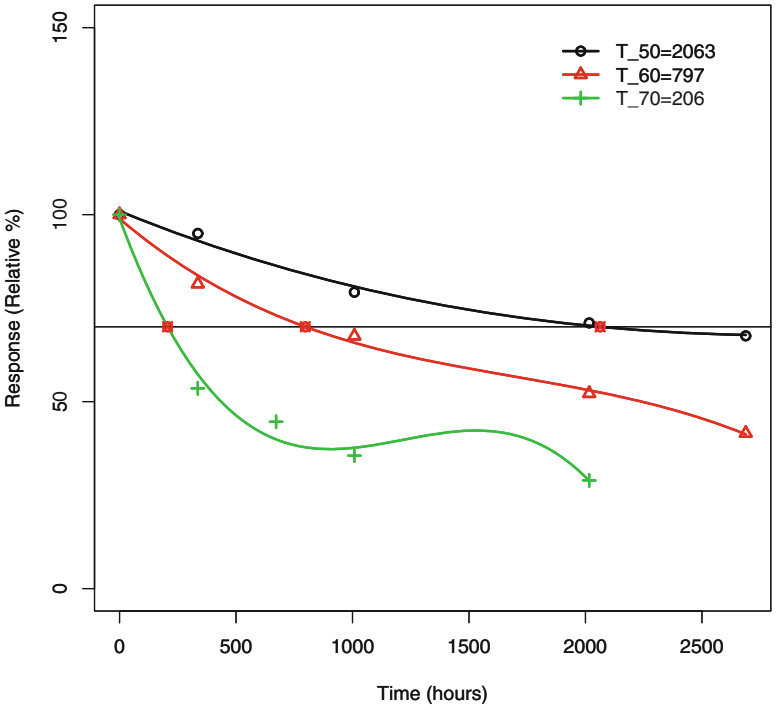

Fig. 14.2 Plot of the fitted polynomial curves for each temperature level, and the corresponding interpolated time to failures. The *horizontal dark line* presents the failure threshold. The y-axis shows the relative value of material strength

```
Interpolation time:
     Temp       Time
[1,]   50 2063.0924
[2,]   60  797.1901
[3,]   70  206.1681
```

The *summary* function for *proc* = *"LS"* provides the parameter estimates and interpolated mean time to failure for the corresponding temperature levels. In the Adhesive Bond B example, the parameter estimates are $\hat{\beta}_0 = -13.7805$ and $\hat{\beta}_1 = 5535.0907$ for the temperature-time relationship. Estimated mean time to failure for temperature level at 50 °C, 60 °C, and 70 °C, are 2063.092, 797.190 and 206.168 h, respectively. The estimated TI is 22 °C in this example. Figure 14.2 shows the fitted polynomial curves for each temperature levels and the corresponding interpolated mean time to failure, according to least-squares method. The R code that is used to plot the results is shown below.

```
>plot(addt.fit.lsa, type="LS")
```

14.2.3 The Parametric Method

Different from the two-step approach in the traditional method, for the parametric method, one uses a parametric model to describe the degradation path. The maximum likelihood (ML) method is then used to estimate the unknown parameters in the model. In particular, we assume that degradation measurement y_{ijk} at time t_{ij} for temperature level i follows the model:

$$y_{ijk} = \mu(t_{ij}; x_i) + \epsilon_{ijk}, i = 1, \cdots, n, j = 1, \cdots, n_i, k = 1, \cdots, n_{ij},$$

where

$$x_i = \frac{1}{\text{TempC}_i + 273.16},$$

TempC_i is the temperature level, the value 273.16 is used to convert the temperature to Kelvin temperature scale. Here n is the number of temperature levels, n_i is the number of time points for level i, and n_{ij} is the number samples tested under the temperature time combination and ϵ_{ijk} is the error term. For polymer materials, the following parametric assumption for $\mu(t; x)$ (e.g., [8]) is used

$$\mu(t; x) = \frac{\alpha}{1 + [\frac{t}{\eta(x)}]^{\gamma}}, \tag{14.2}$$

where α represents the initial degradation, and γ is a shape parameter. Here,

$$\eta(x) = \exp(\upsilon_0 + \upsilon_1 x).$$

is the scale factor that is based on the Arrhenius relationship. By the parametric specification, the ML method is then used to estimate the parameters. King et al. [9] performed a comprehensive comparison between the traditional method and the parametric method. Xie et al. [2] performed a comprehensive comparison among the three methods in term of TI estimation.

For the model in (14.2), the TI is calculated as follows:

$$\text{TI} = \frac{\beta_1}{\log_{10}(t_d) - \beta_0} - 273.16,$$

where β_0 and β_1 are defined as:

$$\beta_0 = \frac{\upsilon_0}{\log(10)} + \frac{1}{\gamma \log(10)} \log\left[\frac{1-p}{p}\right], \quad \text{and} \quad \beta_1 = \frac{\upsilon_1}{\log(10)}.$$

To fit the parametric model, one can use the following command:

```
> addt.fit.mla<-addt.fit(Response~TimeH+TempC,data=Adh
esiveBondB,proc="ML", failure.threshold=70)
```

Similar to the "LS" case, here we provide an example of ML method based on the parametric method implemented in R. Using the same dataset Adhesive Bond B, we now change the *proc* argument to *proc* = *"ML"* so that the parametric model is used. The model results are stored in *addt.fit.mla*. Argument setups are almost the same as those in *addt.fit* for the case of *proc* = *"LS"* except for additional arguments: *"starts"* and *"fail.thres.vec"*. In particular,

- *starts*: It provides a set of starting values for the ML estimation procedure. If this value is not supplied, the function will use the least-squares method to estimate for a set of starting values for the ML estimation.
- *fail.thres.vec*: If the user does not specify *starts* argument, the user may instead provide a vector of two different *failure.thresholds*. The least-squares procedure is then used for the two different failure thresholds to produce starting values for the ML procedure.

For the model results in *addt.fit.lma*, we not only have the parameter estimates as in the LS example, but also have confidence intervals for the model parameters and the TI. The following shows the summary information of the model fitting.

```
> summary(addt.fit.mla)

 Maximum Likelihood Approach:
Call:
lifetime.mle(dat = dat0, minusloglik = minus.loglik.ki
netics,  starts = starts, method = method, control =
 list(maxit = 1e+05))

Parameters:
            mean       std  95% Lower  95% Upper
alpha    87.2004    2.5920    82.2653    92.4315
beta0   -37.2360    4.6450   -46.3401   -28.1318
beta1 14913.1628 1561.1425 11853.3235 17973.0022
gamma     0.7274    0.0870     0.5753     0.9195
sigma     8.2017    0.6405     7.0377     9.5581
rho       0.0000    0.0003    -0.0006     0.0006

Temperature-Time Relationship:
     beta0      beta1
 -16.6830   6478.5641

TI:
      est       std 95% Lower 95% Upper
  25.6183    3.0980   19.5465   31.6902

Loglikelihod:
[1] -288.9057
```

By applying *summary* function to the *addt.fit* results, we have the ML estimates for $\alpha, \upsilon_0, \upsilon_1, \gamma, \sigma$, and ρ along with their standard deviation as well as the associated 95% confidence intervals based on large-sample approximations. The log-likelihood values for the final model is also printed for model comparisons.

The summary table will perform the TI estimates and confidence interval calculation automatically by assigning the default confidence level as 95%. Users can change the confidence level to other values by using the function *addt.confint.ti.mle* and specifying the desired value for *conflevel*. In particular,

```
> addt.confint.ti.mle(addt.fit.mla, conflevel = 0.99)
```

provides an example of customizing confidence level for TI estimates. It shows that the 99% confidence interval for TI and the confidence interval is wider than using 95% as the confidence level. The results are shown as follows.

```
      est.       s.e.      lower      upper
  25.618      3.097     17.638     33.598
```

Similar to the LS method, we can visualize the model fitting results. For the ML method, one can plot the fitted lines along with the data by employing *plot.addt.fit*. Figure 14.3 shows the illustration of the fitting results of *plot.addt.fit*.

```
> plot(addt.fit.mla, type="ML")
```

14.2.4 The Semiparametric Method

Different from the traditional method and the parametric method that are introduced in Sects. 14.2.2 and 14.2.3, the semiparametric method is applicable to different materials with a nonparametric form for the baseline degradation path. In addition, the parametric part of the model (i.e., the Arrhenius relationship) retains the extrapolation capacity to the use condition. Similarly to the parametric model, we model the degradation measurement as follows,

$$y_{ijk} = \mu(t_{ij}, x_i; \theta) + \epsilon_{ijk},$$

where

$$x_i = -\frac{11605}{\text{TempC}_i + 273.16},$$

and $\boldsymbol{\theta}$ stands for all the parameters in the model. We use the semiparametric model structure to describe the degradation path. In particular, the degradation path is modeled as

$$\mu(t_{ij}, x_i) = g[\eta_i(t_{ij}; \beta); \gamma], \tag{14.3}$$

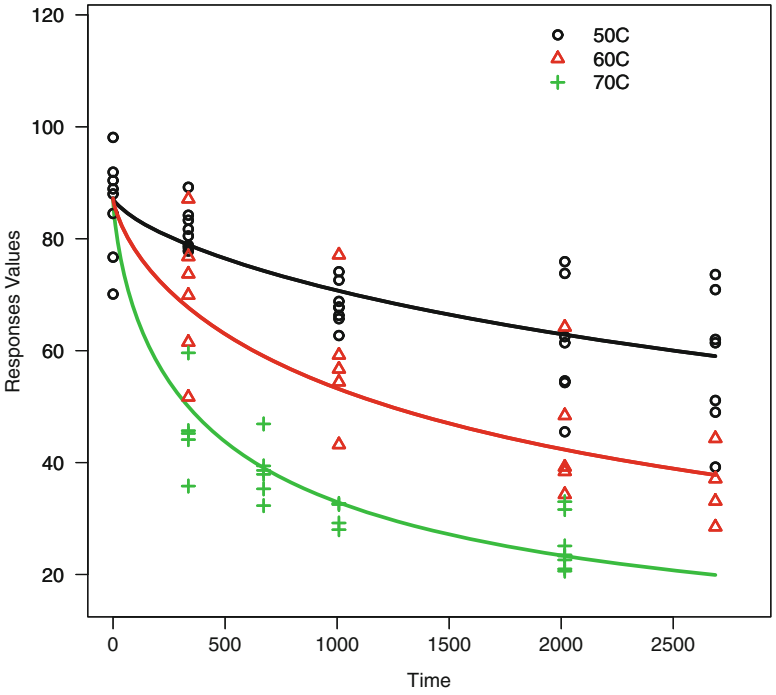

Fig. 14.3 Plot of the original dataset of Adhesive Bond B as well as the fitted degradation paths based on the parametric model. The *black line* with "∘", *red line* with "△" and *green line* with "+" stand for fitted lines at 50, 60 and 70 degree, respectively

and the scale factor is

$$\eta_i(t;\beta) = \frac{t}{\exp(\beta s_i)}, \tag{14.4}$$

with acceleration parameter β. In equation (14.4), we define $s_i = x_{max} - x_i$ where x_{max} is the transformed value of the highest level of temperature. We assume the error terms follow normal distribution with variance σ^2 and the correlations between two error terms are ρ. That is,

$$\epsilon_{ijk} \sim \mathrm{N}(0, \sigma^2),$$

and

$$\mathrm{Corr}(\epsilon_{ijk}, \epsilon_{ijk'}) = \rho. \tag{14.5}$$

We assume that $k \neq k'$ in the error terms correlations in (14.5). In (14.3), $g(\cdot)$ is a monotonically decreasing function modeled by splines with parameter vector γ. See Xie et al. [7] for more details on the semiparametric method.

As a more flexible method designated to a wide variety of materials, the nonparametric component is used to build the baseline degradation path. With inner knots $d_1 \leq d_2 \leq \cdots \leq d_N$ and boundary knots d_0, d_{N+1}, the l-th B-spline basis function with a degree of q can be expressed at z by recursively building the following models:

$$\begin{aligned} B_{0,l}(z) &= 1(d_l \leq z \leq d_{l+1})B_{q,l}(z) \\ &= \frac{z - d_l}{d_{l+q} - d_l} B_{q-1,l}(z) + \frac{d_{l+q+q} - z}{d_{l+q+1} - d_{l+1}} B_{q-1,l+1}(z). \end{aligned}$$

The degradation can be expressed as follows.

$$y_{ijk} = \sum_{l=1}^{p} \gamma_l B_{q,l}[\eta_i(t_{ij}; \beta)] + \epsilon_{ijk},$$

where $\eta(t; \beta)$ accounts for the parametric part while $g(\cdot)$ is the nonparametric component which is constrained to be monotonically decreasing to retain the meanings of the degradation process.

Similarly to the "LS" and "ML" methods, we implement the semiparametric model in R. In *addt.fit*, *proc="SemiPara"* enables users to fit a semiparametric model to the degradation data as we discussed above. In particular,

```
>addt.fit.semi<-addt.fit(Response~TimeH+TempC,data=Adh
esiveBondB,proc="SemiPara",failure.threshold=70)
```

Other than the arguments we introduced for *proc = "LS"* and *proc = "ML"*, there is an other unique option in the *addt.fit* when *proc = "SemiPara"* is called. That is:

- *semi.control*: This argument contains a list of control parameters regarding the *SemiPara* option. Users can specify the model assumptions like correlation *rho*. In *semi.control = list(cor = F, ...)*, the default value is to exclude the correlation term in the model (i.e., $\rho = 0$). If *cor = T*, then there will be a correlation term in the semiparametric model.

Summary results of the semiparametric model object given by *addt.fit* include $\hat{\beta}$, $\hat{\rho}$, knots that were used by the model, log-likelihood and AICc for the final model, which are both model evaluation quantities. Note that in the example shown below, we use the default set up for semiparametric model fit on the Adhesive Bond B data.

```
> summary(addt.fit.semi)

Semi-Parametric Approach:

Parameters Estimates:
betahat
  1.329
```

```
TI estimates:
 TI.semi     beta0     beta1
  26.313   -17.363 6697.074

Model Evaluations:
Loglikelihood       AICC
     -288.135    586.269

B-spline:
Left Boundary    knots    Right Boundary
         0.00   180.66           2016.00
```

We can also call *plot.addt.fit* to present model fitting results.

```
plot(addt.fit.semi, type="SEMI")
```

Figure 14.4 shows the plot of the original dataset of Adhesive Bond B data as well as the fitted degradation mean values using the semiparametric model. Here we assume that there is no correlation ρ between two error terms. Note that for *plot.addt*

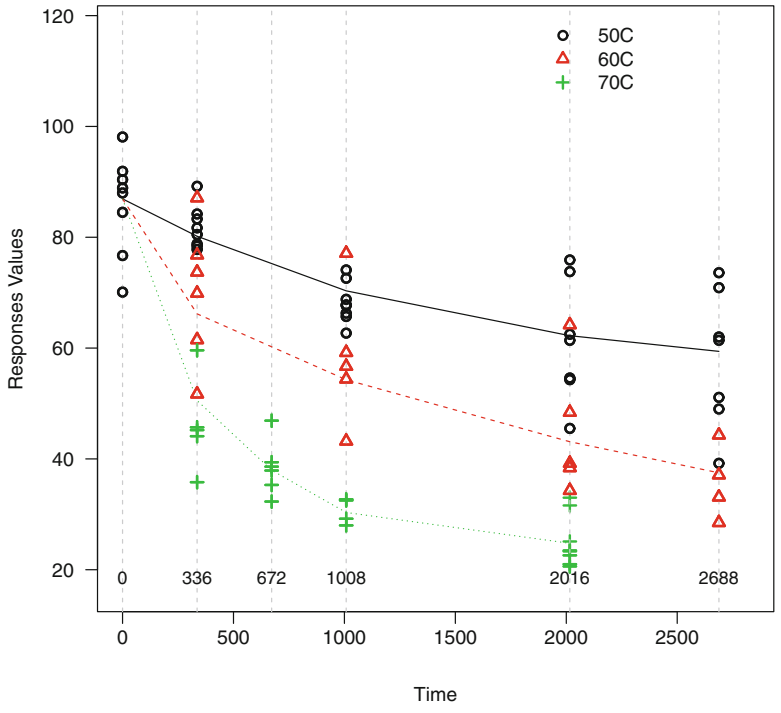

Fig. 14.4 Plot of the original dataset of Adhesive Bond B data as well as the fitted degradation mean values using the semiparametric model

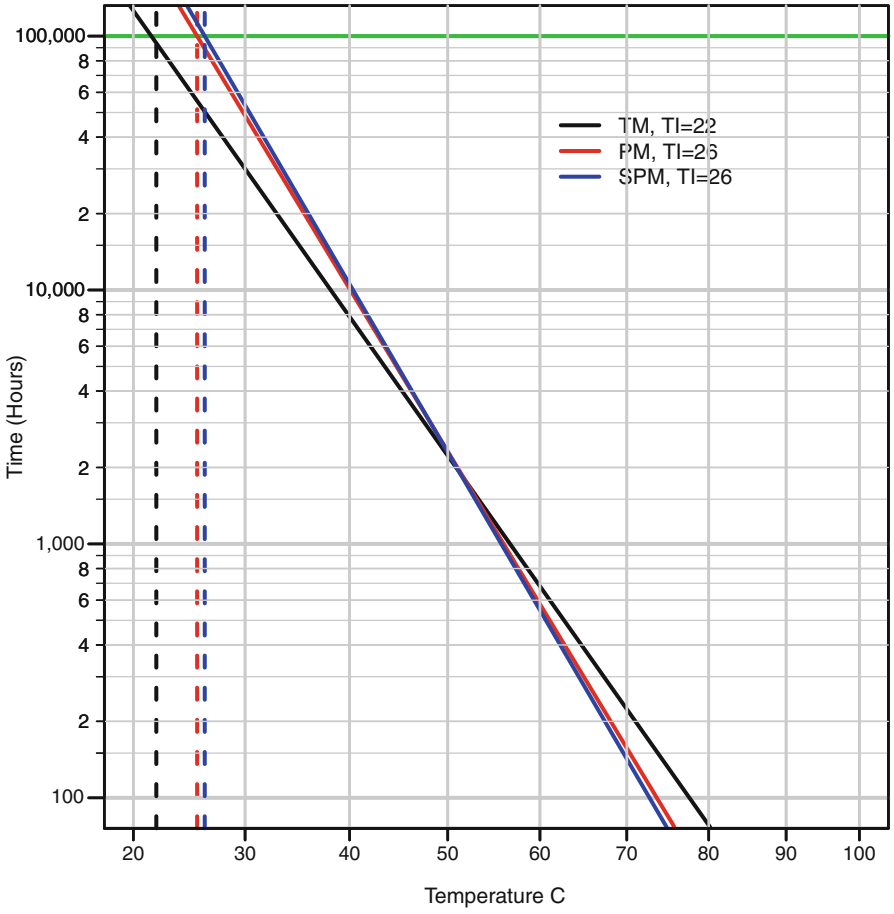

Fig. 14.5 Fitted temperature-time relationship lines for the Adhesive Bond B data from the least-squares, maximum likelihood, and the semiparametric methods. The failure threshold of 70%

function, *type* argument should be compatible with the *addt.obj*, meaning that type used in plot function should be the same with *proc* argument in the function *addt.fit*, otherwise error messages will be generated.

We illustrate the comparisons among the least-squares, maximum likelihood and semiparametric methods in terms of TI estimation in Fig. 14.5. Temperature-time relationship lines are plotted for all three methods in black, red and blue lines correspondingly.

14.3 Data Analysis

In this section, we present a complete ADDT data analysis using the Seal Strength data to illustrate the use of functions in Sect. 14.2. The details of the Seal Strength data is available in Li and Doganaksoy [5]. The first ten observations are listed below. Note that in the Seal Strength data, temperatures at time point 0 are modified to 200 degrees while those in the original Seal Strength dataset in Table 14.2 are 100 degrees. Changing temperatures at time point 0 to the lowest temperature is a computing trick that will not affect fitting results, because at time 0, the temperature effect has not kicked in yet.

```
>head(SealStrength, n=10)
    TempC TimeH Response
1     200     0    28.74
2     200     0    25.59
3     200     0    22.72
4     200     0    22.44
5     200     0    29.48
6     200     0    23.85
7     200     0    20.24
8     200     0    22.33
9     200     0    21.70
10    200     0    27.97
```

A graphical representation of the data is useful for users to obtain a general idea of the degradation paths. Using the *addt.fit.mla* object from *addt.fit* with *proc="ML"*, one can plot the degradation paths using option *type="data"*.

```
>plot(addt.fit.mla, type="data")
```

Figure 14.6 shows the plot of the Seal Strength data, in which the degradations were measured at six different time points under three different temperatures. For Seal Strength data, we observe a decreasing trend in degradation measurements as time increases. Degradation measurements decrease with the accelerating variable, temperature as well.

Three different *addt.fit* models can be fitted, which are *proc ="LS"*, *proc = "ML"*, and *proc= "SemiPara"*.

```
>addt.fit.lsa<-addt.fit(Response~TimeH+TempC,data=Seal
Strength,proc="LS",failure.threshold=70)

>addt.fit.mla<-addt.fit(Response~TimeH+TempC,data=Seal
Strength,proc="ML",failure.threshold=70)

>addt.fit.semi<-addt.fit(Response~TimeH+TempC,data=Seal
Strength,proc="SemiPara",failure.threshold=70)
```

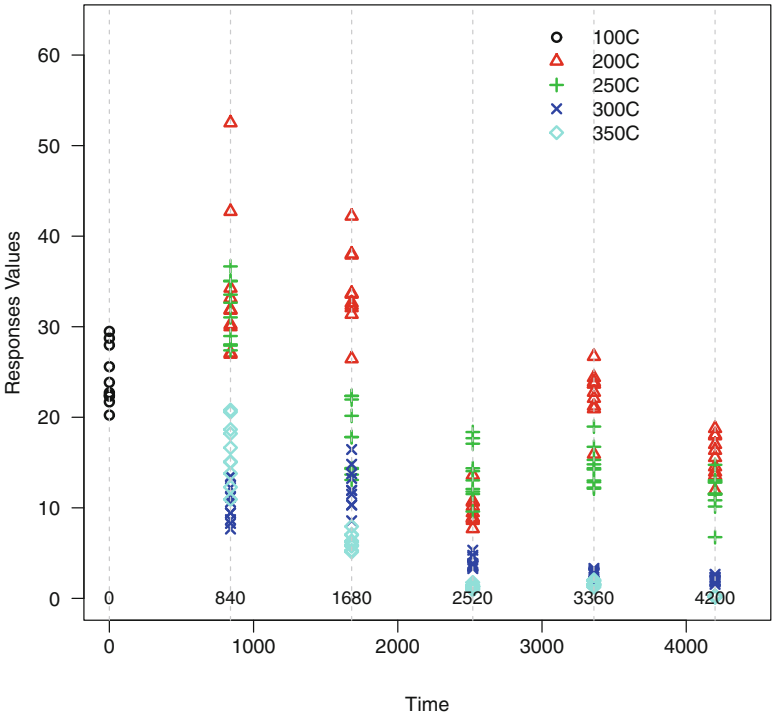

Fig. 14.6 Plot of the Seal Strength data. Degradations were measured at six different time points under three different temperatures

Alteratively, users can specify all three methods via one call of *addt.fit* by setting *proc = "All"*. The returned object for the three methods is stored in *addt.fit.all*.

```
> addt.fit.all<-addt.fit(Response~TimeH+TempC,data=Seal
Strength,proc="All",=failure.threshold=70)
```

To view the results of all three models, users can call the *summary* function:

```
> summary(addt.fit.all)
Least Squares Approach:
    beta0     beta1
   0.1934 1565.1731
est.TI: 52
Interpolation time:
     Temp      Time
[1,]  200 2862.3430
[2,]  250 2282.3303
[3,]  300  509.2084
[4,]  350  622.0857
```

```
 Maximum Likelihood Approach:
Call:
lifetime.mle(dat = dat0, minusloglik = minus.
loglik.kinetics, starts = starts, method =
method, control = list(maxit = 1e+05))

Parameters:
           mean      std 95% Lower 95% Upper
alpha   30.5898   3.4550   24.5152   38.1697
beta0    0.2991   1.7013   -3.0355    3.6337
beta1 3867.7170 899.5312 2104.6360 5630.7981
gamma    1.6556   0.4171    1.0105    2.7127
sigma    5.5456   0.6521    4.4041    6.9831
rho      0.7306   0.0664    0.6004    0.8607

Temperature-Time Relationship:
    beta0     beta1
  -0.0942 1680.4055

TI:
      est       std 95% Lower 95% Upper
  56.6920   28.1598    1.4997  111.8842

Loglikelihood:
[1] -555.0169

 Semi-Parametric Approach:

Parameters Estimates:
betahat
  0.282

TI estimates:
 TI.semi    beta0    beta1
  32.768    0.362 1418.833

Model Evaluations:
Loglikelihood          AICC
     -639.206      1288.412

B-spline:
 Left Boundary  knots  knots  knots   knots
          0.00 268.60 527.17 840.00 1394.55
 Right Boundary
 4200.00
```

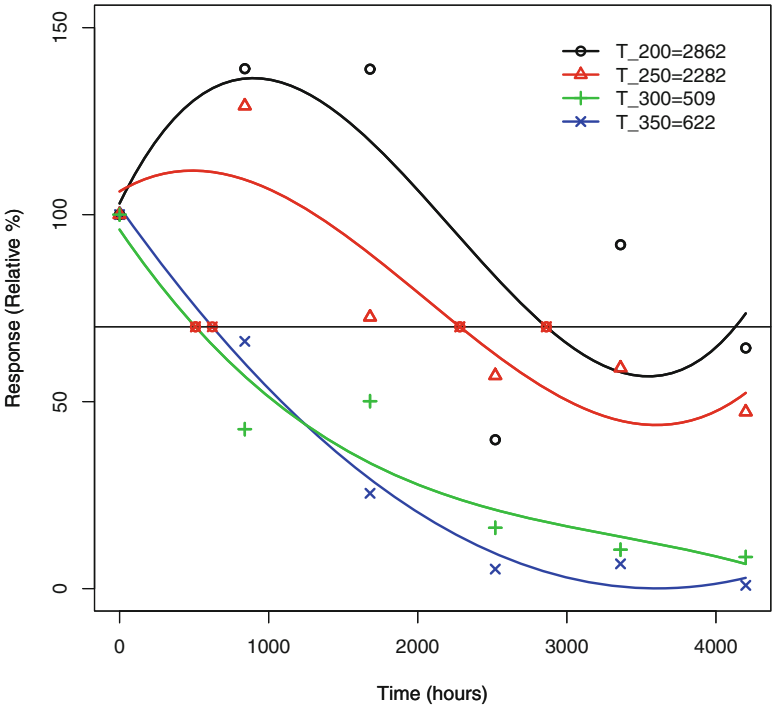

Fig. 14.7 Plot of the Seal Strength parametric lines with the least-squares method. The *black line* with "∘", *red line* with "△", *green line* with "+", and *blue line* with "×" represent 200, 250, 300 and 350 °C interpolated curves, respectively

Results shown here are the same when users call *summary* for three different models separately. The *add.fit.all* and *summary* for *addt.fit.all* provides an alternative way to analyze the data simultaneously.

Similar to Sect. 14.2, we illustrate the results from the least-squares, the maximum likelihood, and the semiparametric methods in Figs. 14.7, 14.8, 14.9, and 14.10, respectively. Note that in Figs. 14.9 and 14.10, we show the results for models without ρ and with ρ, respectively.

In addition, users can specify the *semi.control* argument in the *SemiPara* fit option. The *semi.control* contains a list of arguments that regards the *SemiPara* option in the model. For example, whether or not to include a correlation ρ in the model. When *semi.control = list(cor = T)*, the model will fit the correlation model with ρ. Otherwise, when default value *semi.control = list(cor = F)* or *semi.control* is not specified, the no-correlation model will be fitted. Note that for the option *SemiPara* in the function *addt.fit*, including the correlation ρ in the model may require more computing time, but potentially it will provide a better fit.

Here we compare the model results from the traditional method, the parametric method, and the semiparametric method for the Seal Strength data. In the results

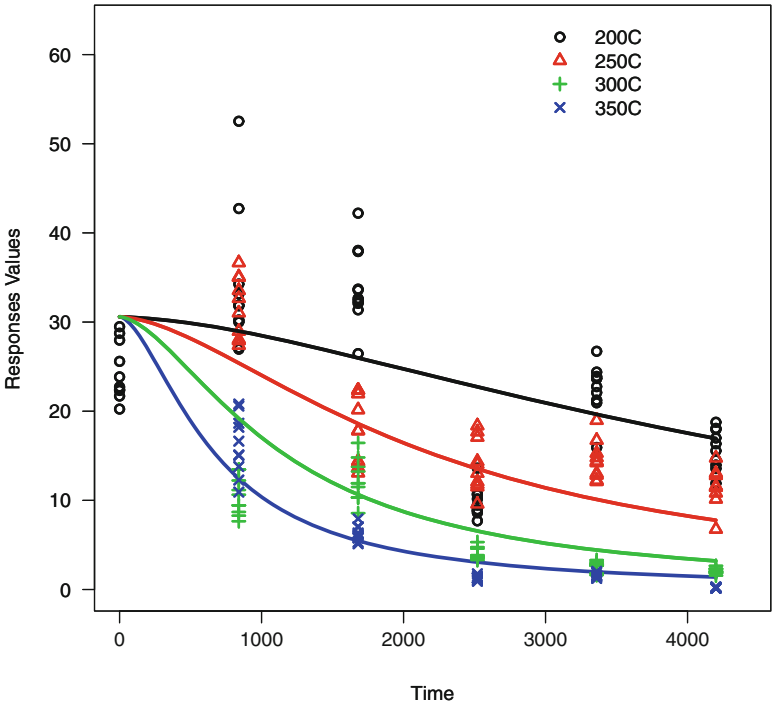

Fig. 14.8 Plot of the fitted mean function using maximum likelihood method for the Seal Strength data. The 200, 250, 300 and 350 °C estimated curves are represented. The *black line* with "∘", *red line* with "△", *green line* with "+", and *blue line* with "×", respectively

from *summary*, TI estimates are 52 °C, 56 °C and 47 °C, respectively. With β_0 and β_1 estimates, the TI plot is presented in Fig. 14.11. The black line is the TI from the traditional model, the red line is the parametric model TI estimates, and the blue line stands for the results from the semiparametric method.

In the results from two methods, without and with the correlation ρ, $\hat{\beta}$ are 0.282 and 0.323, while TI estimates are 32.768 and 47.338, respectively. The differences come from the assumption of ρ in the model. From the AICc value, the model with correlation provides a better fit to the data because it provides a smaller AICc value. The details of the model outputs are shown as follows.

```
>addt.fit.semi.no.cor<-addt.fit(Response~TimeH
+TempC,data=SealStrength,proc="SemiPara",
failure.threshold=70)

>addt.fit.semi.cor<-addt.fit(Response~TimeH
+TempC,data=SealStrength,proc="SemiPara",
failure.threshold=70, semi.control = list(cor=T))
```

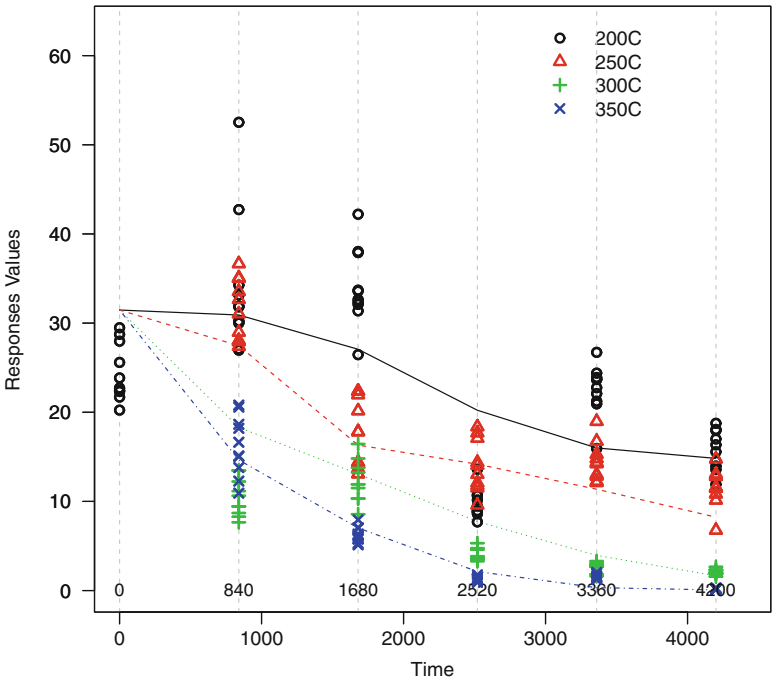

Fig. 14.9 Plots of fitted lines using the semiparametric method for the Seal Strength data, for the model without ρ

- Model without correlation ρ:

```
> summary(addt.fit.semi.no.cor)

 Semi-Parametric Approach:

Parameters Estimates:
betahat
  0.282

TI estimates:
 TI.semi    beta0    beta1
  32.768    0.362 1418.833

Model Evaluations:
Loglikelihood     AICC
     -639.206 1288.412
```

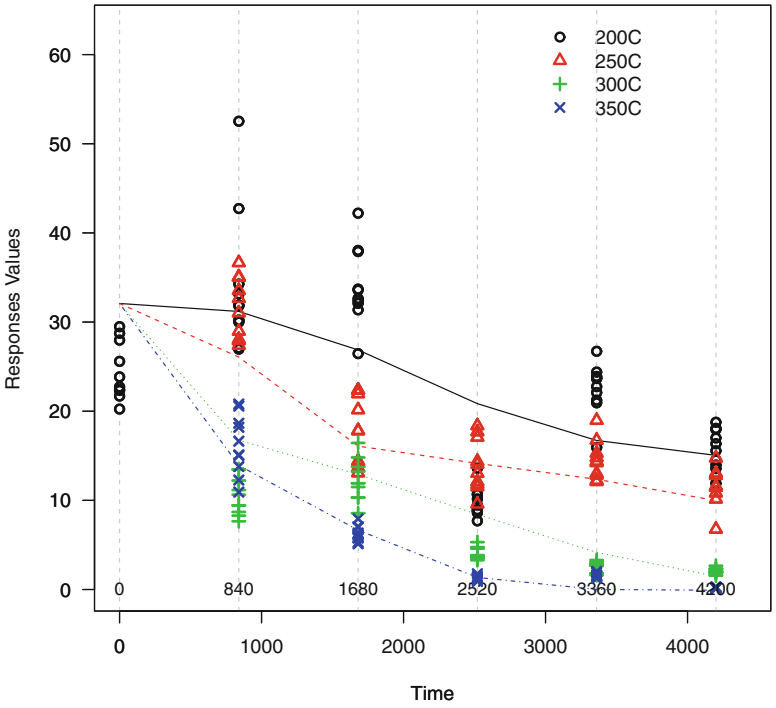

Fig. 14.10 Plots of fitted lines using the semiparametric method for the Seal Strength data, for the model with ρ

```
B-spline:
 Left Boundary    knots     knots     knots    knots
          0.00   268.60    527.17    840.00  1394.55
 Right Boundary
 4200.00
```

- Model with correlation ρ:

```
> summary(addt.fit.semi.cor)

 Semi-Parametric Approach:

Parameters Estimates:
betahat      rho
  0.323    0.714

TI estimates:
 TI.semi     beta0     beta1
  47.338    -0.087 1630.282
```

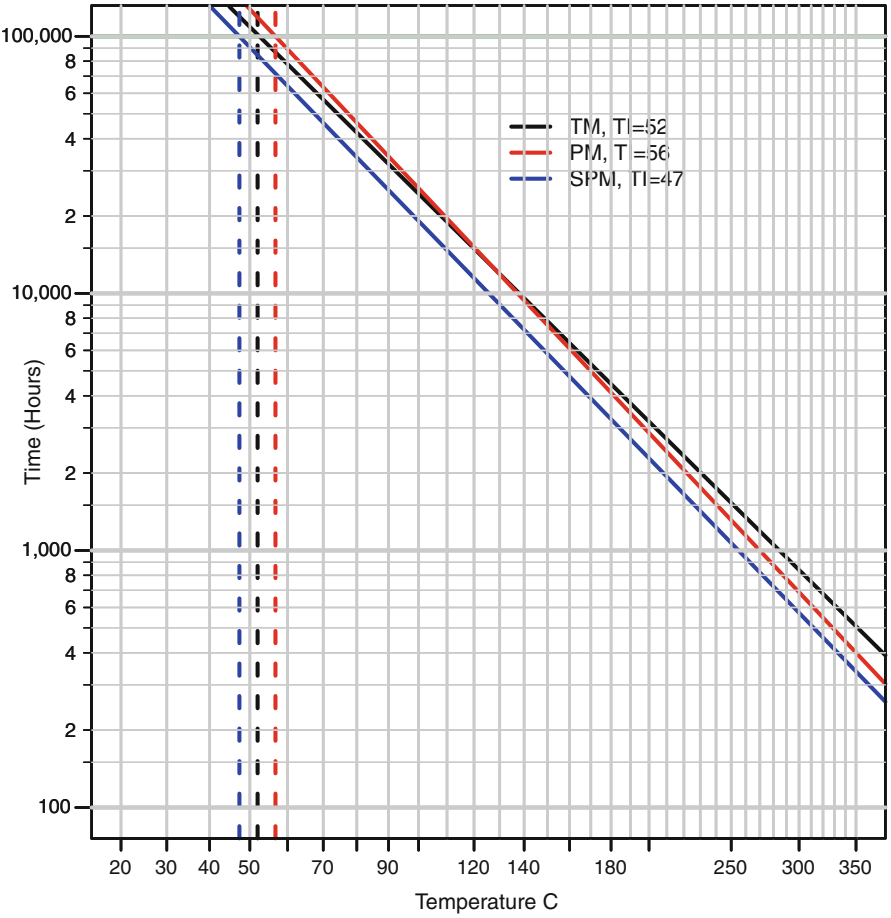

Fig. 14.11 Fitted temperature-time relationship lines for Seal Strength data using the traditional, maximum likelihood, and semiparametric methods. The failure threshold is 70%

```
Model Evaluations:
Loglikelihood     AICC
     -552.662  1117.323

B-spline:
 Left Boundary    knots     knots     knots     knots
          0.00   265.59    520.02    840.00   2483.29
 Right Boundary
 4200.00
```

14.4 Concluding Remarks

In this chapter, we provide a comprehensive description with illustrations for the ADDT methods implemented in the ADDT package. Functions such as the *addt.fit* and *summary* are illustrated for the traditional method, the parametric method, and the semiparametric method. We also show R examples using the Adhesive Bond B data and the Seal Strength data under various function options like *proc* and *semi.control*. Results from three different models are discussed. Users can consult the reference manual [3] for further details regarding the software package.

Acknowledgements The authors acknowledge Advanced Research Computing at Virginia Tech for providing computational resources. The work by Hong was partially supported by the National Science Foundation under Grant CMMI-1634867 to Virginia Tech.

References

1. UL746B (2013) Polymeric materials – long term property evaluations, UL 746B. Underwriters Laboratories, Incorporated
2. Xie Y, Jin Z, Hong Y, Van Mullekom JH (2017) Statistical methods for thermal index estimation based on accelerated destructive degradation test data. In: Chen DG, Lio YL, Ng HKT, Tsai TR (eds) Statistical modeling for degradation data, chap. 12. Springer, New York
3. Hong Y, Xie Y, Jin Z, King C (2016) ADDT: a package for analysis of accelerated destructive degradation test data, R package version 1.1. [Online]. Available: http://CRAN.R-project.org/package=ADDT
4. Escobar LA, Meeker WQ, Kugler DL, Kramer LL (2003) Accelerated destructive degradation tests: data, models, and analysis. In: Lindqvist BH, Doksum KA (eds) Mathematical and statistical methods in reliability, chap. 21. World Scientific Publishing Company, River Edge
5. Li M, Doganaksoy N (2014) Batch variability in accelerated-degradation testing. J Qual Technol 46:171–180
6. Tsai C-C, Tseng S-T, Balakrishnan N, Lin C-T (2013) Optimal design for accelerated destructive degradation tests. Qual Technol Quantitative Manag 10:263–276
7. Xie Y, King CB, Hong Y, Yang Q (2017, in press) Semi-parametric models for accelerated destructive degradation test data analysis. Technometrics, doi:10.1080/00401706.2017.1321584
8. Vaca-Trigo I, Meeker WQ (2009) A statistical model for linking field and laboratory exposure results for a model coating. In: Martin J, Ryntz RA, Chin J, Dickie RA (eds) Service life prediction of polymeric materials, chap. 2. Springer, New York
9. King CB, Xie Y, Hong Y, Van Mullekom JH, DeHart SP, DeFeo PA (2017) A comparison of traditional and maximum likelihood approaches to estimating thermal indices for polymeric materials. J Qual Technol

Chapter 15
Modeling and Inference of CD4 Data

Shuang He, Chuanhai Liu, and Xiao Wang

Abstract CD4+ T lymphocyte cell count serves as a surrogate marker of human immunodeficiency virus (HIV) disease progression as it indicates how healthy the immune system is. It is well known that total CD4 counts degrade as the disease progress, so that the CD4 data can be treated as a special type of degradation data. It is also critical to identify HIV-resistant subjects with non-decreasing CD4 curves, i.e., the HIV/AIDS nonprogressors. In this paper, we propose a comprehensive statistical analysis strategy motivated and designed specifically to fit unique characters of the CD4 count data, which includes steps such as exploratory data analysis, modeling, and statistical inference. We formulate the problem into a simultaneous hypothesis testing problem. The inferential model (IM), a new prior-free probabilistic inference framework, is applied to locate HIV nonprogressors with the uncertainty characterization. We compared it with the classical Benjamini-Hochberg procedure for controlling false discovery rate (FDR).

Keywords Exploratory data analysis • FDR • Functional data analysis • Inferential model

15.1 Introduction

As one of the most important prognostics of the infections of human immunodeficiency virus (HIV), CD4 count has been investigated by many scientists from different aspects. It is well known that total CD4 counts degrade as the disease progress, so that the CD4 data can be treated as a special type of degradation data. Biologists make use of a wide range of well-established statistical methods, including but not limited to, univariate and multivariate regression, categorical data

S. He
Eli Lilly and Company, Lilly Corporate Center, 46285, Indianapolis, IN, USA
e-mail: he_shuang@lilly.com

C. Liu • X. Wang (✉)
Department of Statistics, Purdue University, 250 N University Street, 47907, West Lafayette, IN, USA
e-mail: chuanhai@purdue.edu; wangxiao@purdue.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_15

analysis and longitudinal data analysis, to model the CD4 count data and draw practical conclusions [1, 3, 5, 6]. In particular, some scientists are interested in studies the association between the time to developing AIDS and the baseline CD4 counts immediately after HIV infections, while others are interested in investigate the effectiveness of antirevtoviral therapy by looking at the slope of the change of CD4 counts. For example, [10, 11] developed mechanistic models using mathematical differential equations in succession to summarize dynamic changes of CD4 counts after antiretroviral treatment. Statisticians [4, 12, 14] used the CD4 count data to validate the capability and efficiency of intuitive methodologies designed specifically for sparse longitudinal data with measurement errors. In this paper, our objective is to identify HIV-resistant subject with non-decreasing CD4 curves, i.e., the HIV/AIDS nonprogressors. We propose a comprehensive statistical analysis strategy motivated and designed specifically to fit unique characters of the CD4 count data, which includes steps such as exploratory data analysis, modeling, and statistical inference.

Statistical modeling is an iterative process. It is fully reflected in this study focusing on investigating the CD4 data, which is a longitudinal data set with a random number of sparsely distributed repeated measurements of the CD4 counts taken at non-equal time intervals for each sample object. Tools of functional data analysis (FDA) are first employed to explore features of the CD4 data, both numerically and visually. In particular, we apply the nonparametric repeated Hanning method [13] to smooth the CD4 percentage, which is the sample CD4 count as compared to the whole number of lymphocyte cells. The amount of smoothing is determined when the estimated autocorrelation coefficients are stable or negligible [7]. By doing so, we have obtained conditionally independent CD4 percentage measurements and uncovered the real shape of CD4 curves by removing the random errors. After the CD4 percentage trend has been shown to be linear after smoothing, a simple linear model is fitted for each subject. A simultaneous multiple hypotheses of testing is formulated to test whether the CD4 slopes are positive or negative, with the intention to locate HIV-resistant subjects with non-decreasing CD4 curves, i.e., the HIV/AIDS nonprogressors. The inferential model (IM) [8], a new prior-free probabilistic inference framework, is applied to identify HIV nonprogressors with the uncertainty characterization. We compared it with the classical Benjamini-Hochberg procedure [2] for controlling false discovery rate (FDR).

This paper is organized as follows. In Sect. 15.2, we apply various exploratory data analysis techniques to understand, clean, and summarize the CD4 data. In Sect. 15.3, we make simultaneous inference on the slopes using both the IM and the well-know Benjamini-Hochberg procedure. The paper ends with conclusions in Sect. 15.4.

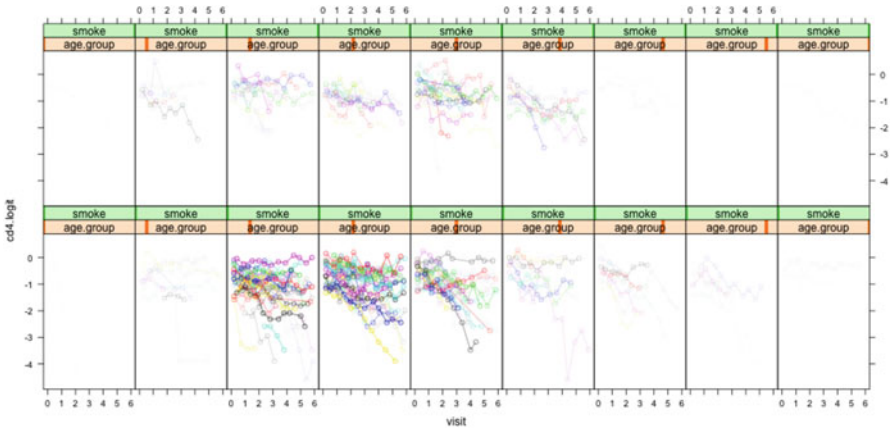

Fig. 15.1 Spaghetti plot of curves of the logit scale of the CD4 percentages over visits in years, stratified by age groups and smoke. The *top panels* and the *bottom panels* stand for smokers and nonsmokers, respectively

15.2 Exploratory Data Analysis

The CD4 dataset is a longitudinal collection of 1817 observations from 283 HIV-infected volunteers followed for up to 5.9 years. For each volunteer with a unique, arbitrary assigned id, the follow-up time in year post infection (visit), the smoking status of this id (smoke), the age of the id at the beginning of the study, the post-infection CD4 percentage (precd4), the CD4 percentage at the current visit (cd4), the time in year when the last CD4 count measurement taken (lt), the time in year when current CD4 count measurement taken (rt), and the CD4 percentage at previous visit were recorded. Among the 283 individuals, 68 (about one third) of them were smokers, and the youngest one and the oldest one were 18 years old and 60 years old at the time of HIV infection, respectively. The post-infection CD4 percentage ranged from 15 to 69, while at the end of the follow-up period for this CD4 data, the CD4 percentage decreased to the range of 7 to 61. The scatter plot of the data under different age and smoke status groups is shown in Fig. 15.1. To uncover the true function shapes of these curves, they have to be deconvoluted. For this purpose, we need to smooth the curves of the CD4 percentages over visits for each volunteer separately. As 3 is the minimum number to capture the pattern of a simple curve, id's with less than 3 visits are excluded from this study, resulting a sample size of 1766 visits from 223 unique id's. In addition, it is natural to apply a logit transformation on the CD4 percentage. The logit transformation will be particularly helpful later on when we performing simulation in that it won't be difficult to randomly generate the CD4 percentage values in a logit scale and ensure their values in the Euclidean space restricted between 0 and 1. In the following, whenever we mention the CD4 percentage, it stands for the logit scale of it.

The repeated Hanning method [13] was used for nonparametric smoothing due to its simplicity and flexibility. For volunteer with id j ($j = 1, 2, \ldots, 223$), denote CD4 percentages measured at visit i ($i = 1, 2, \ldots, n_j$) as $(y^0_{j,1}, y^0_{j,2}, \ldots, y^0_{j,n_j})$ and smooth it by repeated weighted averaging using the following formula

$$y^m_{j,i} = 0.25y^{m-1}_{j,i-1} + 0.5y^{m-1}_{j,i} + 0.25y^{m-1}_{j,i+1} \tag{15.1}$$

at the mth iteration, where the two endpoints are arbitrarily set as

$$y^{m-1}_{j,0} = y^{m-1}_{j,1}, \quad y^{m-1}_{j,n_j+1} = y^{m-1}_{j,n_j}. \tag{15.2}$$

On average, the CD4 percentages at visit 0, visit 1 and the last visit were 42, 35 and 26, respectively, suggesting a decreasing trend. In order to visualize the details of this decreasing trend, we first generated a spaghetti plot of curves of CD4 percentages over visits for each individual volunteers (Fig. 15.1). Volunteers were assigned into 9 age groups, including 15 to 20, 21 to 25, 26 to 30, 31 to 35, 36 to 40, 41 to 45, 46 to 50, 51 to 55 and 56 to 60, to stratify the change of CD4 percentage among volunteers with different ages. At the first impression, an overwhelming majority of the volunteers exhibited jagged decreasing curves of the CD4 percentages, while only a small amount of them remained flat or even increased. It is nearly impossible to tell whether these curves changed differently among different age groups and/or between smokers and nonsmokers from the spaghetti plot of raw CD4 percentage curves.

It is critical to denoise as much as possible to reveal the underlying functional form. On the other hand, we do not want to smooth too much in the way that some features of the function would be lost. Therefore, we borrow the concept of the conditional independence [7]. Let $(s_{j,1}, s_{j,2}, \ldots, s_{j,n_j})$ be the means of the observed CD4 percentages $(y_{j,1}, y_{j,2}, \ldots, y_{j,n_j})$ at each visit i for each id j, then the conditional independence is satisfied when

$$Pr\{y_{j,1}, y_{j,2}, \ldots, y_{j,n_j} | s_{j,1}, s_{j,2}, \ldots, s_{j,n_j}\} = \prod_{i=1}^{n_j} Pr\{y_{j,i} | y_{j,1}, y_{j,2}, \ldots, y_{j,n_j}\}. \tag{15.3}$$

To find the smallest number of iterations that ensure conditional independence, the first 5 sample autocorrelation coefficients are calculated and plotted against the Hanning iteration number. Iteration procedures were terminated whenever the first sample autocorrelation coefficient increased to positive. The maximum number of iterations required for this study was 164. Figures 15.1 and 15.2 are the spaghetti plots of the raw curves and the smoothed curves of the logit scale of the CD4 percentages over visits in years. As shown in Fig. 15.2, the spaghetti plot of the smoothed curves suggests a linear trend of the change of the CD4 percentage. While most of the linear curves are decreasing, it is still easy to identify some of lines which are flat or even increasing. However, it is not straightforward to quantify such assertions.

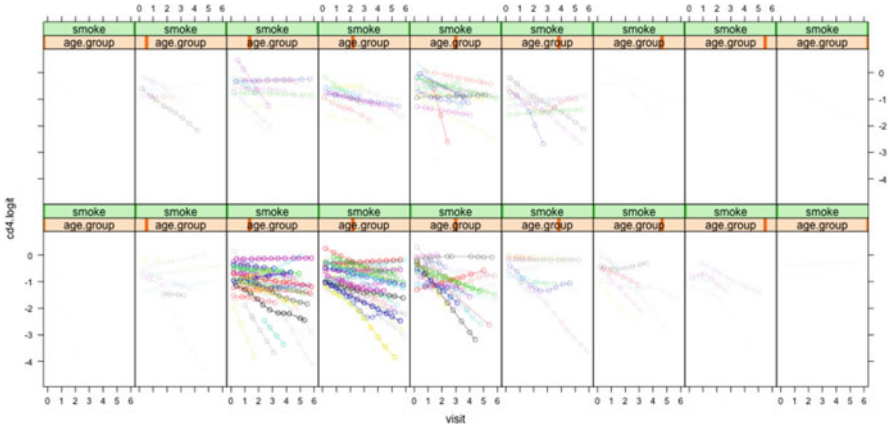

Fig. 15.2 Spaghetti plot of the logit scale of the smoothed curves of CD4 percentages over visits in years, stratified by age groups and smoke. The *top panels* and the *bottom panels* stand for smokers and nonsmokers, respectively

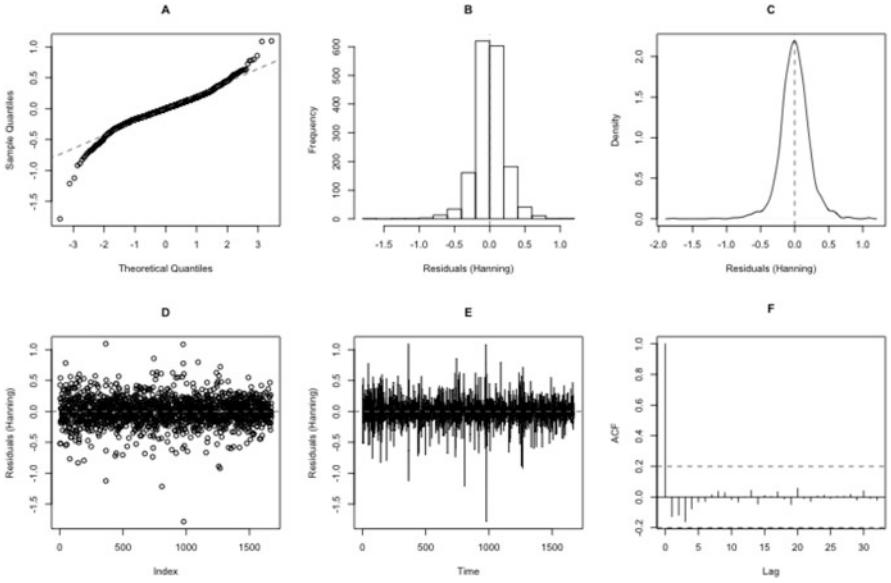

Fig. 15.3 QQ plot (**a**), histogram (**b**), density plot (**c**), index plot (**d**), time series plot (**e**) and plot of autocorrelation function (**f**) of residuals from smoothing the logit scale of the CD4 percentage against visit using the repeated Hanning Method. The average of the residuals are confirmed to be 0 (panel **b**, panel **c**, panel **d** and panel **e**). No sequential correlations among residuals (panel **d**, panel **e** and panel **f**) are observed

Figure 15.3 summarizes the distribution of the smoothing residuals, which are the differences between the logit scale of the CD4 percentage and their smoothed counterparts. Although it is not stringent to check the pooled residuals from all 223 id's, as the variance of errors could be different from individual to individual, Fig. 15.3 provides useful information regarding the CD4 data. As expected from the nature of the smoothing technique and conditional independence, this is a bell-shaped symmetric distribution with mean zero (panel B and panel C), and no obvious series dependency is observed (panel D, panel E and panel F). We observe a heavy tail of the smoothing residuals (panel A), suggesting the existence of some extreme values, which is also reasonable as the measurements of CD4 percentage values vary from people to people.

Based on the exploratory data analysis, we have reasoned that the linear function is appropriate to represent the degradation change of the CD4 percentage over time for each HIV-infected participant. We are moving forward to using the simple linear model to obtain the slope estimate for each linear function, so that we can make inference about HIV non-progressors by identifying those with non-negative slopes of the linear function. For a volunteer with id j ($j = 1, 2, \ldots, 223$), denote the repetitive measurements of CD4 percentages at the ith visit at year $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,i}, \ldots, x_{j,n_j})$ as the vector $\boldsymbol{y}_j = (y_{j,1}, y_{j,2}, \ldots, y_{j,i}, \ldots, y_{j,n_j})$. Then a simple linear model can be written as the follows

$$y_{j,k} = \beta_{j0} + \beta_{j,1} x_{j,k} + \epsilon_{j,k}, \quad k = 1, \ldots, n_j, \tag{15.4}$$

where $\boldsymbol{\beta}_j = (\beta_{j,0}, \beta_{j,1})$ are intercept and slope, respectively, and $\boldsymbol{\epsilon}_j = (\epsilon_{j,1}, \epsilon_{j,2}, \ldots, \epsilon_{j,n_j})$ is a vector of *i.i.d.* normally distributed errors with mean zero and equal variance σ_j^2. The slope $\beta_{j,1}$ and the intercept $\beta_{j,0}$ can be estimated by, respectively,

$$\hat{\beta}_{j,1} = \frac{\hat{Cov}(\mathbf{x}_j, \mathbf{y}_j)}{\hat{Var}(\mathbf{x}_j)}, \quad \hat{\beta}_{j,0} = \bar{y}_j - \hat{\beta}_{j,1}\bar{x}_j. \tag{15.5}$$

The residuals $\epsilon_{j,1}, \epsilon_{j,2}, \ldots, \epsilon_{j,n_j}$ can be obtained by

$$\hat{\epsilon}_{j,i} = y_{j,i} - (\hat{\beta}_{j,0} + \hat{\beta}_{j,1} x_{j,i}), \tag{15.6}$$

for each $i = 1, 2, \ldots, n_j$. The marginal distribution of intercept and that of slope are plotted in left and middle panels of Fig. 15.4 respectively. On one hand, intercept is symmetrically distributed about -0.5. The distribution of slope is skewed to the left, with a median of -0.128, showing that the majority of the HIV-infected volunteers experienced a decrease in the plasma CD4 count measurement during the follow-up period of the study, which is consistent with the smoothed Spaghetti curves shown in Fig. 15.2. Figure 15.4 also reflects the fact that a small portion of the CD4 percentage remained non-decreasing or even increasing during the study. The right panel of Fig. 15.4 shows the joint distribution of estimated intercepts and estimated slopes.

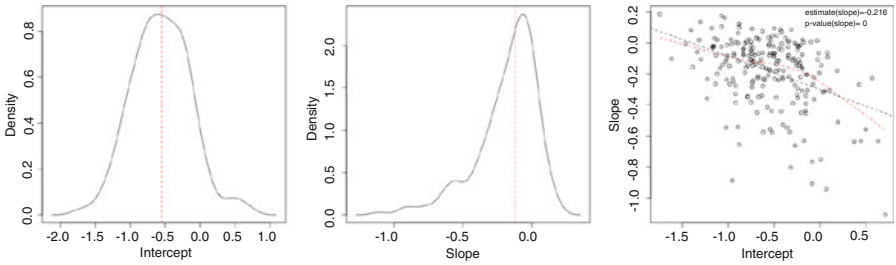

Fig. 15.4 *Left*: The marginal distribution of the estimated intercepts without 4 outliers. *Middle*: The marginal distribution of the estimated slopes without 4 outliers. *Right*: The joint distribution of the estimated slopes and the estimated intercepts without 4 outliers. The *red dashed line* is the LOWESS fitted curve and the *black dashed line* is the *straight line* fitted by a simple linear regression model

It is statistically significant that the slope is negatively correlated with the intercept, i.e., the higher the post-infection CD4 percentage, the slower the decreasing of the CD4 percentage. This reflects an implicit time misalignment phenomenon, that is, the time a patient's blood sample was confirmed to be infected by HIV in the lab was later than the true time. Since HIV/AIDS is a slow progression disease and usually shows no peculiar symptom other than a flu-like illness, a volunteer usually did not notice that he already infected by HIV. In fact, the true event of HIV infection could happen any time between visit 0 and the visit before visit 0, i.e., the true time after HIV infection should be longer than the recorded visit time. In general, the longer the disease had progressed before visit 0, the smaller the number of the intercept, which is the post-infection CD4 percentage. However, this time misalignment problem does not affect the inference of our objective in this study. We are interested in identifying nonprogressors whose CD4 count measurements remained nondecreasing during the clinical latency phase of the natural course of HIV/AIDS progression.

15.3 Statistical Inference

We have elaborated the validity of using a linear function to model the curve of CD4 percentage. The objective of identifying HIV/AIDS nonprogressors can be converted into a statistical simultaneous multiple testing problem for each slope of the CD4 percentage curve. Specifically, for each HIV-infected volunteer with id j ($j = 1, 2, \ldots, n$), assume a simple linear regression model

$$y_{j,i} = \beta_{j,0} + \beta_{j,1} x_{j,i} + \sigma_j Z_{j,i} \tag{15.7}$$

for each visit i ($i = 1, 2, \ldots, n_j$) and assume $Z_{j,i}$ follows a $N(0, 1)$ distribution i.i.d. Note that σ_j's do not need to be equal, allowing measurement errors for each

id to have different variances. Let $\beta = (\beta_{1,1}, \ldots, \beta_{n,1})$. Our objective is to make simultaneous inference for n assertions

$$A_j = \{\beta : \beta_{j,1} < 0\}, \quad j = 1, \ldots, n. \tag{15.8}$$

15.3.1 Inferential Model

It has been a long debate regarding valid statistical inference between frequentist and Bayesian. While one can argue that the frequentist's point estimate is meaningfulness, others can question that the noninformative Bayesian prior is actually informative. Many statisticians have tried to propose alternative approaches to generate probabilistic inference without using a prior. However, none of them can provide a solution of easy interpretation.

Recently, [8] have developed a new framework of probabilistic prior-free statistical inference, known as the inferential model (IM), through a three-step construction [9]. Let X be the observable data and θ be the unknown parameter of interest, the IM starts with introducing a random auxiliary variable U with a probability measure P_U and define the association $X = a(U, \theta)$. The form of the probability measure P_U is flexible. While we are sampling u from P_U, a collection of sets of candidate random values, $\Theta_x(u)$ can be obtained through each possible pair of (x, u) and the association a, i.e.,

$$\Theta_x(u) = \{\theta : x = a(u, \theta)\}. \tag{15.9}$$

Then a predictive random set (PRS) $S(U)$ is employed to predict u^* followed by the third step, which is to combine the association $X = a(U, \theta)$, the observed $X = x$, and the PRS $S(U)$ to transform the available information about u^* to the parameter space. Finally the inference about θ can be made by calculating the probability that the expanded set $\Theta_x(S) = \cup_{u \in S} \Theta_x(u)$ is a subset of the assertion A, which is defined as the belief function

$$bel_x(A) = P_S\{\Theta_x(S) \subseteq A | \Theta_x(S) \neq \emptyset\}. \tag{15.10}$$

Note that $bel_x(A)$ is not an ordinary additive probability measure, i.e., $bel_x(A) + bel_x(A^c) \leq 1$ with equality if and only if $\Theta_x(S)$ is a singleton with probability 1. The related plausibility function is introduced as the following definition

$$pl_x(A) = 1 - bel_x(A^c) = P_S\{\Theta_x(S) \nsubseteq A^c | \Theta_x(S) \neq \emptyset\}. \tag{15.11}$$

From Eqs. (15.10) and (15.11), it is easy to conclude that the value of $bel_x(A)$ is always no bigger than that of $pl_x(A)$. In other words, the belief function summarizes the evidence supporting A, while the plausibility function summarizes the evidence against A^c. If the sum of these two numbers does not equal to 1, the left represents

the "don't know". In together, the belief function $bel_x(A)$ and the plausibility function $pl_x(A)$ provide a post-data, prior-free, probabilistic-adjusted measure of uncertainty of the assertion.

Let $\boldsymbol{Y}_j$ be the vector of $(y_{j,1}, y_{j,2}, \ldots, y_{n_j})^T$ and $\boldsymbol{X}_j \in \mathbb{R}^{2\times n_j}$ be the design matrix of the above regression model. The hat matrix is $\boldsymbol{H}_j = \boldsymbol{X}_j^T(\boldsymbol{X}_j\boldsymbol{X}_j^T)^{-1}\boldsymbol{X}_j$. Define

$$S_j = \frac{\|(\boldsymbol{I} - \boldsymbol{H}_j)\boldsymbol{Y}_j\|_2}{\sqrt{n_j - 2}}, \quad \boldsymbol{B}_j = (\boldsymbol{X}_j^T\boldsymbol{X}_j)^{-1}\boldsymbol{X}_j^T\boldsymbol{Y}_j \tag{15.12}$$

to be the least squares estimates of σ_j and $\boldsymbol{\beta}_j = (\beta_{j,0}, \beta_{j,1})$, respectively. Set

$$M_j = \frac{\|(\boldsymbol{I} - \boldsymbol{H}_j)\boldsymbol{Z}_j\|_2}{\sqrt{n_j - 2}}, \quad \boldsymbol{T}_j = \frac{(\boldsymbol{X}_j^T\boldsymbol{X}_j)^{-1}\boldsymbol{X}_j^T\boldsymbol{Z}_j}{M_j}. \tag{15.13}$$

It is known that M_j and $\boldsymbol{T}_j$ are independent and $(n_j - 2){M_j}^2$ follows a chi-square distribution with n_j-2 degrees of freedom and $\boldsymbol{T}_j$ follows a 2-d Student-t distribution with the scale matrix $(\boldsymbol{X}_j^T\boldsymbol{X}_j)^{-1}$ and $n_j - 2$ degrees of freedom. We may write the equivalent association

$$\boldsymbol{B}_j = \boldsymbol{\beta}_j + S_j\boldsymbol{T}_j, \quad S_j = \sigma_j M_j. \tag{15.14}$$

Therefore, we focus on the association

$$B_{j,1} = \beta_{j,1} + S_j T_{j,1}, j = 1, 2, \ldots, n. \tag{15.15}$$

For simplicity, we rewrite Eq. (15.15) as $B_j = \beta_j + S_j T_j, j = 1, 2, \ldots, n$. For each id j, consider the association

$$B_j = \beta_j + S_j {F_j}^{-1}(U_j), \tag{15.16}$$

where the (B_j, S_j) are observables, the unknown parameter $\beta = (\beta_1, \beta_2, \ldots, \beta_n)^T \in \mathbb{R}^n$, $U_1, U_2, \ldots, U_n$ are i.i.d. uniform auxiliary variable on $[0, 1]$, and F_j is the cumulative distribution function of the Student t-distribution with $n_j - 2$ degrees of freedom.

If $U_1, U_2, \ldots, U_n$ are known, the values of β_j are determined by

$$\beta_j = B_j - S_j {F_j}^{-1}(U_j). \tag{15.17}$$

So it is very critical to predict these unknown quantities $U_1, U_2, \ldots, U_n$. Under the frame work of IM, we use the predictive random set (PRS) to predict U_j's. The criterion to choose PRS is to balance across multiple assertions [8]. The optimal PRS satisfying the balance condition is the square PRS:

$$\varsigma = \{u \in [0, 1]^n : max\{|u_1|, \ldots, |u_n|\} \geq max\{|U_1|, \ldots, |U_n|\}\}. \tag{15.18}$$

Table 15.1 Nonprogressors identified by the IM, as compared by the Benjamini-Hochberg procedure at level $\alpha = 0.05$

j	**id**	$bel(A_j)$	$pl(A_j)$	$ph(A_j)$
13	1631	0.00	1.00	0.00
70	3708	0.23	1.00	0.00
87	4122	0.00	1.00	0.00
106	4868	0.99	1.00	0.01
119	5208	0.00	1.00	0.00
151	7143	0.00	1.00	0.00
161	7317	0.00	1.00	0.00
162	7398	0.00	1.00	0.00
185	8341	0.00	1.00	0.00
211	9451	0.00	1.00	0.00

where $U_1, U_2, \ldots, U_n$ are i.i.d. with $Unif(0, 1)$. With this PRS, let

$$\Theta(\varsigma) = \{\beta : \beta_j = B_j - S_j F_j^{-1}(U_j), j = 1, 2, \ldots, n\}. \tag{15.19}$$

If we are willing to accept that the PRS ς is satisfactory for predicting U, then $\Theta(\varsigma)$ will do equally well at capturing β. For a common PRS ς, IM results are given by the set of the n belief functions and n plausibility functions for each A_j. For any given $\beta \in \mathbb{R}^n$, the belief function and the plausibility function are given by

$$bel(A_j) = Pr_\varsigma(\Theta(\varsigma) \subseteq A_j) \tag{15.20}$$

and

$$pl(A_j) = Pr_\varsigma(\Theta(\varsigma) \nsubseteq A_j^c), \tag{15.21}$$

respectively. There is no closed form to compute $bel(A_j)$ or $pl(A_j)$ directly from the above two equations. Instead, we randomly generated 10,000 n-dimensional samples $u \in [0, 1]^n$ to construct a realization of the PRS ς and then calculate the corresponding $\Theta(\varsigma)$. Next, we calculate the belief functions $bel(A_j)$'s and the plausibility functions $pl(A_j)$'s by constructing 1000 realizations of $\Theta(\varsigma)$ and computing the proportion of simulated $\Theta(\varsigma)$'s with all elements no less than 0 and the proportion of simulated $\Theta(\varsigma)$'s with at least one element less than 0, respectively. All $bel(A_j)$'s are less than or equal to $pl(A_j)$'s. providing the lower probabilities and the upper probabilities to evaluate the uncertainties of assertion A_j's. Among all of the ids, 8 of them, including 1631, 4122, 5208, 7143, 7317, 7398, 8341 and 9451, are identified as nonprogressors by evaluating their belief functions $bel(A_j) \leq \alpha = 0.05$ (Table 15.1).

15.3.2 False Discovery Rate (FDR)

Benjamini and Hochberg [2] proposed a less conservative definition of the error rates, the false discovery rate (FDR), and described the procedure of how to obtain the FDR. The FDR procedure has been shown to associated with greater power, though, at the meanwhile, it also inflates the rates of false discoveries. The FDR, which is defined as the expected proportion of false positives among all rejected null hypotheses, has been widely applied in multiple testing problems, particularly in those high-throughput data. The reason for the prevalence of FDR is that it is associated with greater power as compared to the classical approaches that controlling the family-wise error rate, which is the rate of reject even a single true null hypothesis by mistake. The FDR was proven less or equal to the significance level α. In this section, we elaborate the use of the Benjamini and Hochberg procedure to identify HIV/AIDS nonprogressors.

Define V and R the number of false positives and the number of total rejected null hypothesis, then the FDR is defined as

$$FDR = \mathbb{E}[\frac{V}{R}], \tag{15.22}$$

where $\frac{V}{R} = 0$ when $R = 0$. For n independent hypotheses H_m $(m = 1, 2, \ldots, n)$, the Benjamini-Hochberg procedure controls the FDR at a given level α through the following steps: (1). Order the p-values for each test as $p_{(1)}, p_{(2)}, \ldots, p_{(n)}$; (2). Find the largest k such that $p_{(k)} \leq \frac{k}{m}\alpha$; (3). Reject all null hypotheses with $p_{(i)}$ for $i = 1, \ldots, k$. It is proved that $FDR \leq \frac{n_0}{n}\alpha \leq \alpha$.

For each HIV-infected volunteer with id $j(j = 1, 2, \ldots, n)$, we first conducted the one-sided test of slope

$$H_{0,j} : \beta_j \leq 0 \qquad vs \qquad H_{a,j} : \beta_j > 0. \tag{15.23}$$

The observed values of the test statistic $t_{obs,j} = B_j/S_j$ and the corresponding p-values p_j based on its asymptotic Student-t distribution with $(n_j - 2)$ degrees of freedom are computed. We applied the Benjamini-Hochberg procedure by first ordering the p-values for each test as $p_{(1)}, p_{(2)}, \ldots, p_{(n)}$, and then calculate the adjusted p-values by

$$ph(A_j) = min\{1, cummin(np_{(j)}/j)\}. \tag{15.24}$$

Ten out of all HIV-infected volunteers, including those with id equal to 1631, 3708, 4122, 4868, 5208, 7143, 7317, 7398, 8341 and 9451, are identified as nonprogressors at significance level $\alpha = 0.05$, i.e., $ph(A_j) \leq \alpha = 0.05$ (Table 15.1).

It is observed that the n FDR corrected p-values $ph(A_j)$'s are always no less than the belief functions $bel(A_j)$'s and no bigger than the plausibility functions $bel(A_j)$'s (Table 15.1). The belief functions $bel(A_j)$'s, the plausibility functions $pl(A_j)$'s,

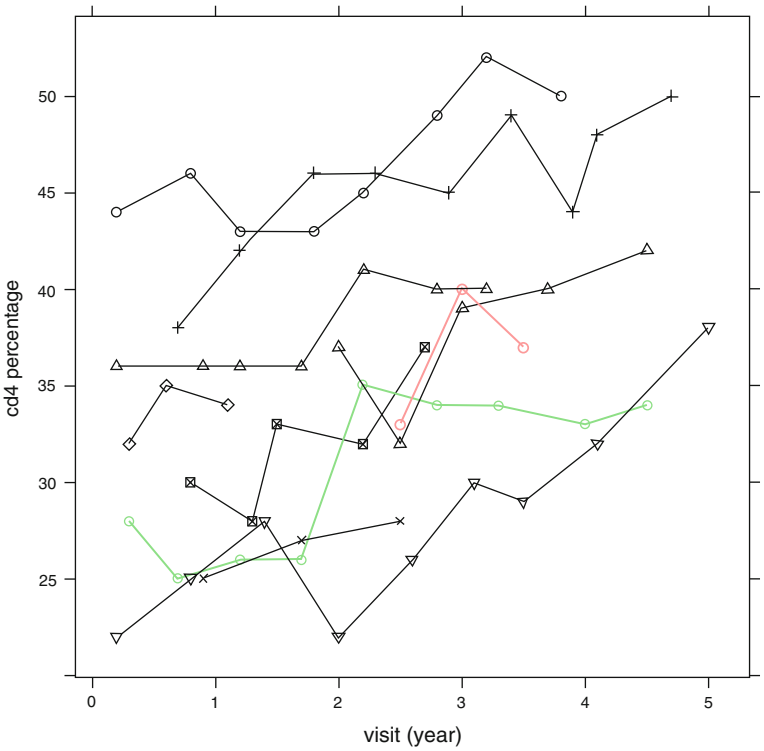

Fig. 15.5 Spaghetti plots of the CD4 percentages (in original scale, not logit scale) of HIV/NIDS nonprogressors identified by IM and/or FDR. All 8 curves in *black* were nonprogressors identified by using both the IM procedure and the Benjamini-Hochberg procedure. The curve in *red* and the curve in *green* show the 2 nonprogressors, with $id = 4868$ and $id = 3708$, respectively, stand for nonprogressors identified by the Benjamini-Hochberg procedure as nonprogressors, but not the IM procedure. Visualization of this plot confirms that all the 10 nonprogressors identified by IM and/or FDR express increasing CD4 percentages

and the FDR corrected p-values $ph(A_j)$'s are also reported for the 10 HIV/AIDS nonprogressors identified using FDR with level $\alpha = 0.05$ (Table 15.1). The plausibility functions of all 10 of them were all 1. The belief functions of 8 patients including 1631, 4122, 5208, 7143, 7317, 7398, 8341 and 9451, were less than 0.05. In fact, they were all equal to 0.00. The rest two of them, with $id = 3708$ and $id = 4868$, have the values of the belief functions equaled to 0.23 and 0.99, respectively. If we also choose level $\alpha = 0.05$ to identify HIV/AIDS nonprogressors based on the plausibility function, similar but not exactly the same conclusions would be drawn as compared to the Benjamini-Hochberg procedure. That is, the 8 id's mentioned above were also identified by the Benjamini-Hochberg procedure. However, the Benjamini-Hochberg procedure identified two extra volunteers as nonprogressor. The CD4 percentage of the above 8 plus 2 HIV/AIDS nonprogressors was plotted in black and red in Fig. 15.5, respectively. All of them show increasing trends of

their CD4 percentage curves. The key difference between the IM and the FDR is that the IM always provides a way of measuring uncertainties of the assertion directly through the belief functions and the plausibility functions, while as a post hoc adjustment of multiple p-values, the FDR procedure does not provide any uncertainty justification.

15.4 Conclusion

This paper comprises of two main parts, the data and objective driven analysis strategy and the application of the inferential model in functional data analysis. As illustrated in this study, a good exploratory data analysis, along with a thorough background study, can not only help understand the data structure, visualize the distribution of random variables and check model assumptions, but also be able to reveal the problems such as missing value and curve registration. In addition, the exploratory data analysis is even more important in functional data analysis, particularly in that the application of nonparametric smoothing techniques is critical in both uncovering the underlying function of curves and hence affecting the selection of data analysis and inference methods.

The construction of the inferential model is novel in the way that intrinsically, both observations and parameters are fixed and auxiliary variables count for the source of randomness, so that we are not arbitrarily coerce the sample and/or the parameter to be random. The thinking of IM is more of the nature of statistical inference. Furthermore, it is easy to see from the construction of the IM that it targets on the scientific problem directly and provides the uncertainties about multiple assertions simultaneously, through the lower and upper probabilities, $bel_x(A)$ and $pl_x(A)$. It separates the uncertainty and decision. Scientists will make the decision, not statisticians. On the other hand, the Benjamini-Hochberg procedure is a post hoc correction for multiple p-values controlling the false discovery rate, which is not a direct evaluation of the probability of uncertainty itself. Other advantages of the IM include, but not limited to, that the IM is finite sample inference which comes from data with a prior-free probabilistic calibration. In a word, the IM provides a post-data and prior-free probabilistic measure of the uncertainty of an assertion or a set of assertions simultaneously.

There are many potentials to other extensions, including: (1) Identifying long-term nonprogressors, also known as elite progressors, by applying to a larger CD4 data set which contains longer duration of follow-up, with or without medical intervention; (2) Investigating the associations between HIV/AIDS nonprogressors with a variety of demographical, socioeconomics, and/or health variables as the independent variables by performing a logistic regression with nonprogressors as the dependent variable and other covariates as independent variables.

Acknowledgements Xiao Wang's research is supported by National Science Foundation under grant DMS-1613060.

References

1. Anastos K, Kalish LA, Hessol N, Weiser B, Melnick S, Burns D, Delapenha R, DeHovitz J, Cohen M, Meyer W, Bremer J, Kovacs A (1999) The relative value of CD4 cell count and quantitative HIV-1 RNA in predicting survival in HIV-1-infected women: results of the women's interagency HIV study. AIDS 13:1717–1726
2. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300
3. Bofill M, Janossy G, Lee CA, MacDonald-Burns D, Phillips AN, Sabin C, Timms A, Johnson MA, Kernoff PB (1992) Laboratory control values for CD4 and CD8 T lymphocytes. Implications for HIV-1 diagnosis. Clin Exp Immunol 88(2):243–52
4. Boscardin WJ, Taylor JM, Law N (1998) Longitudinal models for AIDS marker data. Stat Methods Med Res 7:13–27
5. Castilla J, del Romero J, Hernando V, Marincovich B, Garcia S, Rodriguez C (2005) Effectiveness of highly active antiretroviral therapy in reducing heterosexual transmission of HIV. J Acquir Immune Defic Syndr 40:96–101
6. Easterbrook PJ, Yu LM, Goetghebeur E, Boag F, McLean K, Gazzard B (2000) Ten-year trends in CD4 cell counts at HIV and AIDS diagnosis in a London HIV clinic. AIDS 14(5):561–571
7. Lambert D, Liu C (2006) Adaptive thresholds: monitoring streams of network counts. J Am Stat Assoc 101(473):78–88
8. Liu C, Martin R (2015) Inferential models: reasoning with uncertainty. Monographs in statistics and applied probability series. Chapman-Hall/CRC Press, Boca Raton
9. Martin R, Liu C (2013) Inferential models: a framework for prior-free posterior probabilistic inference. J Am Stat Assoc 108(501):201–313
10. Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, Markowitz M, Ho DD (1997) Decay characteristics of HIV-1-infected compartments during combination therapy. Nature 387:188–191
11. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582–1586
12. Shi M, Weiss R, Taylor J (1996) An analysis of pediatric CD4 counts for acquired immune deficiency syndrome using flexible random curves. Appl Stat 45:151–164
13. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading
14. Yao F, Müller HG, Wang JL (2005) Functional data analysis for sparse longitudinal data. J Am Stat Assoc 100:577–590

Chapter 16
State Space Models Based Prognostic Methods for Remaining Useful Life Prediction of Rechargeable Batteries

Dong Wang and Kwok-Leung Tsui

Abstract Rechargeable batteries are widely used in many electronic products and systems to provide power sources. Because of the influence of charge/discharge cycling and some significant battery degradation factors, such as discharge rate, temperature, depth of charge, etc., on battery health condition, battery degrades over time. In this chapter, several state space models based prognostic methods are proposed to predict battery remaining useful life. Firstly, a particle filtering based state space model for battery remaining useful life prediction at a constant discharge rate is introduced. Secondly, to improve particle filtering and its application to battery prognostics, spherical cubature Kalman filtering is introduced to provide an importance function for the use of particle filtering at a constant discharge rate. Thirdly, to extend battery prognostics at a constant discharge rate to battery prognostics at different discharge rates, a more general battery degradation model is presented. Based on the developed model, a battery prognostic method at different discharge rates is designed. Some discussions are made at last.

Keywords Prognostics and health management • Rechargeable batteries • State space model • Remaining useful life • Filtering • Degradation

16.1 Introduction

Prognostics and health management (PHM) is an enabling discipline that aims to predict the time when a component or system will no longer satisfy its functionality required by users, to schedule a necessary maintenance and to keep the high reliability of the component or system [1, 2]. Being different from traditional handbooks based reliability prediction methods, the PHM incorporates timely measurements collected from the component or system to update parameter distributions

D. Wang (✉) • K.-L. Tsui
Department of Systems Engineering and Engineering Management,
City University of Hong Kong, Hong Kong, China
e-mail: dongwang4-c@my.cityu.edu.hk; kltsui@cityu.edu.hk

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_16

of empirical degradation models established from historical data or physical models provided by domain knowledge, and to infer remaining useful life (RUL) of the product or system [3]. Here, the RUL [4] is defined as the period from the current time to the end of useful lifetime. Rechargeable batteries, especially lithium-ion batteries, are one of the most common energy storage devices used in electronic vehicles, electronic products, unmanned aerial vehicles, etc. [5, 6, 7]. Battery failures may cause unexpected system breakdowns, increase economic losses and even catastrophe. To ensure the high reliability of rechargeable batteries, their RUL, namely remaining battery capacity or remaining charge/discharge cycles, should be accurately and immediately predicted before their failures [8].

At the very early beginning, Burgess [9] experimentally investigated the float service life of a valve regulated lead acid battery and found that it has two distinct stages. The former stage of the float service life of the valve regulated lead acid battery has a relatively small battery capacity loss, while the latter stage of the float service life of the valve regulated lead acid battery has a linear battery degradation trend that can be modeled by using a linear and Gaussian state space model. Then, Burgess used Kalman filtering [10] to solve the linear and Gaussian state space model and predicted RUL of the valve regulated lead acid battery. However, for many other different types of rechargeable batteries, such as lithium-ion batteries, battery degradation trends are not purely linear and they can not be well modeled by the linear and Gaussian state space model, which limits the use of Kalman filtering for battery prognostics. To solve this problem, Saha et al. [11] combined relevance vector machine with particle filtering (a general sequential Monte Carlo method) to predict RUL of lithium-ion batteries. To show the superiority of their proposed method, comparisons with autoregressive integrated moving average and extended Kalman filtering were conducted. In their finding, particle filtering has higher prediction accuracies than the other prognostic methods [12]. The main idea of particle filtering [13] is to use an amount of random particles sampled from an importance function and their associated weights to estimate posterior distributions of parameters or states. Consequently, particle filtering is more flexible than Kalman filtering and its variants for solving nonlinear and non-Gaussian state space models. Since the work of Saha et al. was reported for battery prognostics, particle filtering based prognostic methods became popular in the research community of diagnostics and prognostics [14, 15]. He et al. [16] used the sum of two exponential functions, namely a bi-exponential function, as an empirical battery capacity degradation model to fit battery degradation data and employed particle filtering to predict RUL of lithium-ion battery. However, it should be noted that the sum of the two exponential functions is too flexible to fit many degradation data and thus may cause an overfitting problem that connotes poor prediction accuracies [17]. Following the work of He et al., Xing et al. [18] employed the combination of an exponential function and a polynomial function with an order of 2 to fit battery degradation data and used particle filtering to predict battery RUL. Even though a new empirical battery degradation model was developed, no evidence can be used to support that such new empirical battery degradation model can be applied to general battery degradation data. Xian et al. [19] combined a Verhulst model,

particle swarm optimization and particle filtering to predict battery RUL. Li and Xu [20] mixed Gaussian process models with particle filtering to predict battery RUL. Dong et al. [21] used support vector regression to resample particle filtering to predict battery RUL. Liu et al. [22] employed autoregressive model and particle filtering to predict battery RUL. Hu et al. [23] used kernel smoothing to increase the diversity of random particles used in particle filtering for battery prognostics. Besides the above work, a recent comparison [24] showed that particle filtering is able to provide higher prediction accuracies than unscented Kalman filtering and non-linear least squares for battery prognostics [25]. Although particle filtering has proven a good prediction ability in battery prognostics [11, 12, 16, 18–22, 24], it is often assumed that its importance function is equal to a prior distribution that does not consider the recent battery capacity degradation data to update its own distribution and the weights of random particles [26]. Moreover, it should be noted that the above prognostic methods can only work at a fixed operating condition, such as a fixed discharge rate. In other words, the authors of the above works did not propose a unique prognostic model that is able to work at different or varying discharge rates. In real applications, the unique prognostic model working at different or varying discharge rates is more preferable because many electric and hybrid electric systems, such as electric vehicles, require different or varying discharge rates. Consequently, it is necessary to develop such unique prognostic model for battery prognostics.

For the rest of the book chapter, three cutting-edge state space models based prognostic methods are introduced to predict battery RUL and to solve the two challenging problems mentioned in the previous paragraph. In Sect. 16.2, based on historical battery degradation data, a state space model is presented to describe battery degradation over cycle. How to solve the state space model is then detailed by using particle filtering. Extrapolations of the determined state space model to a soft failure threshold are used to infer battery RUL. Here, the soft failure threshold is a flexible failure threshold defined by users according to historical degradation data. A case study is investigated to illustrate how the presented prognostic method works for battery RUL prediction. However, even though the particle filtering based prognostic method is effective in predicting battery RUL, as mentioned in the previous paragraph, an important function must be properly designed so that the recent battery degradation data can be incorporated to update the weights of random particles for battery RUL prediction. In Sect. 16.3, spherical cubature Kalman filtering is introduced to provide a proper importance function for use of particle filtering. Based on this idea, a spherical cubature particle filtering based prognostic method is presented to predict RUL of lithium-ion batteries. Comparisons with the particle filtering based prognostic method are conducted to demonstrate that the proper selection of an importance function is able to enhance prediction accuracy of battery RUL. In Sect. 16.4, besides the two prognostic methods that can only work at a constant discharge rate, a state space model based prognostic method is presented to predict battery RUL at different discharge rates. In other words, the third prognostic method is able to use a unique state space model to represent battery capacity degradation at different discharge rates, which is much more attractive in

practice because electric products and systems work at different or varying discharge rates. Some discussions are made in Sect. 16.5. Conclusion remarks are given at last.

16.2 A Particle Filtering Based State Space Model for Battery Remaining Useful Life Prediction at a Constant Discharge Rate

Before particle filtering is introduced to predict battery RUL at a constant discharge rate, a state space model should be developed to describe battery degradation over cycle. In this section, the capacity degradation samples used by Park et al. [27] are investigated in this section. 26 lithium-ion batteries were randomly chosen from manufactured lots and they went through cycle-life tests of 400 cycles under a constant-current/constant-voltage mode at a room temperature. The capacity of each battery was calculated by integrating current over time in the battery discharge process. The normalized capacities (NC) of 26 battery degradation samples are plotted in Fig. 16.1, where the 26 battery degradation samples are divided into training samples (blue dots) and testing samples (red stars). The training samples are used to establish an empirical battery degradation model and construct a state space model for battery degradation assessment. The testing samples are used to validate the effectiveness of the particle filtering based prognostic method for battery RUL

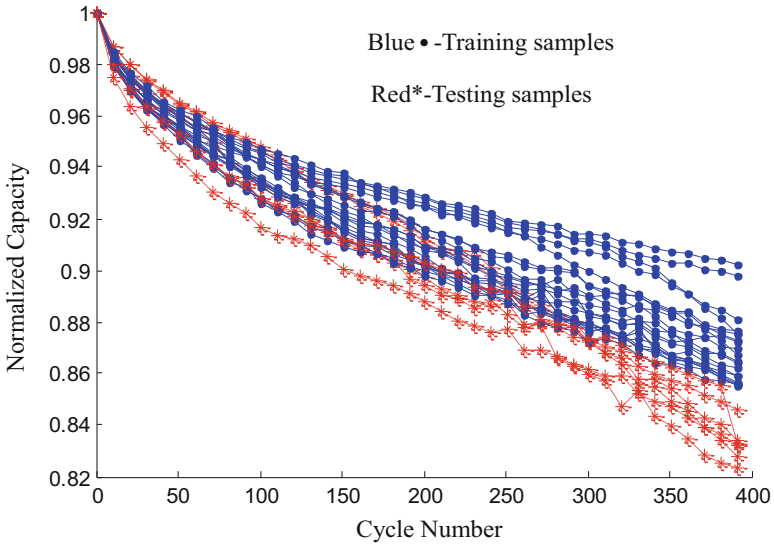

Fig. 16.1 26 lithium-ion batteries capacity degradation samples under a constant discharge rate

Table 16.1 Goodness of fit of the exponential function for the 20 training samples

Training sample number	R-squared value	Training sample number	R-squared value	Training sample number	R-squared value
1	0.98	8	0.97	15	0.96
2	0.98	9	0.97	16	0.96
3	0.98	10	0.97	17	0.94
4	0.98	11	0.96	18	0.96
5	0.97	12	0.96	19	0.96
6	0.97	13	0.96	20	0.99
7	0.98	14	0.94		

prediction. Because only 6 samples are able to reach a soft failure threshold of 0.85, the 6 capacity degradation samples are used as the testing samples.

In Fig. 16.1, we observe that these battery capacity degradation samples have similar degradation trends along with different degradation rates. A simple and useful function, namely an exponential model with the amplitude a and the slope b, is used to fit the training samples and describe the battery capacity degradation. The goodness of fit of the exponential model is indicated by the R-square values tabulated in Table 16.1, where we observe that the exponential function can describe the 20 training samples well. Thus, we can treat the exponential function as part of a measurement function used in a state space model. The boxplots of the fitted amplitudes and slopes of the exponential function are plotted in Fig. 16.2, in which we observe that different battery degradation samples have different amplitudes and slopes. To describe the diversity of the fitted exponential functions, the amplitude a and the slope b should be the two states/parameters used in a state space model. Considering the above two points, the state space model consisting of the measurement function (the nonlinear exponential function $f(\cdot)$ plus a noise term) and the two states is constructed as follows:

$$\mathrm{NC}_k = f(k) + v_1 = a_k \times \mathrm{e}^{b_k \times k} + v_1. \tag{16.1}$$

$$a_k = a_{k-1} + v_2, \tag{16.2}$$

$$b_k = b_{k-1} + v_3, \tag{16.3}$$

where v_1, v_2 and v_3 follow additive Gaussian noises with zero means and standard deviations σ_1,σ_2 and σ_3, respectively; and NC_k is the kth measured normalized capacity as mentioned in introduction of the experiment in this section. The initial states a_0 and b_0 are calculated as 0.9692 and −0.3028, respectively, by taking the medians of the a and b values shown in Fig. 16.2. σ_1 , σ_2 and σ_3 are set to 5×10^{-3}, 1×10^{-3} and 1×10^{-6}, respectively, by considering the scales and uncertainties of the normalized capacities, and the a and b values shown in Fig. 16.2.

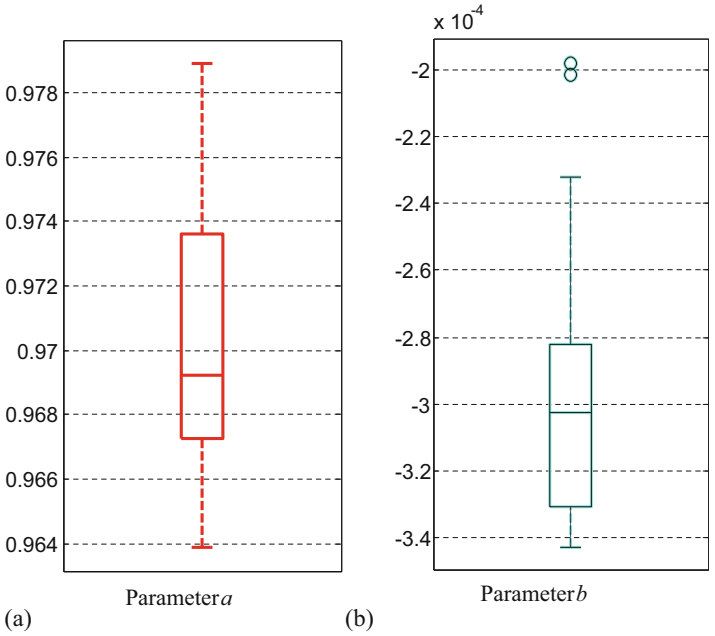

Fig. 16.2 Boxplots of the two states/parameters obtained by fitting the 20 training samples with the exponential function: (**a**) the parameter a; (**b**) the parameter b

Once the state space model is initialized, its states including the amplitude and slope of the exponential function should be posteriorly updated when a new battery capacity measurement is available. Therefore, Bayesian filtering, such as particle filtering, is able to posteriorly estimate the distribution of the states used in the state space model. The main idea of particle filtering is to use some random particles sampled from an importance function and their associated weights to represent a posterior probability density function. We will introduce how to use particle filtering to estimate the posterior distribution of the states of the exponential function in the following paragraphs.

Step 1. Generate N_s initial random particles $\left\{a_0^i\right\}_{i=1}^{N_s}$ and $\left\{b_0^i\right\}_{i=1}^{N_s}$ from the prior Gaussian distributions $\mathrm{p}\left(a_0\right) = \mathrm{N}\left(a_0, \sigma_2^2\right)$, and $\mathrm{p}\left(b_0\right) = \mathrm{N}\left(b_0, \sigma_3^2\right)$, respectively. N_s is artificially equal to 100 here. As the number of random particles increases, an approximation to a true posterior state distribution is improved. Nevertheless, associated computational time also increases. It is necessary to make a tradeoff between approximation accuracy and computational time. Then, each of the $\left\{\omega_0^i\right\}_{i=1}^{N_s}$ is set to $1/N_s$ and their sum is equal to 1. The random particles and their associate weights are able to approximate the prior distributions of the amplitude a and slope b. According to the theory of particle filtering [25], once a new measurement NC_1, namely the first measured normalized capacity, is available, the weight updating is calculated by:

$$\omega_1^i = \frac{\omega_0^i \dfrac{\mathrm{p}\left(\mathrm{NC}_1 \left| a_1^i, b_1^i \right.\right) \mathrm{p}\left(a_1^i \left| a_0^i \right.\right) \mathrm{p}\left(b_1^i \left| b_0^i \right.\right)}{\mathrm{q}\left(a_1^i, b_1^i \left| \mathrm{NC}_1, a_0^i, b_0^i \right.\right)}}{\sum_{i=1}^{N_s} \omega_0^i \dfrac{\mathrm{p}\left(\mathrm{NC}_1 \left| a_1^i, b_1^i \right.\right) \mathrm{p}\left(a_1^i \left| a_0^i \right.\right) \mathrm{p}\left(b_1^i \left| b_0^i \right.\right)}{\mathrm{q}\left(a_1^i, b_1^i \left| \mathrm{NC}_1, a_0^i, b_0^i \right.\right)}}. \tag{16.4}$$

where $\mathrm{p}\left(a_1^i \left| a_0^i \right.\right)$ and $\mathrm{p}\left(b_1^i \left| b_0^i \right.\right)$ are the state transition probability density functions established by Eqs. (16.2) and (16.3); $\mathrm{q}\left(a_1^i, b_1^i \left| \mathrm{NC}_1, a_0^i, b_0^i \right.\right)$ is the importance function designed by users; and $\mathrm{p}\left(\mathrm{NC}_1 \left| a_1^i, b_1^i \right.\right)$ is the measurement probability density function. The random particles and their associated weights provided by Eq. (16.4) are able to express the posterior distributions of the amplitude a and slope b. Here, it should be noted that ω_1^i can not be calculated if a_1^i and b_1^i are not predicted from a_0^i and b_0^i according to Eqs. (16.2) and (16.3), respectively. This is because ω_1^i is posteriorly and iteratively updated. If the importance function is chosen as the prior state transition probabilities, Eq. (16.4) is simplified as:

$$\omega_1^i = \frac{\omega_0^i \mathrm{p}\left(\mathrm{NC}_1 \left| a_1^i, b_1^i \right.\right)}{\sum_{i=1}^{N_s} \omega_0^i \mathrm{p}\left(\mathrm{NC}_1 \left| a_1^i, b_1^i \right.\right)}. \tag{16.5}$$

It should be noted that the importance function used in Eq. (16.5) does not consider the recent battery degradation data NC_1 to update the weights of random particles. According to the size of ω_1^i, a systematic resampling [28] is used to re-draw random particles $\left\{a_1^i\right\}_{i=1}^{N_s}$ and $\left\{b_1^i\right\}_{i=1}^{N_s}$ from the posterior distributions of the amplitude a and slope b. First, construct the cumulative distribution function of the weights $\omega_1^i, i = 1, 2, \ldots, N_s$. Let $c_1 = 0$ and $c_i = c_{i\text{-}1} + \omega_1^i$. Draw a starting point d_1 from a uniform distribution $\mathrm{U}\left[0, N_s^{-1}\right]$. For each increased point $d_j = d_1 + N_s^{-1}(j-1), j = 1, 2, \ldots, N_s$ moving along the cumulative distribution function of the weights, if $d_j \geq c_i$ is satisfied, $i = i + 1$; and then, $a_1^j = a_1^i$ and $b_1^j = b_1^i$. After the systematic resampling is conducted, each of the weights is set to $1/N_s$. The posterior probability density functions of the amplitude a and slope b at iteration 1 are respectively expressed as follows:

$$\begin{aligned} \mathrm{p}\left(a_1 \left| \mathrm{NC}_1 \right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(a_1 - a_1^i\right)/N_s \\ \mathrm{p}\left(b_1 \left| \mathrm{NC}_1 \right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(b_1 - b_1^i\right)/N_s \end{aligned}, \tag{16.6}$$

where $\delta(\cdot)$ is the Kronecker Delta and it is equal to 1 if two integers in the Kronecker Delta are equal to each other; and 0 otherwise.

Step 2. Generate N_s random particles $\left\{a_{k-1}^i\right\}_{i=1}^{N_s}$ and $\left\{b_{k-1}^i\right\}_{i=1}^{N_s}$ from $\mathrm{p}\left(a_{k-1} \left| \mathrm{NC}_{1:k-1} \right.\right) = \mathrm{N}\left(a_{k-1}, \sigma_2^2\right)$, and $\mathrm{p}\left(b_{k-1} \left| \mathrm{NC}_{1:k-1} \right.\right) = \mathrm{N}\left(b_{k-1}, \sigma_3^2\right)$, respectively. According to the theory of particle filtering [25], once a new measurement NC_k, namely the k^{th} measured normalized capacity, is available, the weight updating is calculated by:

$$\omega_k^i = \frac{\omega_{k-1}^i \dfrac{\mathrm{p}\left(\mathrm{NC}_k \left| a_k^i, b_k^i\right.\right) \mathrm{p}\left(a_k^i \left| a_{k-1}^i\right.\right) \mathrm{p}\left(b_k^i \left| b_{k-1}^i\right.\right)}{\mathrm{q}\left(a_k^i, b_k^i \left| \mathrm{NC}_k, a_{k-1}^i, b_{k-1}^i\right.\right)}}{\sum_{i=1}^{N_s} \omega_{k-1}^i \dfrac{\mathrm{p}\left(\mathrm{NC}_k \left| a_k^i, b_k^i\right.\right) \mathrm{p}\left(a_k^i \left| a_{k-1}^i\right.\right) \mathrm{p}\left(b_k^i \left| b_{k-1}^i\right.\right)}{\mathrm{q}\left(a_k^i, b_k^i \left| \mathrm{NC}_k, a_{k-1}^i, b_{k-1}^i\right.\right)}}. \tag{16.7}$$

where $\mathrm{p}\left(a_k^i \left| a_{k-1}^i\right.\right)$ and $\mathrm{p}\left(b_k^i \left| b_{k-1}^i\right.\right)$ are the state transition probability density functions established by Eqs. (16.2) and (16.3); $\mathrm{q}\left(a_k^i, b_k^i \left| \mathrm{NC}_k, a_{k-1}^i, b_{k-1}^i\right.\right)$ is the importance function; and the probability density function of the measurement $\mathrm{p}\left(\mathrm{NC}_k \left| a_k^i, b_k^i\right.\right)$. In a number of applications related to battery prognostics [16, 18, 29], the importance function is often selected as a prior distribution so as to simplify Eq. (16.7):

$$\omega_k^i = \frac{\omega_{k-1}^i \mathrm{p}\left(\mathrm{NC}_k \left| a_k^i, b_k^i\right.\right)}{\sum_{i=1}^{N_s} \omega_{k-1}^i \mathrm{p}\left(\mathrm{NC}_k \left| a_k^i, b_k^i\right.\right)}. \tag{16.8}$$

It should be noted that the importance function used in Eq. (16.8) does not consider the recent battery degradation data NC_k to update the weights of random particles. According to the size of ω_k^i, the same systematic resampling [28] previously introduced in this section is used to re-draw random particles $\left\{a_k^i\right\}_{i=1}^{N_s}$ and $\left\{b_k^i\right\}_{i=1}^{N_s}$. After the systematic resampling is conducted, each of the weights is set to $1/N_s$. The posterior probability density functions of the amplitude a and slope b are respectively expressed as follows:

$$\begin{aligned} \mathrm{p}\left(a_k \left| \mathrm{NC}_{1:k}\right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(a_k - a_k^i\right)/N_s \\ \mathrm{p}\left(b_k \left| \mathrm{NC}_{1:k}\right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(b_k - b_k^i\right)/N_s \end{aligned}. \tag{16.9}$$

Step 3. Let $k = k + 1$ and return to Step 2 until $k > M$is satisfied. Here, M is the length of the available normalized capacity degradation data for a specific lithium-ion battery. At iteration M, the posterior probability density functions of the two parameters are given as follows:

$$\begin{aligned} \mathrm{p}\left(a_M \left| \mathrm{NC}_{1:M}\right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(a_M - a_M^i\right)/N_s \\ \mathrm{p}\left(b_M \left| \mathrm{NC}_{1:M}\right.\right) &\approx \sum_{i=1}^{N_s} \delta\left(b_M - b_M^i\right)/N_s \end{aligned}. \tag{16.10}$$

Extrapolations of the nonlinear function $f(\cdot)$ with the parameters distributions provided by (16.10) to a specified failure threshold, such as 0.85 used in Sects. 16.2 and 16.3, are used to calculate the probability density function $\mathrm{RUL}(c_M)$ of the RUL of lithium-ion batteries at cycle M:

$$\mathrm{RUL}\left(c_M\right) = \sum_{i=1}^{N_s} \frac{\delta\left(c_M - \inf\left(k \in \mathrm{int} : a_M^i \times \mathrm{e}^{b_M^i \times k} \leq 0.85\right) + M\right)}{N_s}, \tag{16.11}$$

where the int is the set of all integers and the inf takes the greatest lower bound of a set. The prediction of RUL and its lower and upper bounds are 50th, 5th, and 95th percentiles of Eq. (16.11), respectively. Moreover, the failure probability density function (FPDF) is calculated by removing the constant term M from Eq. (16.11).

16.3 A Spherical Cubature Particle Filtering Based State Space Model for Battery Remaining Useful Life Prediction at a Constant Discharge Rate [30]

It should be noted that the importance functions used in Eqs. (16.5) and (16.8) only depend on the current system states and does not incorporate the recent battery degradation data, such as the recent battery degradation data. When the recent battery degradation data has tangible influences on the importance function, Eqs. (16.5) and (16.8) for updating the weights become inaccurate and unreliable. Therefore, the recent battery degradation data should be explicitly incorporated in the design of the importance function. To achieve this goal, spherical cubature Kalman filtering is used in particle filtering to design a proper importance function which incorporates both the previous system states and the recent battery degradation data. When a new NC is available, the posterior probability density functions of the amplitude a and slope b are iteratively updated by the following steps.

Step 1. Generate N_s initial random particles $\{a_0^i\}_{i=1}^{N_s}$ and $\{b_0^i\}_{i=1}^{N_s}$ from the prior distributions $\mathrm{p}(a_0) = \mathrm{N}(a_0, \sigma_2^2)$, and $\mathrm{p}(b_0) = \mathrm{N}(b_0, \sigma_3^2)$, respectively. Each of $\{\omega_0^i\}_{i=1}^{N_s}$ is set to $1/N_s$ and their sum is equal to 1. For each random particle, an initial distribution for the use of the spherical cubature Kalman filtering is $\mathrm{p}(\boldsymbol{x}_0^i) = \mathrm{p}(a_0^i, b_0^i) \approx \mathrm{N}(a_0^i, b_0^i | \boldsymbol{m}_0, \boldsymbol{P}_0)$. The initial mean vector $\boldsymbol{m}_0$ and the initial covariance matrix $\boldsymbol{P}_0$ can be constructed by using the means and standard deviations stated in Sect. 16.2. Recalling spherical cubature integration, only four sigma points $\boldsymbol{\chi}_0^j = \boldsymbol{m}_0 + \sqrt{2(\boldsymbol{P}_0 + \boldsymbol{Q})}\boldsymbol{u}^j, j = 1, \ldots, 4$, are propagated to the nonlinear function $f(\cdot)$:

$$Y_0^j = f\left(\boldsymbol{\chi}_0^j\right), \quad j = 1, \ldots, 4, \tag{16.12}$$

where $\boldsymbol{Q}$ is a noise covariance established by standard deviations of additive Gaussian state noises.

The predicted mean u_{U}^0, the predicted covariance S_{U}^0, and the cross-covariance $\boldsymbol{C}_{\mathrm{U}}^0$ of the predicted state and the predicted measurement are calculated as follows:

$$\begin{aligned} \mathrm{E}[f(\boldsymbol{\chi}_0)] &= u_{\mathrm{U}}^0 = \textstyle\sum_{j=1}^4 Y_0^j/4 \\ \mathrm{Cov}[f(\boldsymbol{\chi}_0) + v_1] &= S_{\mathrm{U}}^0 = \textstyle\sum_{j=1}^4 \left(Y_0^j - u_{\mathrm{U}}^0\right)\left(Y_0^j - u_{\mathrm{U}}^0\right)^{\mathrm{T}}/4 + \sigma_1^2. \\ \mathrm{Cov}[\boldsymbol{\chi}_0, f(\boldsymbol{\chi}_0)] &= \boldsymbol{C}_{\mathrm{U}}^0 = \textstyle\sum_{j=1}^4 \left(\boldsymbol{\chi}_0^j - \boldsymbol{m}_0\right)\left(Y_0^j - u_{\mathrm{U}}^0\right)^{\mathrm{T}}/4 \end{aligned} \tag{16.13}$$

The $\mathrm{q}\left(a_1^i, b_1^i \,|\mathrm{NC}_1\,, a_0^i, b_0^i\right)$ is approximated by the multi-dimensional Gaussian distribution with the updated mean vector $\boldsymbol{m}_1$ and the updated covariance matrix $\boldsymbol{P}_1$:

$$\begin{aligned} \boldsymbol{m}_1 &= \boldsymbol{m}_0 + \boldsymbol{C}_{\mathrm{U}}^0\left(S_{\mathrm{U}}^0\right)^{-1}\left(\mathrm{NC}_1 - u_{\mathrm{U}}^0\right) \\ \boldsymbol{P}_1 &= \boldsymbol{P}_0 - \boldsymbol{C}_{\mathrm{U}}^0\left(S_{\mathrm{U}}^0\right)^{-1} S_{\mathrm{U}}^0\left(\boldsymbol{C}_{\mathrm{U}}^0\left(S_{\mathrm{U}}^0\right)^{-1}\right)^{\mathrm{T}}, \end{aligned} \tag{16.14}$$

where $\boldsymbol{C}_{\mathrm{U}}^0\left(S_{\mathrm{U}}^0\right)^{-1}$ is the filtering gain used in the Kalman filtering. Equations (16.12), (16.13), and (16.14) are called the spherical cubature Kalman filtering. The equation for the calculation of the weight used in the spherical cubature Kalman filtering is derived as

$$\omega_1^i = \frac{\omega_0^i \dfrac{p\left(\mathrm{NC}_1 \left|a_1^i, b_1^i\right.\right) p\left(a_1^i \left|a_0^i\right.\right) p\left(b_1^i \left|b_0^i\right.\right)}{q\left(a_1^i, b_1^i \,|\mathrm{NC}_1\,, a_0^i, b_0^i\right)}}{\sum_{i=1}^{N_s} \omega_0^i \dfrac{p\left(\mathrm{NC}_1 \left|a_1^i, b_1^i\right.\right) p\left(a_1^i \left|a_0^i\right.\right) p\left(b_1^i \left|b_0^i\right.\right)}{q\left(a_1^i, b_1^i \left|\mathrm{NC}_1, a_0^i, b_0^i\right.\right)}}. \tag{16.15}$$

According to the size of ω_1^i, the same systematic resampling [28] introduced in the previous section is used to re-draw random particles. Then, each of the associated weights equals to $1/N_s$. Resulting posterior probability density functions of the two parameters are derived as

$$\begin{aligned} \mathrm{p}\left(a_1 \,|\mathrm{NC}_1\right) &\approx \textstyle\sum_{i=1}^{N_s} \delta\left(a_1 - a_1^i\right)/N_s \\ \mathrm{p}\left(b_1 \,|\mathrm{NC}_1\right) &\approx \textstyle\sum_{i=1}^{N_s} \delta\left(b_1 - b_1^i\right)/N_s \end{aligned}. \tag{16.16}$$

Step 2. Generate N_s random particles $\left\{a_{k-1}^i\right\}_{i=1}^{N_s}$ and $\left\{b_{k-1}^i\right\}_{i=1}^{N_s}$ from Eq. (16.16). Suppose that the posterior probability density function $\mathrm{q}\left(\boldsymbol{x}_{k-1}^i \,|\mathrm{NC}_{k-1}\,, \boldsymbol{x}_{k-2}^i\right) = \mathrm{q}\left(a_{k-1}^i, b_{k-1}^i \,|\mathrm{NC}_{k-1}\,, a_{k-2}^i, b_{k-2}^i\right) \approx \mathrm{N}\left(a_{k-1}^i, b_{k-1}^i \,|\boldsymbol{m}_{k-1}, \boldsymbol{P}_{k-1}\right)$at iteration k-1 is estimated by using the spherical cubature Kalman filtering. Only four sigma points $\boldsymbol{\chi}_{k-1}^j = \boldsymbol{m}_{k-1} + \sqrt{2\boldsymbol{P}_{k-1}}\,\boldsymbol{u}^j, j = 1\,,\ \ldots\ ,4$, are used to calculate the mean and the covariance of the nonlinear function $f(\cdot)$:

$$Y_{k-1}^j = f\left(\boldsymbol{\chi}_{k-1}^j\right),\ \ j = 1, \ldots, 4. \tag{16.17}$$

The predicted mean u_{U}^{k-1}, the predicted covariance S_{U}^{k-1}, and the cross-covariance $\boldsymbol{C}_{\mathrm{U}}^{k-1}$ of the predicted state and the predicted measurement are calculated as follows:

$$\begin{aligned}
\mathrm{E}\left[f\left(\boldsymbol{\chi}_{k-1}\right)\right] &= u_{\mathrm{U}}^{k-1} = \textstyle\sum_{j=1}^{4} Y_{k-1}^{j}/4 \\
\mathrm{Cov}\left[f\left(\boldsymbol{\chi}_{k-1}\right)+v_1\right] &= S_{\mathrm{U}}^{k-1} = \textstyle\sum_{j=1}^{4}\left(Y_{k-1}^{j}-u_{\mathrm{U}}^{k-1}\right)\left(Y_{k-1}^{j}-u_{\mathrm{U}}^{k-1}\right)^{\mathrm{T}}/4+\sigma_1^2 . \\
\mathrm{Cov}\left[\boldsymbol{\chi}_{k-1}, f\left(\boldsymbol{\chi}_{k-1}\right)\right] &= \boldsymbol{C}_{\mathrm{U}}^{k-1} = \textstyle\sum_{j=1}^{4}\left(\boldsymbol{\chi}_{k-1}^{j}-\boldsymbol{m}_{k-1}\right)\left(Y_{k-1}^{j}-u_{\mathrm{U}}^{k-1}\right)^{\mathrm{T}}/4
\end{aligned} \tag{16.18}$$

Once a new NC_k is available, the $\mathrm{q}\left(a_k^i, b_k^i \left|\mathrm{NC}_k\right., a_{k-1}^i, b_{k-1}^i\right)$ is approximated by a multi-dimensional Gaussian distribution with the mean vector $\boldsymbol{m}_k$ and the covariance matrix $\boldsymbol{P}_k$:

$$\begin{aligned}
\boldsymbol{m}_k &= \boldsymbol{m}_{k-1} + \boldsymbol{C}_{\mathrm{U}}^{k-1}\left(S_{\mathrm{U}}^{k-1}\right)^{-1}\left(\mathrm{NC}_k - u_{\mathrm{U}}^{k-1}\right) \\
\boldsymbol{P}_k &= \boldsymbol{P}_{k-1} - \boldsymbol{C}_{\mathrm{U}}^{k-1}\left(S_{\mathrm{U}}^{k-1}\right)^{-1} S_{\mathrm{U}}^{k-1}\left(\boldsymbol{C}_{\mathrm{U}}^{k-1}\left(S_{\mathrm{U}}^{k-1}\right)^{-1}\right)^{\mathrm{T}} .
\end{aligned} \tag{16.19}$$

The weight updating formula is derived as

$$\omega_k^i = \frac{\omega_{k\text{-}1}^i \dfrac{\mathrm{p}\left(\mathrm{NC}_k \left|a_k^i, b_k^i\right.\right)\mathrm{p}\left(a_k^i \left|a_{k-1}^i\right.\right)\mathrm{p}\left(b_k^i \left|b_{k-1}^i\right.\right)}{\mathrm{q}\left(a_k^i, b_k^i \left|\mathrm{NC}_k, a_{k-1}^i, b_{k-1}^i\right.\right)}}{\sum_{i=1}^{N_s}\omega_{k\text{-}1}^i \dfrac{\mathrm{p}\left(\mathrm{NC}_k \left|a_k^i, b_k^i\right.\right)\mathrm{p}\left(a_k^i \left|a_{k-1}^i\right.\right)\mathrm{p}\left(b_k^i \left|b_{k-1}^i\right.\right)}{\mathrm{q}\left(a_k^i, b_k^i \left|\mathrm{NC}_k\right., a_{k-1}^i, b_{k-1}^i\right.\right)}} . \tag{16.20}$$

According to the size of ω_k^i, the same systematic resampling is conducted to redraw random particles from the posterior probability density function. After that, all of their associated weights are equal with each other.

Step 3. Let $k = k + 1$ and return to Step 2 until $k > M$is satisfied. Here, M is the length of the available capacity degradation data for a specific lithium-ion battery. At iteration M, the posterior probability density functions of the two parameters are given as follows:

$$\begin{aligned}
\mathrm{p}\left(a_M \left|\mathrm{NC}_{1:M}\right.\right) &\approx \textstyle\sum_{i=1}^{N_s}\delta\left(a_M - a_M^i\right)/N_s \\
\mathrm{p}\left(b_M \left|\mathrm{NC}_{1:M}\right.\right) &\approx \textstyle\sum_{i=1}^{N_s}\delta\left(b_M - b_M^i\right)/N_s
\end{aligned} . \tag{16.21}$$

Extrapolations of the nonlinear function $f(\cdot)$ with the parameters distributions provided by Eq. (16.21) to a specified failure threshold, such as 0.85 used in this section, are used to calculate the probability density function of the RUL of lithium-ion batteries at cycle number M:

$$\mathrm{RUL}\left(c_M\right) = \sum_{i=1}^{N_s}\frac{\delta\left(c_M - \inf\left(k \in \mathrm{int} : a_M^i \times \mathrm{e}^{b_M^i \times k} \le 0.85\right) + M\right)}{N_s}, \tag{16.22}$$

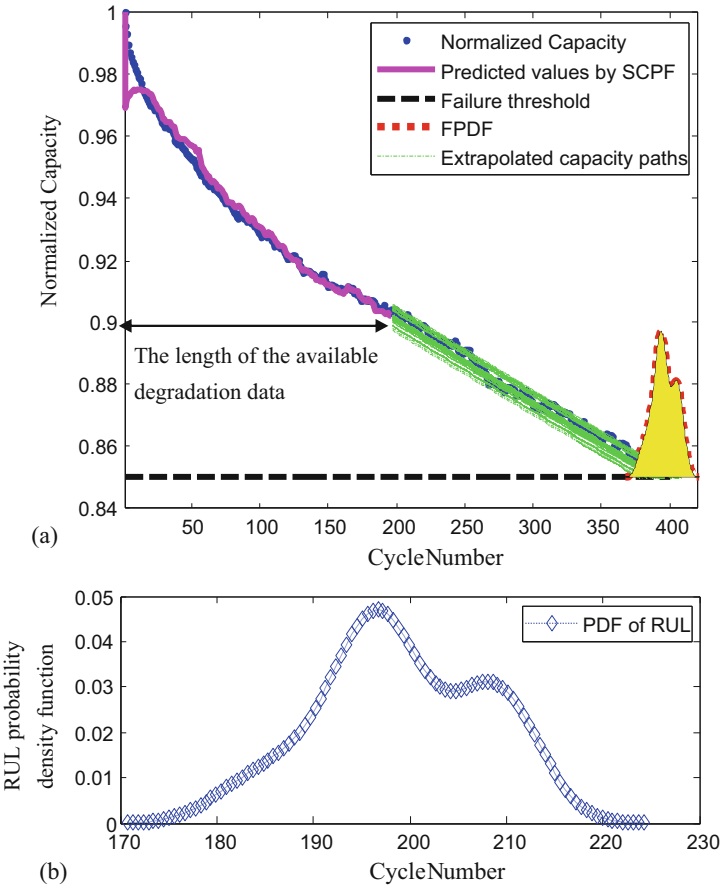

Fig. 16.3 RUL prediction of testing sample 5 by using the proposed spherical cubature particle filtering based prognostic method at the 50% of the AL. (**a**) Predicted capacity values and extrapolated capacity paths; (**b**) the probability density function of battery RUL

The prediction of RUL and its lower and upper bounds are 50th, 5th, and 95th percentiles of Eq. (16.22), respectively.

In this section, RUL of lithium-ion batteries is predicted at the 50%, 75% and 90% of the actual life (AL) of lithium-ion batteries. Here, the AL is defined as the cycle time to the capacity reaching the failure threshold of 85% of an initial capacity value.

Taking testing sample 5 for example, at the 50% of the actual life, the predictions of the RUL are plotted in Fig. 16.3. In Fig. 16.3a, we observe that the predictive values estimated by the spherical cubature particle filtering track the true (normalized) capacities well (extrapolations of the nonlinear function are represented by the green dash lines). The PDF of the RUL is shown in Fig. 16.3b, where 5th, 50th and 95th percentiles of the RUL are 183, 198 and 211 cycles, respectively. Note that the

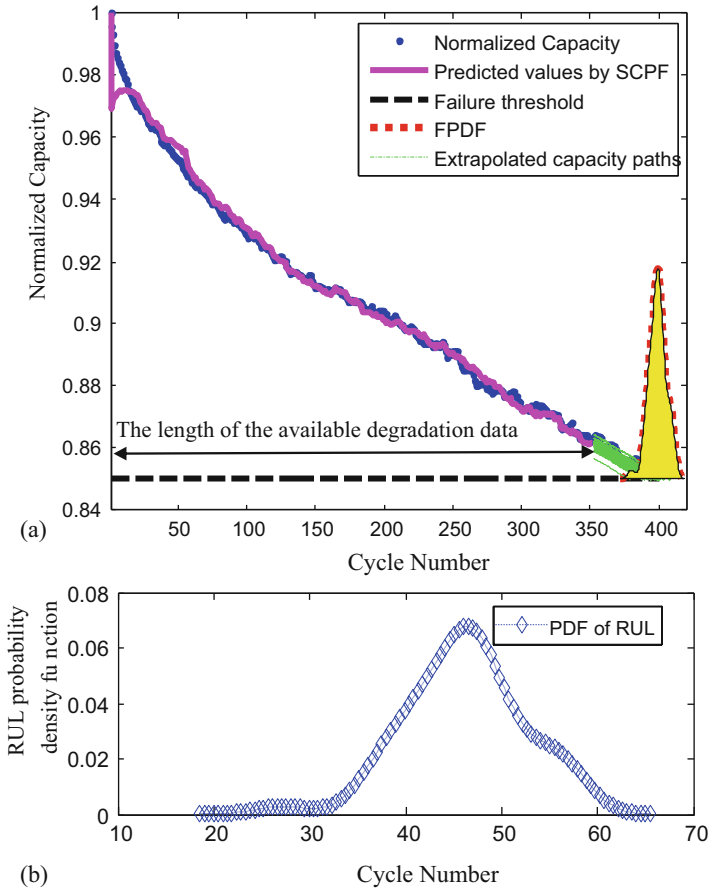

Fig. 16.4 RUL prediction of testing sample 5 by using the proposed spherical cubature particle filtering based prognostic method at the 90% of the AL. (**a**) Predicted capacity values and extrapolated capacity paths; (**b**) the probability density function of battery RUL

absolute error between the predicted RUL and true RUL is 3 cycles. The prognostic results obtained by using the spherical cubature particle filtering at the 90% of the AL are presented in Fig. 16.4, respectively. From the results shown in Fig. 16.4a, we observe that the predictive normalized capacities are well matched with the true normalized capacities. At the prediction cycle of the 90% of the actual life, the 5th, 50th and 95th percentiles of the RUL are 37, 45 and 55 cycles, respectively. The absolute error between the predicted RUL and true RUL is 6 cycles. In this instance, the prediction accuracies are high because the capacity degradation curve of sample 5 has a good accordance with the exponential function as used in the state space model.

The prediction results by using the particle filtering based prognostic method and the spherical cubature particle filtering based prognostic method for all the testing samples 1 to 6 are respectively tabulated in Table 16.2. The differences of

Table 16.2 RUL predictions by using the particle filtering based prognostic method and the spherical cubature particle filtering based prognostic method (Unit: Cycle)

Testing sample	Prediction time (50%, 75% and 90% of AL)	True RUL	The percentile of RUL and absolute error obtained by using the spherical cubature Particle Filtering Prognostic Method				The percentile of RUL and absolute error obtained by using the standard PF based prognostic method				The difference of the absolute error of the two methods
			The percentile of predicted RUL			Absolute error	The percentile of predicted RUL			Absolute error	
			5%	50%	95%		5%	50%	95%		
1	50%	177	238	243	253	66	235	254	283	77	−11
	75%	88	120	128	134	40	120	138	155	50	−10
	90%	35	27	49	63	14	46	62	76	27	−13
2	50%	173	252	257	269	84	251	270	298	97	−13
	75%	86	131	147	155	61	142	157	179	71	−10
	90%	35	46	56	66	21	59	74	87	39	−18
3	50%	160	205	212	229	52	212	226	244	66	−14
	75%	80	109	123	135	43	126	138	153	58	−15
	90%	32	40	53	61	21	51	67	83	35	−14
4	50%	187	163	183	189	4	170	191	215	4	0
	75%	94	81	97	103	3	90	105	125	11	−8
	90%	37	32	35	52	2	28	43	63	6	−4
5	50%	195	183	198	211	3	184	203	228	8	−5
	75%	98	82	96	113	2	94	107	116	9	−7
	90%	39	37	45	55	6	35	51	72	12	−6
6	50%	166	149	165	168	1	163	176	194	10	−9
	75%	83	78	89	98	6	79	94	108	11	−5
	90%	33	23	38	52	5	36	46	63	13	−8

the absolute errors of the two prognostic methods in the last column of Table 16.2 are almost less than 0, which means that, in all the instances, the spherical cubature particle filtering based prognostic method improves the prediction accuracies of the particle filtering based prognostic method. On the other hand, it is found that the prediction accuracies for the testing samples 1 to 3 are lower than those for the testing samples 4 to 6. This is caused by the relatively large fluctuation of the testing samples 1 to 3 in Fig. 16.5. Specifically, due to some sudden changes inside the cells, such as surface temperature increasing, internal resistance dropping, together with battery-to-battery variations, the capacity degradation samples 1 to 3 in Fig. 16.5a are fluctuated. Compared with those testing samples in Fig. 16.5a, testing samples 4 to 6 in Fig. 16.5b have moderate fluctuations. Therefore, if predictions are made at these unexpected fluctuated capacity degradation points, the prediction results of testing samples 1 to 3 are not as good as those of testing samples 4 to 6 because battery capacity degradation does not well follow the exponential function used in the state space model.

16.4 A Particle Filtering Based State Space Model for Battery Remaining Useful Life Prediction at Different Discharge Rates [31]

In Sects. 16.2 and 16.3, two prognostic methods have been proposed to predict battery RUL at a constant discharge rate. In this section, we will introduce another particle filtering based prognostic method working at different discharge rates, which is more attractive in practice because many electronic products work at different discharge rates. Before the proposed prognostic method will be introduced, we will illustrate our battery experiment designs and how to collect battery degradation data at different discharge rates. In our experiments, four cylindrical BK 18650 batteries rated with 1 Ah were tested using an Arbin BT2000 tester. The battery test bench shown in Fig. 16.6a composed of the Arbin BT2000 tester, a host computer with an Arbin MITS Pro Software, and a computer with the Matlab R2012b Software. An operating profile for each battery was made of a sequence of repetitive C/2, 1C, 3C, and 5C constant current discharge regimes to a 2 V cutoff voltage. Here, C is a measure of the rate at which a battery would be discharged. 1 C means that a battery's rated capacity would be delivered in 1 h. Each of the four batteries was recharged with a schedule recommended by manufacturer, which comprised a 1C constant current charging step to 3.6 V followed by a constant voltage step until a cutoff current of C/20 was reached. Figure 16.6b shows the measured current and voltage profile at a four-cycle rotation interval. In the Arbin testing system, the discharge current was denoted as a negative value, vice versa. For the rest of this book chapter, cycle is used instead of rotation for battery RUL prediction. The relationship between the discharge curve and the accumulated capacity is plotted in Fig. 16.6c, where the maximum releasable capacities at

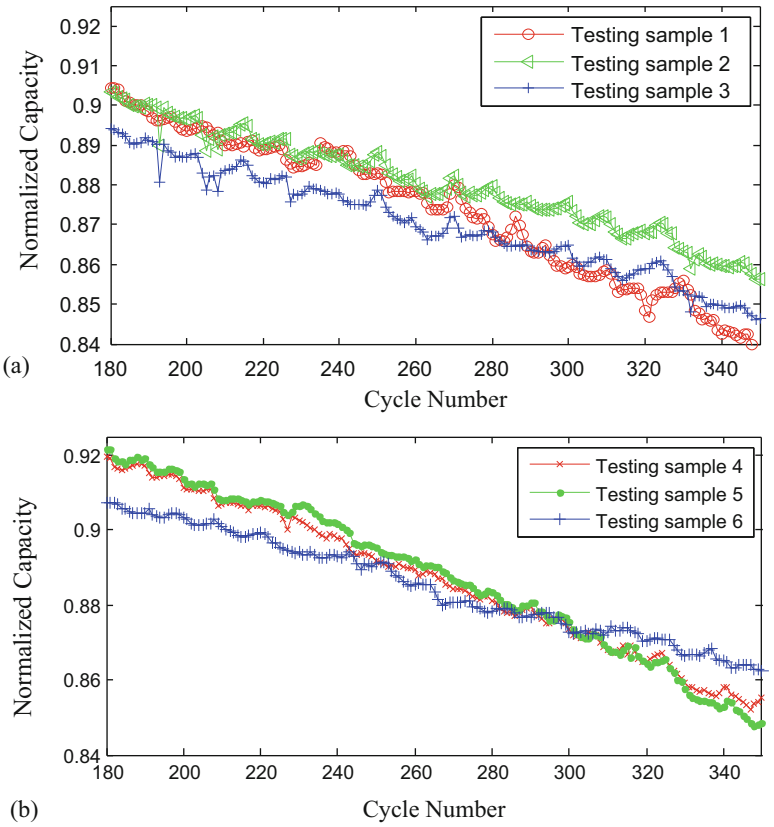

Fig. 16.5 Capacity degradation curves of samples 1–6 from cycles 180 to 350. (**a**) Testing samples 1–3; (**b**) Testing samples 4–6

different discharge rates are different. The four testing batteries were named as batteries 1 to 4, respectively and their associated degradation profiles are plotted in Figs. 16.7a–d, respectively, where the increasing capacities at some initial cycles are removed because they are not useful to describe battery degradation. Besides, it should be noted that the battery capacity degradation data at discharge rate 5 C are more fluctuated than those at discharge rates 0.5 C, 1 C and 3 C, which is caused by a high battery temperature in the case of a high discharge rate.

In terms of Fig. 16.7, it is not difficult to find that battery degradation data are totally dependent on different discharge rates. Therefore, it is necessary to propose a discharge rates-dependent state space model based prognostic method before battery RUL prediction at the different discharge rates is conducted. Firstly, we need to use historical degradation data to find the relationship between the discharge rates-dependent state space model and different discharge rates. In this section, lithium-ion batteries 1 to 3 are used as historical degradation data and lithium-ion battery 4 is used as testing degradation data. Moreover, a soft failure threshold is

defined as the 80% of the maximum capacity of a battery. Consequently, only the capacity data from the very early beginning to the cycle reaching the soft failure threshold are used for our analyses. Being similar with Sects. 16.2 and 16.3, we found that a simple exponential function with the amplitude a and the slope b is able to fit the battery capacity degradation at different discharge rates. To validate such conclusion, goodness of fit is also used here. The boxplots of the estimated parameters of the exponential function for fitting battery degradation data 1 to 3 at different discharge rates and their associated R-squared values are plotted in Figs. 16.8a–c, respectively, where most of R-squared values approach to 1 and they connote that all capacity degradation data at the different discharge rates are well fitted by the exponential function.

To establish a unique degradation equation for battery degradation at the different discharge rates, a discharge rates-dependent exponential function is proposed. The main idea of the discharge rates-dependent exponential function is to correlate the amplitude and slope of the exponential function with the different discharge rates. The estimated amplitudes and slopes at the different discharge rates for the three batteries are plotted in Figs. 16.9a–f, in which we observe that the amplitude and the slope have a linear relationship with the discharge rate, respectively. As a result, a simple linear function is used to describe the relationship between the amplitude/slope and the different discharge rates. The specific formulas for the linear relationships for the three different batteries are given in Figs. 16.9a–f, respectively.

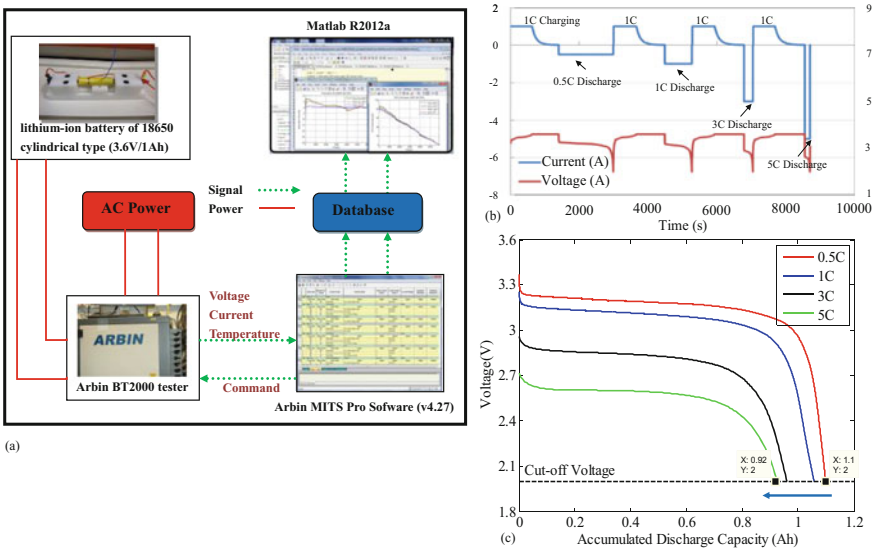

Fig. 16.6 Experiment designs and data collection at four-cycle rotation intervals: (**a**) the battery test bench; (**b**) the measured current and voltage at a four-cycle rotation interval; (**c**) the maximum releasable capacity at different discharge rates

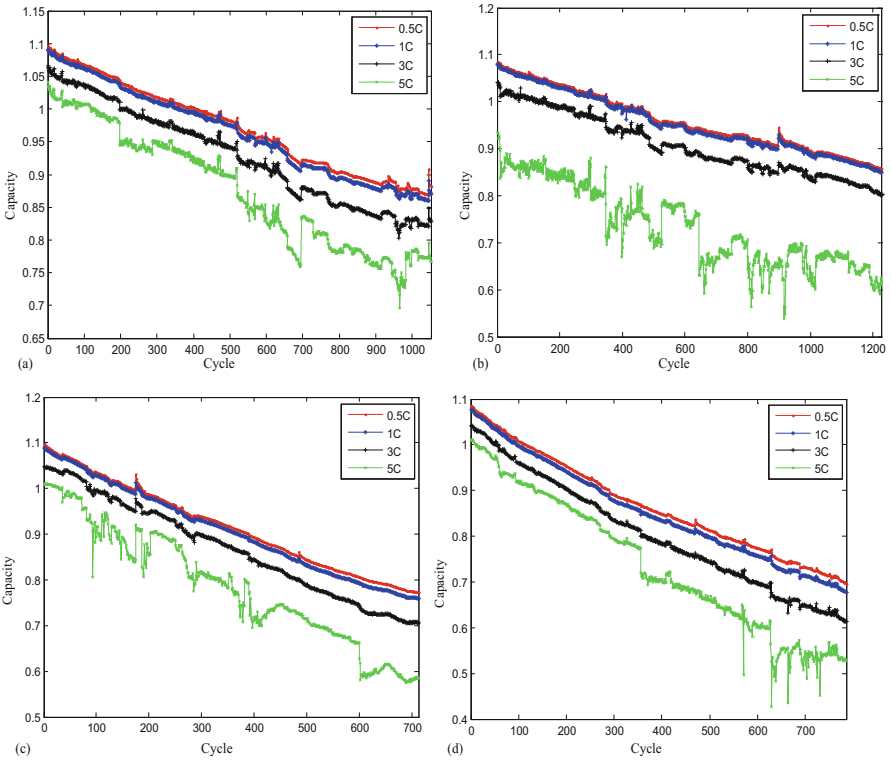

Fig. 16.7 Lithium-ion battery degradation data at different discharge rates 0.5C, 1 C, 3C and 5C: (**a**) battery 1; (**b**) battery 2; (**c**) battery 3; (**d**) battery 4

From the aforementioned analyses, the following discharge rates-dependent state space model is proposed:

$$\begin{aligned} a1_i &= a1_{i-1} + v_1 \\ a2_i &= a2_{i-1} + v_2 \\ b1_i &= b1_{i-1} + v_3 \\ b2_i &= b2_{i-1} + v_4 \\ x_i &= a(D) \times \mathrm{e}^{b(D)\times i} + v_5 = (a1_i + a2_i \times D) \times \mathrm{e}^{(b1_i + b2_i \times D)\times i} + v_5, \end{aligned} \tag{16.23}$$

where $a(D) = (a1_i + a2_i \times D)$ is used to express the relationship between the amplitude a and the discharge rate D; $b(D) = (b1_i + b2_i \times D)$ is used to express the relationship between the amplitude b and the discharge rate D; each of thev_1, v_2, v_3, v_4 and v_5 follows a Gaussian distribution with a zero mean and its associated deviation σ_1, σ_2, σ_3, σ_4 and σ_5, respectively; x_i is the current capacity degradation data at cycle i. Here, x_i is not normalized and it is different from the normalized capacity used in the previous two sections because the influence of different

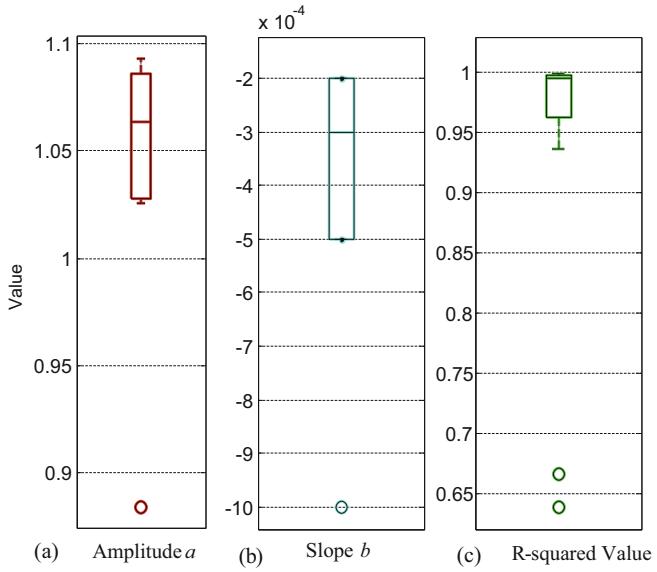

Fig. 16.8 Boxplots of the amplitudes and slopes of an exponential function and its associated R-squared values at the different discharge rates in the case of lithium-ion batteries 1, 2 and 3: (**a**) the amplitude a; (**b**) the slope b; (**c**) the R-squared value

discharge rates on battery capacity degradation is considered in this section. The initialization of Eq. (16.23) can be established by taking the mean of the estimated amplitudes and slopes as shown in Fig. 16.9. Specifically, $a1_0$, $a2_0$, $b1_0$ and $b2_0$ are respectively equal to 1.1026, -0.0231, -2.6503×10^{-4} and -4.19×10^{-5}. According the scale of each of the parameters used in Eq. (16.23), the σ_1, σ_2, σ_3, σ_4 and σ_5 are empirically set to 10^{-4}, 10^{-4}, 10^{-5}, 10^{-5}, and 10^{-2}, respectively.

For a specific battery, given its historical measurements up to date $X(0:M) = \{x_0, x_1, \ldots, x_i, \ldots, x_M\}$ and a specific discharge rate $D = \mathrm{d}$, such as $\mathrm{d} = 0.5$ C or 1 C or 3 C or 5 C, the parameters $\boldsymbol{\phi} = [a1_i, a2_i, b1_i, b2_i]$ and their distributions should be posteriorly estimated by solving Eq. (16.23). Here, M is the length of the available battery degradation data. The specific details how the proposed prognostic method works at different discharge rates are illustrated in the following paragraphs.

Step 1. $N_s = 2000$ random particles $\left\{a1_0^k\right\}_{k=1}^{N_s}$, $\left\{a2_0^k\right\}_{k=1}^{N_s}$, $\left\{b1_0^k\right\}_{k=1}^{N_s}$ and $\left\{b2_0^k\right\}_{k=1}^{N_s}$ are initially drawn from the Gaussian distributions $\mathrm{p}\left(a1\,|x_0\right) = \mathrm{N}\left(a1_0, \sigma_1^2\right)$, $\mathrm{p}\left(a2\,|x_0\right) = \mathrm{N}\left(a2_0, \sigma_2^2\right)$, $\mathrm{p}\left(b1\,|x_0\right) = \mathrm{N}\left(b1_0, \sigma_3^2\right)$ and $\mathrm{p}\left(b2\,|x_0\right) = \mathrm{N}\left(b2_0, \sigma_4^2\right)$, respectively. Compared with the number of random particles used in Sects. 16.2 and 16.3, the number of random particles used in this section increases because more

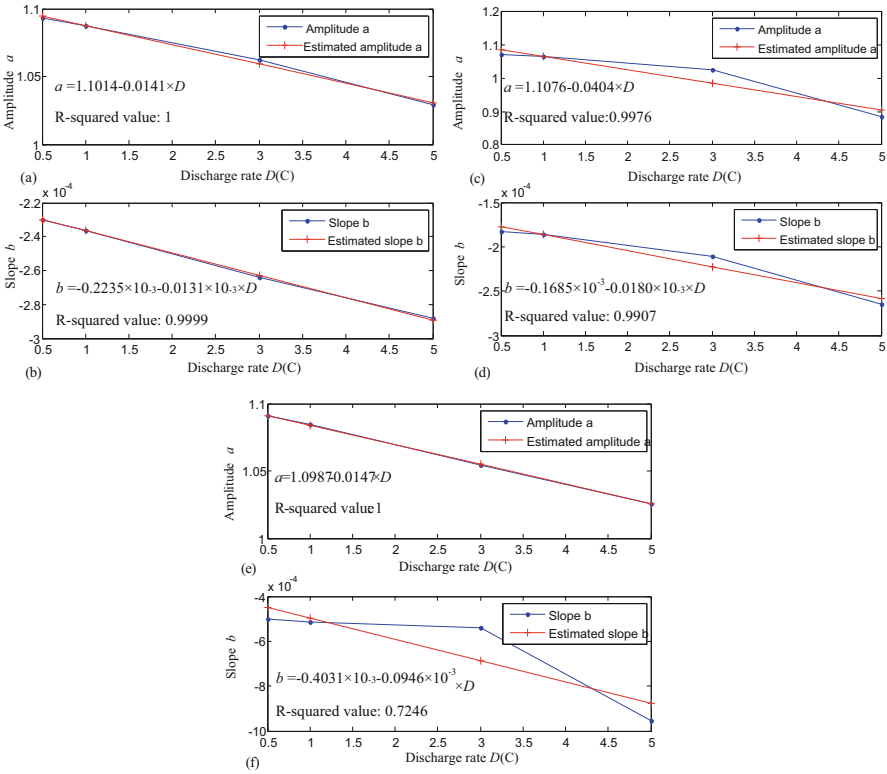

Fig. 16.9 The amplitudes and slopes of an exponential function at the different discharge rates in the case of lithium-ion batteries 1, 2 and 3: (**a**) the relationship between the amplitude a and the different discharge rates in the case of battery 1; (**b**) the relationship between the slope b and the different discharge rates in the case of battery 1; (**c**) the relationship between the amplitude a and the different discharge rates in the case of battery 2; (**d**) the relationship between the slope b and the different discharge rates in the case of battery 2; (**e**) the relationship between the amplitude a and the different discharge rates in the case of battery 3; (**f**) the relationship between the slope b and the different discharge rates in the case of battery 3. (Note: D is the symbol used for representing a discharge rate)

random particles are required to track battery degradation at different discharge rates. Additionally, all random particles have the same initial weights $\{\omega_0^k\}_{k=1}^{N_s}$ of $1/N_s$. According to the fundamental theory of particle filtering [25] and the specification of Eq. (16.23), the formula for updating each weight at iteration $i = 1$ is provided as follows:

$$\omega_1^k = \omega_0^k \times \frac{1}{\sigma_5 \times 2\pi} \mathrm{e}^{\frac{-\left(x_1 - \left(a1_0^k + a2_0^k \times D\right) \times \mathrm{e}^{\left(b1_0^k + b2_0^k \times D\right) \times i}\right)^2}{2\sigma_5^2}} \Bigg/ \left(\sum_{k=1}^{N_s} \omega_0^k \times \frac{1}{\sigma_5 \times 2\pi} \mathrm{e}^{\frac{-\left(x_1 - \left(a1_0^k + a2_0^k \times D\right) \times \mathrm{e}^{\left(b1_0^k + b2_0^k \times D\right) \times i}\right)^2}{2\sigma_5^2}} \right), k = 1, 2, \ldots, N_s. \tag{16.24}$$

From Eq. (16.24), it is not difficult to find that only random particles which produce better estimated capacity values for x_1 have higher weights. After several iterations, most of the weights will become negligible and the variances of the weights increase. To alleviate such problem, the same systematic resampling algorithm [32] introduced in the previous two sections is used here when the following condition is satisfied:

$$\left(\sum_{k=1}^{N_s} \left(\omega_i^k\right)^2 \right)^{-1} < 0.9 \times N_s. \tag{16.25}$$

Step 2. Suppose that the posterior distributions of $a1_{i-1}$, $a2_{i-1}$, $b1_{i-1}$ and $b2_{i-1}$ at iteration $i-1$ are available. $N_s = 2000$ random particles $\left\{a1_{i-1}^k\right\}_{k=1}^{N_s}$, $\left\{a2_{i-1}^k\right\}_{k=1}^{N_s}$, $\left\{b1_{i-1}^k\right\}_{k=1}^{N_s}$ and $\left\{b2_{i-1}^k\right\}_{k=1}^{N_s}$ are drawn from the Gaussian distributions $\mathrm{p}\left(a1 \left| x_{i-1} \right.\right) = \mathrm{N}\left(a1_{i-1}, \sigma_1^2\right)$, $\mathrm{p}\left(a2 \left| x_{i-1} \right.\right) = \mathrm{N}\left(a2_{i-1}, \sigma_2^2\right)$, $\mathrm{p}\left(b1 \left| x_{i-1} \right.\right) = \mathrm{N}\left(b1_{i-1}, \sigma_3^2\right)$ and $\mathrm{p}\left(b2 \left| x_{i-1} \right.\right) = \mathrm{N}\left(b2_{i-1}, \sigma_4^2\right)$, respectively. Moreover, their associated weights are $\left\{\omega_{i-1}^k\right\}_{k=1}^{N_s}$. Given a new capacity measurement x_i, each of the weight of the random particles is updated by:

$$\omega_i^k = \omega_{i-1}^k \times \frac{1}{\sigma_5 \times 2\pi} \mathrm{e}^{\frac{-\left(x_i - \left(a1_{i-1}^k + a2_{i-1}^k \times D\right) \times \mathrm{e}^{\left(b1_{i-1}^k + b2_{i-1}^k \times D\right) \times i}\right)^2}{2\sigma_5^2}} \Bigg/ \left(\sum_{k=1}^{N_s} \omega_{i-1}^k \times \frac{1}{\sigma_5 \times 2\pi} \mathrm{e}^{\frac{-\left(x_i - \left(a1_{i-1}^k + a2_{i-1}^k \times D\right) \times \mathrm{e}^{\left(b1_{i-1}^k + b2_{i-1}^k \times D\right) \times i}\right)^2}{2\sigma_5^2}} \right), k = 1, 2, \ldots, N_s. \tag{16.26}$$

Conduct the same systematic resampling algorithm with those used in the previous sections, if Eq. (16.25) is satisfied.

Step 3. Increase $i = i + 1$ and repeat Step 2 until $i > M$. The posterior distributions of $a1_M$, $a2_M$, $b1_M$ and $b2_M$ at iteration M are represented as follows:

$$\begin{aligned}
\mathrm{p}\left(a1\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta\left(a1 - a1_M^k\right) \\
\mathrm{p}\left(a2\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta\left(a2 - a2_M^k\right) \\
\mathrm{p}\left(b1\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta\left(b1 - b1_M^k\right) \\
\mathrm{p}\left(b2\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta\left(b2 - b2_M^k\right).
\end{aligned} \tag{16.27}$$

Given a failure threshold $x_{\text{threshold}}$, such as the 80% of the maximum capacity of a battery, the historical measurements up to date $X(0:M) = \{x_0\,, x_1\,, \ \ldots\ , x_i, \ \ldots\ , x_M\}$, and the specific DCR $D = \mathrm{d}$, the probability density function of the battery RUL is predicted by using the following equation:

$$\begin{aligned}
\mathrm{p}\left(\kappa\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta \\
&\left(\kappa - \inf\left(j \in \mathrm{int} : \left(a1_M^k + a2_M^k \times \mathrm{d}\right) \times \mathrm{e}^{\left(b1_M^k + b2_M^k \times \mathrm{d}\right) \times j} \le x_{\text{threshold}}\right) + M\right), \\
&\qquad k = 1, 2, \ldots, N_s.
\end{aligned} \tag{16.28}$$

Consequently, the prediction of RUL and its lower and upper bounds are 50th, 5th, and 95th percentiles of Eq. (16.28), respectively.

Additionally, the failure probability density function of the battery is defined as follows:

$$\begin{aligned}
\mathrm{p}\left(\psi\,|\{x_0\,, x_1, \ldots, x_M\}\,, D = \mathrm{d}\right) &= \textstyle\sum_{k=1}^{N_s} \omega_M^k \delta \\
&\left(\psi - \inf\left(j \in \mathrm{int} : \left(a1_M^k + a2_M^k \times \mathrm{d}\right) \times \mathrm{e}^{\left(b1_M^k + b2_M^k \times \mathrm{d}\right) \times j} \le x_{\text{threshold}}\right)\right), \\
&\qquad k = 1, 2, \ldots, N_s.
\end{aligned} \tag{16.29}$$

At last, capacity degradation data oflithium-ion battery 4 at the different discharge rates are used to validate the effectiveness of the proposed discharge rates-dependent state space model based prognostic method. Firstly, it is necessary to define the actual life (AL) of battery 4 as the lifetime from an initial cycle to the cycle associated with the capacity reaching the 80% of its maximum capacity, namely the soft failure threshold. Secondly, we make RUL predictions at the 50%, 75% and 90% of the AL of battery 4 at the different discharge rates. Once the capacity degradation data and a specific discharge rate are available, the probability density function of the battery RUL can be calculated by using the proposed discharge rates-dependent state space model based prognostic method. To visualize the prognostic results, the raw capacity degradation data, predicted capacity, FPDF, predicted capacity degradation paths and PDFs of RUL at the different discharge rates and prediction cycles are plotted in Fig. 16.10. It is observed that the predicted capacity degradation paths well cover the raw capacity degradation data (blue dots). Specifically, 50th, 5th, and 95th percentiles of RUL for predictions at the different discharge rates and prediction cycles are tabulated in Table 16.3, where the prediction accuracies at discharge rate 5 C are not as good

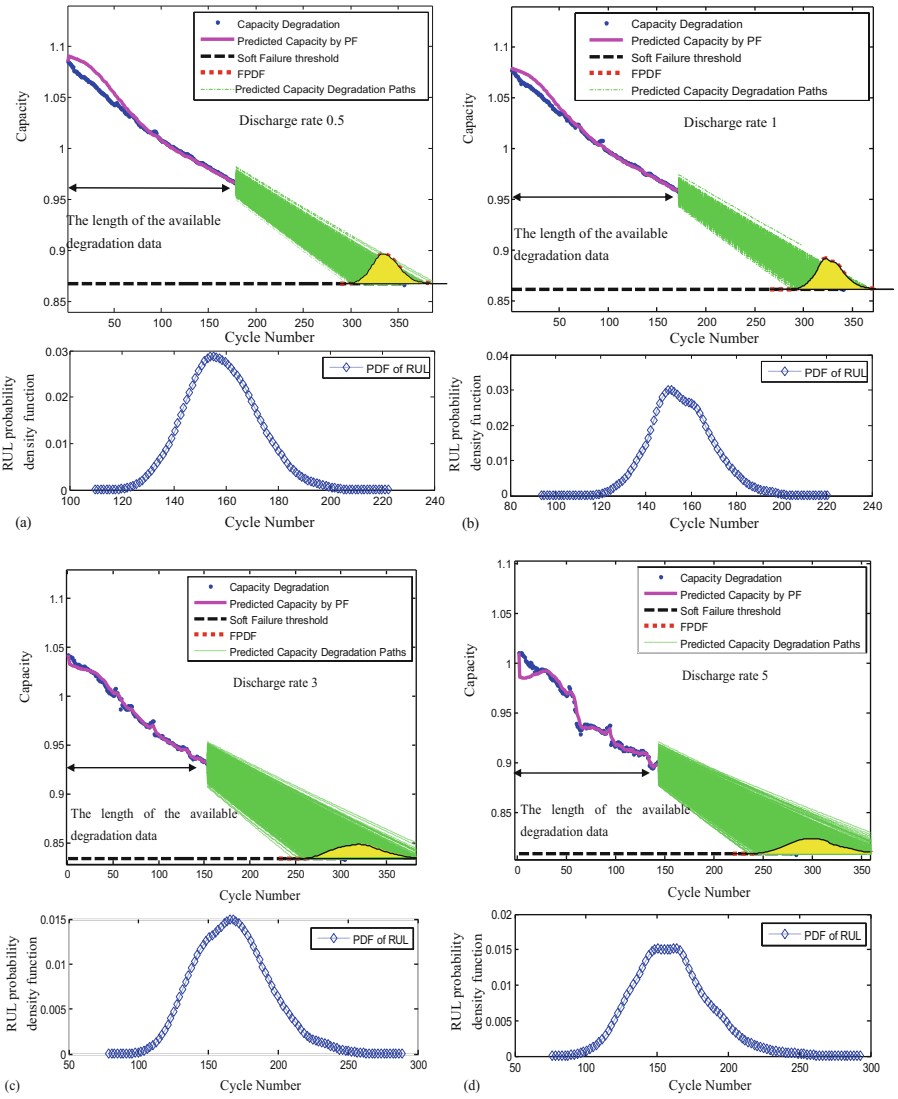

Fig. 16.10 RUL prediction at the different discharge rates by using the proposed discharge rates-dependent state space model based prognostic method in the case of battery 4: (**a**) RUL prediction at 50% of the AL and discharge rate 0.5 C; (**b**) RUL prediction at 50% of the AL and discharge rate 1 C; (**c**) RUL prediction at 50% of the AL and discharge rate 3 C; (**d**) RUL prediction at 50% of the AL and discharge rate 5 C

as the prediction accuracies at discharge rates 0.5 C, 1 C and 3 C because the raw capacity degradation data at discharge rate 5 C are more fluctuated. According to our literature review, because RUL predictions at the different discharge rates are seldom reported, no fair comparisons with other prognostic methods are conducted.

Table 16.3 RUL predictions obtained by using the proposed discharge rates-dependent state space model based prognostic method at the different discharge rates (Unit: Cycle)

Lithium-ion battery number	Discharge rate (Unit: C)	Prediction time (50%, 75% and 90% of AL)	True RUL	The percentile of RUL and absolute error obtained by using the proposed prognostic method			
4	0.5			5%	50%	95%	Absolute error
		178 (50%)	178	137	158	182	20
		267 (75%)	89	66	85	108	4
		320 (90%)	36	6	24	43	12
	1	171 (50%)	170	135	155	179	15
		256 (75%)	85	63	86	111	1
		307 (90%)	34	5	21	37	13
	3	152 (50%)	152	128	166	213	14
		228 (75%)	76	64	87	117	11
		274 (90%)	30	14	36	60	6
	5	142 (50%)	141	121	159	207	18
		212 (75%)	71	63	91	128	20
		255 (90%)	28	28	56	90	28

16.5 Discussions

Regardless of our prognostic methods introduced in Sects. 16.3 and 16.4, in which the two problems of the state space model based battery prognostics, including the design of a proper importance function for the use of particle filtering and the RUL prediction at the different discharge rates, were partly solved, there remain many important challenges.

Firstly, a physical battery capacity degradation model rather than empirical battery capacity degradation models generated by fitting historical battery degradation data should be proposed to well describe battery capacity degradation. The main reason is that an empirical battery capacity degradation model may be only suitable to specific battery capacity degradation data generated from specific battery types rather than general battery capacity degradation data generated from many other battery types. Even though an empirical battery capacity degradation model with more parameters, such as the sum of two exponential functions, is able to fit various battery degradation data, it suffers from the overfitting problem and causes inaccurate battery RUL prediction. On the other hand, if the physical battery capacity degradation model is difficult to be established, it is necessary to fix a specific battery type and investigate which empirical degradation model is the best choice for the specific battery type at some significant degradation factors, such as discharge rates, temperatures, charge rates, depth of discharge and time intervals between full charge cycles. Only in this way, we can know how good an empirical model for a specific battery type is for such battery prognostics.

Secondly, how to systematically and automatically initialize state space models is not addressable. Initialization of state space models include initial state parameters, initial state noises variances and initial measurement noise variances. If initial parameters are improperly set, it is difficult to timely track the current battery capacity degradation. For example, if a large measurement noise variance is used, the convergence to true measurements is difficult to achieve; if a small state noise variance is used, the convergence to true states is very slow. Moreover, it is necessary to investigate whether these noise variances should be posteriorly updated over cycles and these posterior estimates are beneficial to battery RUL prediction.

Thirdly, battery prognostics should take different operating conditions, such as different discharge rates and different temperatures, into consideration. In other words, a unique prognostic model at different operating conditions should be built. For example, a unique model at different discharge rates and a fixed temperature could be built as a preliminary battery prognostic model. Indeed, such prognostic model matches with real applications, such as electric vehicles and hybrid electric vehicles, in which different or varying discharge rates are used. Consequently, our research strategy is given as follows: prognostic models at fixed operating conditions ➔ prognostic models at different operating conditions ➔ prognostic models at varying operating conditions. Here, different operating conditions mean some discrete operating conditions, such as different discharge rates at 0.5 C, 1 C, 3 C

and 5 C. Varying operating conditions represent more general operating conditions, such as varying discharge rates from 0.5 C to 5 C with an increment of 0.1 C.

From the discussions made in this section, it is not difficult to see that there are many gaps required to be filled in the research community of battery prognostics.

16.6 Conclusion Remarks

In the introduction, many works highly related to battery prognostics were reviewed. Based on our literature review, the two unsolved problems in the research community of battery prognostics were pointed out. To solve these two problems, three state space model based prognostic methods for the battery RUL prediction at the constant and different discharge rates were respectively introduced in Sects. 16.2, 16.3, and 16.4, where all battery degradation data were well fitted by the exponential function. Consequently, the exponential function was used as the empirical battery degradation model. In Sect. 16.2, the particle filtering based state space model was developed to predict battery RUL at a constant discharge rate. Because the importance function used in the particle filtering did not consider the recent battery degradation data to update the weights of random particles used in the particle filtering, the importance function might reduce battery RUL prediction accuracies in some cases, where battery degradation data had tangible influences on the importance function. To solve this problem, in Sect. 16.3, the importance function provided by the spherical cubature Kalman filtering was introduced to the particle filtering so as to form the spherical cubature particle filtering. The comparisons between the particle filtering based prognostic method and the spherical cubature particle filtering based prognostic method were conducted. The results showed that for all the testing samples, the spherical cubature particle filtering based prognostic method has higher RUL prediction accuracies, which connoted that the importance function provided by spherical cubature Kalman filtering was able to improve battery RUL prediction accuracy. In Sect. 16.4, the third prognostic method was proposed to predict battery RUL at the different discharge rates, including 0.5 C, 1C, 3C and 5C. According to the analyses of our experimental battery degradation data at the different discharge rates, we found the relationship between the amplitude/slope of the exponential function and the different discharge rates was linear. This finding was useful and interesting because we could establish the unique state space model called the discharge rates-dependent state space model to describe the battery degradation data at the different discharge rates. Then, particle filtering was introduced to solve the discharge rates-dependent state space model for the battery RUL prediction. The testing results showed that our new discharge rates-dependent state space model based prognostic method is able to predict battery RUL at the different discharge rates. At last, we made more discussions in the research community of battery prognostics and pointed out our future research direction.

Acknowledgement This work was supported in part by General Research Fund of City University of Hong Kong under Project 11216014, in part by the National Natural Science Foundation of China under Project 11471275 and Project 51505307), and in part by the Research Grants Council Theme-Based Research Scheme under Project T32-101/15-R. The authors would like to thank the reviewer for his/her valuable comments on this book chapter.

References

1. Tsui KL, Chen N, Zhou Q et al (2015) Prognostics and health management: a review on data driven approaches. Math Probl Eng 2015:1
2. Zhang J, Lee J (2011) A review on prognostics and health monitoring of Li-ion battery. J Power Sources 196:6007–6014
3. Pecht M (2008) Prognostics and health management of electronics. Wiley-Interscience, London
4. Si X-S, Wang W, Hu C-H et al (2011) Remaining useful life estimation – a review on the statistical data driven approaches. Eur J Oper Res 213:1–14
5. Farmann A, Waag W, Marongiu A et al (2015) Critical review of on-board capacity estimation techniques for lithium-ion batteries in electric and hybrid electric vehicles. J Power Sources 281:114–130
6. Xing Y, Ma EWM, Tsui KL et al (2011) Battery management systems in electric and hybrid vehicles. Energies 4:1840–1857
7. Ng SSY, Xing Y, Tsui KL (2014) A naive Bayes model for robust remaining useful life prediction of lithium-ion battery. Appl Energy 118:114–123
8. Rezvanizaniani SM, Liu Z, Chen Y et al (2014) Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility. J Power Sources 256:110–124
9. Burgess WL (2009) Valve regulated lead acid battery float service life estimation using a Kalman filter. J Power Sources 191:16–21
10. Meinhold RJ, Singpurwalla ND (1983) Understanding the Kalman filter. Am Stat 37:123–127
11. Saha B, Goebel K, Poll S et al (2009) Prognostics methods for battery health monitoring using a Bayesian framework. IEEE T Instrum Meas 58:291–296
12. Saha B, Goebel K, Christophersen J (2009) Comparison of prognostic algorithms for estimating remaining useful life of batteries. T I Meas Control 31:293–308
13. Singpurwalla ND, Polson NG, Soyer R From signal processing to particle filtering (An Incredible Journey).
14. Wang D, Miao Q, Zhou Q et al (2015) An intelligent prognostic system for gear performance degradation assessment and remaining useful life estimation. J Vib Acoust 137:021004–021004
15. Wang D, Miao Q (2015) Some Improvements on a general particle filter based Bayesian approach for extracting bearing fault features. J Vib Acoust 137:041016
16. He W, Williard N, Osterman M et al (2011) Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method. J Power Sources 196:10314–10321
17. Wang D, Miao Q, Pecht M (2013) Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. J Power Sources 239:253–264
18. Xing Y, Ma EW, Tsui K-L et al (2013) An ensemble model for predicting the remaining useful performance of lithium-ion batteries. Microelectron Reliab 53:811–820
19. Xian W, Long B, Li M et al (2014) Prognostics of lithium-ion batteries based on the Verhulst model, particle swarm optimization and particle filter. IEEE T. Instrum. Meas. 63:2–17

20. Li F, Xu J (2015) A new prognostics method for state of health estimation of lithium-ion batteries based on a mixture of Gaussian process models and particle filter. Microelectron Reliab 55:1035–1045
21. Dong H, Jin X, Lou Y et al (2014) Lithium-ion battery state of health monitoring and remaining useful life prediction based on support vector regression-particle filter. J Power Sources 271:114–123
22. Liu D, Luo Y, Liu J et al (2014) Lithium-ion battery remaining useful life estimation based on fusion nonlinear degradation AR model and RPF algorithm. Neural Comput Appl 25:557–572
23. Hu Y, Baraldi P, Di Maio F et al (2015) A particle filtering and kernel smoothing-based approach for new design component prognostics. Reliab Eng Syst Safe 134:19–31
24. Walker E, Rayman S, White RE (2015) Comparison of a particle filter and other state estimation methods for prognostics of lithium-ion batteries. J Power Sources 287:1–12
25. Arulampalam MS, Maskell S, Gordon N et al (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE T Signal Proces 50:174–188
26. Miao Q, Xie L, Cui H et al (2013) Remaining useful life prediction of lithium-ion battery with unscented particle filter technique. Microelectron Reliab 53:805–810
27. Park JI, Baek SH, Jeong MK et al (2009) Dual features functional support vector machines for fault detection of rechargeable batteries. Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on Systems 39:480–485
28. Särkkä S (2013) Bayesian filtering and smoothing. Cambridge University Press, Cambridge
29. Saha B, Goebel K (2009) Modeling Li-ion battery capacity depletion in a particle filtering framework. In: Proceedings of the annual conference of the prognostics and health management society, pp 1–10
30. Wang D, Yang F, Tsui KL et al (2016) Remaining useful life prediction of lithium-ion batteries based on spherical cubature particle filter. IEEE T. Instrum. Meas. 65:1282–1291
31. Wang D, Yang F, Zhao Y et al (2016) Prognostics of lithium-ion batteries at different discharge current rates. **Submitted**
32. Kitagawa G (1996) Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. J Comput Graph Stat 5:1–25

Chapter 17
On System Identification for Accelerated Destructive Degradation Testing of Nonlinear Dynamic Systems

Jacq Crous, Daniel Nicolas Wilke, Schalk Kok, Ding-Geng (Din) Chen, and Stephan Heyns

Abstract Accelerated destructive degradation testing is considered with the objective of reproducing high fatigue incidents for a severely nonlinear system in a lab environment. In the lab, a test specimen is mounted on servo hydraulic actuators which are then used to induce the same response in the system as was measured in field tests. Finding the inputs to the actuators that accurately induce the measured response in the system is crucial to the integrity of the testing procedure. The problem is an inverse problem, and often exhibits ill-posed characteristics. To this end a new method for system identification from time series data is developed and is shown to outperform current methods such as different variants of NARX and Hammerstein-Wiener models. From the results obtained it is concluded that an alternative method of data generation for accelerated destructive degradation on severely nonlinear systems in a lab context is required. Three methods are developed and tested on simulated data and it is shown that a prototype bootstrapping strategy is superior: using 400,000 data points generated by this strategy the input signals were predicted with mean square errors of 5.08e-4.

Keywords Accelerated destructive degradation • Time series analysis • Spanning basis transformation regression • Nonlinear system identification

J. Crous (✉) • D.N. Wilke • S. Kok • S. Heyns
Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa
e-mail: jacq.crous@gmail.com; nico.wilke@up.ac.za; schalk.kok@up.ac.za; stephan.heyns@up.ac.za

D.-G. (Din) Chen
School of Social Work & Department of Biostatistics, University of North Carolina, 27599, Chapel Hill, NC, USA

Department of Statistics, Univeristy of Pretoria, Pretoria, South Africa
e-mail: dinchen@email.unc.edu

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*, ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4_17

17.1 Introduction

Accelerated testing is ubiquitous in engineering manufacturing and certification to rapidly obtain reliability information. Numerous approaches to achieve this have been developed depending on the application under consideration. These include use-rate acceleration, aging-rate acceleration and increased intensity acceleration [14].

Use-rate acceleration is particularly applicable to systems that are either not continuously used or in continuous use but only exposed to discrete events that significantly affect the life of the system. To accelerate this scenario a lab scale test that can be used to expose a system more frequently to these events is often required. This approach is valid as long as the time-scale and cyclic rate through these events do not affect the cycles-to-failure distribution. Although use-rate acceleration is a well established technique in industry it may present challenges, particularly when the required input to the lab scale test is unknown or needs to be identified. Accurately reproducing the measured response of these discrete events that degrade a system is crucial to accurately assessing the degradation induced in the system by these events. An additional complexity often encountered in certification is that small sample sizes or even single products need to be tested, which commences one the product is set up in the lab. This lab test rig then needs to be identified before the actual accelerated degradation testing can commence. Since, system identification degrades the system, it is important to ensure that efficient and accurate approaches are considered when identifying the required input to a lab experiment. The aim of this study is to develop an approach that meet these demands.

Accelerated destructive degradation (ADD), particularly in this work refers to accelerating the process of degradation in the system by reproducing discrete events that contribute significantly to the degradation of the system. These discrete events are considered as high fatigue incidents which cause destructive degradation if the system is repeatedly subjected to it. In practice ADD testing is preformed by taking the system into a lab environment where servo hydraulic actuators are used to reconstruct measured responses corresponding to high fatigue incidents in the system. Accurately reproducing the measured conditions in the system is crucial to accurately assessing the degradation process of the system [20]. To this end the process of ADD can be divided into two distinct phases: Firstly a mathematical model is constructed given available data which is then used to determine the actuator loadings that will generate the desired response in the system. In general ADD testing is applied to a wide range of specimens, even within the same laboratory. For this reason black box models, which require no *a priory* information about the system to construct the required mathematical model, are used. Secondly the constructed model is used in conjunction with control strategies in order to trace the desired response profile by the system [8].

The mathematical modelling of dynamic systems assists in understanding their dynamic characteristics and predicting their development over time. For this purpose mathematical modelling of dynamic systems has been of great interest in system identification (SI) over the past few decades (see [13] and references). This interest

has developed beyond academic problems to real life systems which exhibit complicated nonlinear behaviour [23]. There are two distinct motivations for SI: the first is to identify the system in order to gain further insights into the workings of the system. This requires the model constructed by the SI algorithm to consist of parameters that can be traced to physical variables that capture the system's behaviour. The second motivation is to describe the system dynamics over time in order to make accurate predictions of the system's behaviour. The identification problem of a model to be used for response reconstruction is a more relaxed problem as there is no need to require physically justifiable model parameters or even that the model parameters are unique [3]. These approaches are not always necessarily different. In this study, both motivations are important, since both accurate response reconstructing and understanding the behaviour of the system when subjected to events discussed before are pertinent. To this end an alternative approach to time series analysis is developed to yield both accurate predictions and interpretability of the fitted mathematical model.

SI for nonlinear systems is often an example of an inverse problem which is ill-posed. This problem arises due to trying to find the inputs that would produce the observed responses, which may either not exist or not be unique [12, 21]. In the case of ADD the uniqueness of the solution is not of concern; rather it is the accuracy of the responses induced in the system that is pertinent [8]. The life cycle of a system is determined by low cycle high fatigue incidents [20]. In this work the in-service time of the system and the equivalent test time for a given event is the same. For this reason the severity levels of the measured field responses must be reproduced as accurately as possible [17].[1]

This work is concerned with time-domain methods, where the input-output data are time series. Hence, time domain-methods are considered, of which NARX (Nonlinear Auto-regression with eXogenous input) and NARMAX (nonlinear Auto-regression with Moving Averages and eXogenous input) are the most versatile in modelling nonlinear dynamics [13]. For this reason NARX and Hammerstein-Wiener (HW) models are used as benchmarks to evaluate the performance of the developed approach to SI.

In this work two sets of time series data are employed: The first data set is virtual dataset, obtained by simulating a nonlinear quarter car model driving along an irregular road profile. The input space consists of road profiles (these road profiles are generated randomly from the power spectral density prescribed by the ISO 8608 standard [1]) and the output space consists of the unsprung mass' response to the input. The second data set is an experimental data set where the physical system consists of a monoshock swing-arm motorcycle rear suspension where the motion of the platform on which the wheel rests is controlled by a servo hydraulic actuator. The actuator simulates an irregular road profile. As before the input space is numerous randomly generated road profiles, while the strain on the suspension is measured to constitute the output space of the data set [8].

[1] Method 514.6 Annex A.

17.2 System Identification

Black-box modelling of a dynamic system does not use prior knowledge of the system to choose a model structure, rather a family of structures are used which have been shown to be successful in the past [24]. Given input, $u(t)$, and output, $y(t)$, data the object is to construct an estimation of the system response that has the general form

$$y(t) = g(\phi(t), \theta) + v(t), \tag{17.1}$$

where ϕ is the regression vector, and is constructed by mapping the past input and output observations to the fixed dimensional regression vector up to time t. In general the estimated output is a vector, but in the case of a single input single output system $y(t)$ is a scalar. The finite dimensional parameter vector θ is used to parametrise the model structure. An additional term, v, is added to the estimation which represents the part of the response that is not based on past observations and is assumed to be small. The black-box SI process can thus be decomposed into two problems: construction of the regression vector ϕ and construction of the mapping from the regression vector to the output space.

17.2.1 NARX Models

The NARX method uses the measured input and output data as regressors. The number of observations that are used are selected beforehand. Various ways of building the nonlinear mapping, investigated by Kerschen et al. [13], Sjöberg et al. [24] and Aguirre and Letellier [2], include wavelet networks, neural networks and tree propagation algorithms. These approaches have also been implemented in the MATLAB system identification toolbox and are used to benchmark our work.

The general form of the NARX methods is given by

$$\begin{aligned} y(k) = F[&y(k-1), y(k-2), \ldots, y(k-n_y), \\ &u(k-d), u(k-d-1), \ldots, u(k-d-n_u)] + e(k), \end{aligned} \tag{17.2}$$

where F is a nonlinear functional and, $e(k)$, the error term. The error term is considered to be an independent sequence and thus all cross terms between the polynomial and error terms are neglected [3].

It has been shown that any input-output process can be expressed as in Eq. (17.2) [15, 16]. A system whose non-linearities are of a polynomial nature can be accurately modelled at all levels of excitation [13].

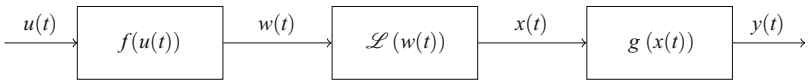

Fig. 17.1 Graphical representation of HW model structure. f and g are the input and output nonlinearities respectively and $\mathscr{L}$ is the linear regressor

17.2.2 Hammerstein-Wiener Models

The HW model uses two nonlinear mappings, one for the inputs f and the other on the outputs g. These two mappings are the so-called input and the output nonlinearity. A linear regression model, $\mathscr{L}$, is used to map the transformed input space to the transformed output space. The HW structure is illustrated in Fig. 17.1. This is also called a static nonlinear model. The input and output mappings can be built with a wide variety of mapping techniques. The nonlinear mappings however are memoryless and therefore these techniques are static nonlinear identification techniques. It can be shown that based on the models selected for the input nonlinearity and the linear regression model, either a NARX or Voltera model can be recovered [19].

17.3 Data Set Generation

In this section we outline the two systems that were used to generate the datasets used in this study.

17.3.1 Road Profile Generation

ISO 8608 (Mechanical Vibration – Road surface profiles – Reporting measured data) provides a means to classify road profiles according to different damage levels. Using these guidelines as a starting point it is possible to backtrack from the specifications given and reconstruct a road profile that adheres to a specified road profile class. The road profile classes specify the lower and upper limits of the values of the Power Spectral Density (PSD) for a specific class of road.

To generate the artificial road profile one uses a stochastic representation, which is a function of the PSD of vertical displacements. This PSD is obtained through a Fourier Transform of the auto-correlation function of the stochastic process describing the road profile [1]. The artificial road profile, $h(x)$, can then be generated from the following equation:

$$h(x) = \sum_{i=1}^{N} \sqrt{\Delta n} 2^k 10^{-3} \left(\frac{n_0}{in} \right) \cos\left(2\pi i \Delta n x + \phi_i\right), \tag{17.3}$$

where $x \in [0, L]$, where L is the total length of the road, ϕ denotes a random phase angle drawn from a uniform distribution over the interval $-\pi$ and π, and k is a constant that depends on the class of the road profile, and varies (discretely) between 3 to 9 corresponding to class A to H. The road profile is sampled at equally spaced distances over its length at a frequency of $\Delta n = \frac{1}{L}$. n is the spacial frequency and n_0 are the number of cycles per meter sampled.

17.3.2 Nonlinear Quarter Car Model

The vertical vibrations of a car can be approximated by means of a reduced model termed the quarter car model. The vehicle model consists of a two body system, a sprung and unsprung mass with two springs and a damper to represent the suspension of the car as well as the elasticity of the wheel [22]. A schematic of the system is shown in Fig. 17.2. This is an approximation of the quarter vehicle road simulator. In this work road profiles with unrealistically large amplitudes are generated. This is done specifically to ensure that a large degree of nonlinearity is manifested by the model.

For the nonlinear case the spring and damper between the sprung and unsprung masses are assumed to be nonlinear. For this case the damper was assumed to be a quadratic damper and the spring was chosen to be third order stiffening. This kind

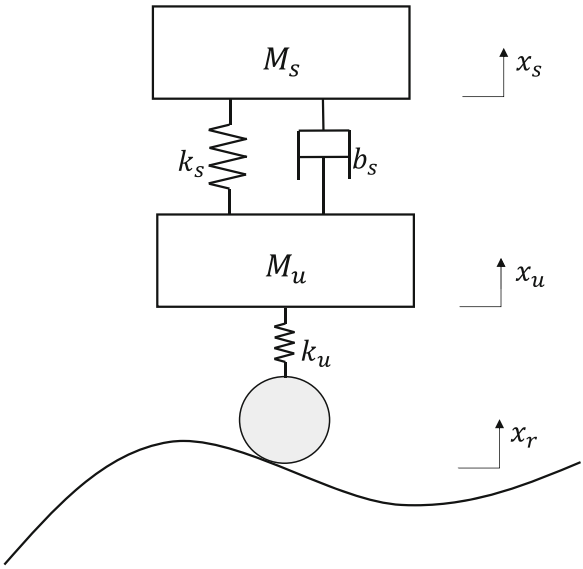

Fig. 17.2 Schematic of quarter car model. This model can be made nonlinear by replacing the spring, k_A, and damper, b_A, with their more realistic nonlinear counterparts

of spring is used with the Duffing oscillator [18]. The equations of motion for a nonlinear quarter car model are as follows:

$$M_A\ddot{x}_A + b_A(\dot{x}_A - \dot{x}_R)^2 + k_A(x_A - x_R) + k_{NL}(x_A - x_R)^3 = 0 \tag{17.4a}$$

$$M_R\ddot{x}_R - b_A(\dot{x}_A - \dot{x}_R)^2 - k_A(x_A - x_R) - k_{NL}(x_A - x_R)^3 + k_R(x_R - x_{road}) = 0, \tag{17.4b}$$

where the parameter k_{NL} can be used to tune the severity of the spring's nonlinearity. In this work the spring was chosen to be highly nonlinear. This produces a severely nonlinear system [3]. The non-linear system in Eq. (17.4) was solved using an exact Newton's method with variable time scaling [4]. For more detail regarding severely nonlinear system and the parameters of the quarter car model associated with it we refer the interested reader to the book by Billings et al. and the article by Agostinacchio et al. both cited in this work.

17.3.3 Quarter Vehicle Road Simulator

The quarter vehicle road simulator consist of a monoshock swing-arm motorcycle rear suspension. The frame of the simulator is fitted with rigidly connected dummy weights. This setup is shown in Fig. 17.3. Driving signals are generated from a PSD, as described in Sect. 17.3.1, and sent to a 40 kN linear actuator which excites the specimen. The data used in this work were the system response measured at the shock's lower bracket connection to the swing-arm. The measurements were taken with a 0°–90° rosette strain gauge. A/D and D/A conversions were done with 24 and 16 bit National Instrument systems respectively [8].

The quarter vehicle road simulator is a severely nonlinear system. The nonlinearities in the system are three fold: the damper and spring are both nonlinear. Secondly the geometry of the suspension introduces further nonlinearities. Finally when the system is tested for high fatigue, sudden impacts on the system causes the wheel to lose contact with the actuator.

17.4 Spanning Basis Transformation Regression for Time Domain System Identification

In this section the Spanning Basis Transformation Regression (SBTR) technique is introduced and strategies of applying SBTR to time series data are then discussed.

Fig. 17.3 Picture of a quarter vehicle road simulator driven by a servo hydraulic actuator

17.4.1 SBTR

The SBTR algorithm constructs a mapping from the principal components of the input space to the principal components of the output space [6]. In order to transform a component in the input space to a component in the output space a scaling and rotation transformation is applied to the components of the input space. Let $X \in \mathbb{R}^{p \times N}$ and $Y \in \mathbb{R}^{p \times N}$ be input and output data, where N are the number of observations and p the number of variables per observation. SBTR solves the following optimisation problem:

$$\begin{aligned} \min_{R,S} \quad & ||Z_x - RSZ_y||_F^2, \\ \text{such that} \quad & RR^T = I. \end{aligned} \tag{17.5}$$

The optimization problem solves for R, an orthogonal rotation matrix and S, a diagonal scaling matrix. Note that $\|\cdot\|_F^2$ is the Frobenius norm.The matrices Z_x and Z_y are composed of the eigenvectors of XX^T and YY^T respectively. These principal components can be interpreted as eigenfeatures [27]: Eigenfeatures are events in the time series that explain the highest amount of variance. Therefore the aim of a regression strategy is to find events in the time series data that have high variance.

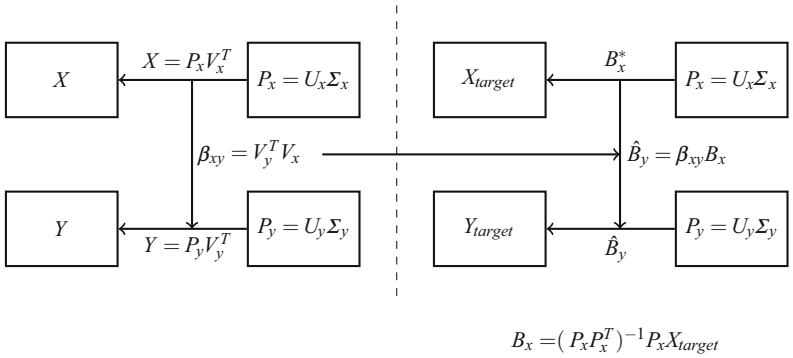

Fig. 17.4 Diagram of applying SBTR to time series data

These events then become features that are ranked according to the variance in the data set they explain. The regression algorithm then attempts to reconstruct a desired output based on these identified events.

The problem is therefore twofold: First the eigenvectors of XX^T and YY^T are found. Secondly these eigenfeatures of the input space are then regressed onto the eigenfeatures of the output space by means of a set of rotations and scalings.

In the context of system identification, the regressors are the eigenfeatures of the input and output spaces and the mapping is constructed by a series of scaling and rotation matrices.

The SBTR algorithm transforms one set of eigenfeatures into another. A diagram showing the general structure of SBTR and how it transforms the eigenfeatures of the input space to the eigenfeatures of the output space is shown in Fig. 17.4.

This application is based on recognising that eigenvectors and their corresponding eigenvalues change from the input to the output space for a nonlinear system. Therefore a series of rotation and scaling transformations can be used to map the input space's principal components to the corresponding output space's principal components. To this end a series of rotation and scaling transformations are sought that would map the data to the principal components in the input and output spaces respectively. In addition, a mapping is sought that would transform the aforementioned mapping in the input space to the corresponding mapping in the output space. Once all these mapping have been found a new mapping can be obtained that maps the principal components in the input space directly to the principal components in the output space.

In order to find the principal components of the input and output space a singular value decomposition (SVD) is performed on both of them. The SVD of a matrix is defined as follows [9]: Let $X \in \mathbb{R}^{m \times n}$, then the SVD of X decomposes it as:

$$X = U \Sigma V^T \tag{17.6}$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix, composed of the so-called singular values. The principal components of the matrix, X, is given by $P_x = U\Sigma$ and their corresponding eigenvalues are given by the square root of the singular values. After the principal components of the input and output spaces, P_x and P_y respectively, have been found, the input and output spaces can be reconstructed from these principal components given a set of transformations. These transformations are given by V_x^T and V_y^T (this follows simply from the definition of the SVD presented in Eq. (17.6)). Since by definition V_x^T and V_y^T are orthogonal matrices, it follows that these matrices are rotation matrices [27]. Since both these sets transform the principal components into their respective spaces by means of a set of rotations, it follows that a mapping can be found from V_x^T to V_y^T. This is done by multiplying these rotation matrices with one another to yield a relative rotation: this single rotation applied to a set of data transforms the data in the same way as when first applying V_x^T followed by V_y^T. This is the first part of the algorithm as described on the left hand side of Fig. 17.4.

The mapping from V_x^T to V_y^T forms the basis of the algorithm, since this mapping is used to predict what the output space transformation should be based on the input space transformation found by reconstructing the input target from P_x. This last step is shown on the right hand side of Fig. 17.4.

17.4.2 An Alternative View of Time Series Data

This section presents an alternative view of time series data. The benefits of this alternative view are not restricted to accurate predictions, but also an intuitive way of viewing time series data: instead of viewing it as two long strings of data we divide the data into short events within the data. This is illustrated in Fig. 17.5.

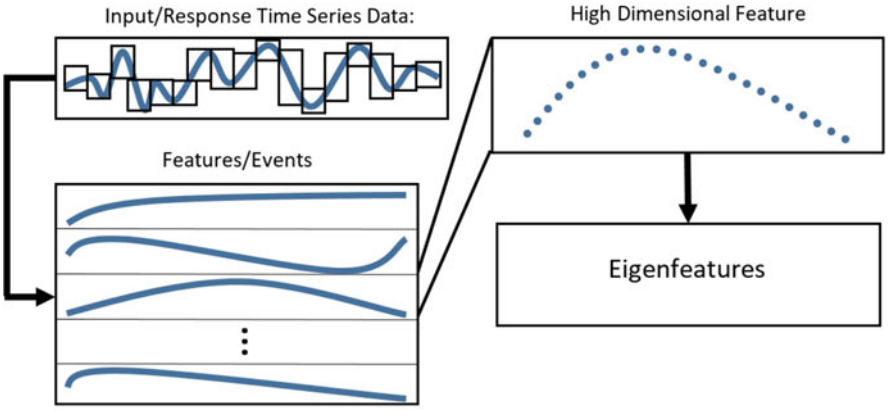

Fig. 17.5 Deriving eigenfeatures from time series data by breaking a long signal into shorter time signals

Once the data has been divided into features/events the object is to find reoccurring events which can then be used to describe the behaviour of the system, and these are called eigenfeatures. These are the principal components in the feature space and therefore correspond to the events that describe the largest amount of variance in the event space (event space is used here synonymously with feature space). Once the data has been processed as outlined above, Reduced-Rank Linear Discriminant Analysis can be used to project the data to low dimensions in order to obtain an informative view of the data [11].

17.4.3 Feature Lengths and Regularisation

The selection of the feature length has a significant impact on the performance of the regression algorithm. The feature length can be related to the properties of the physical system as follows: if the mechanical system has little damping, the effect of past events take longer to dampen out, and in order to account for this larger feature, lengths need to be selected such that the largest part of the feature is not affected by previous events. On the other hand for highly damped mechanical systems the aforementioned effect is minimal and hence shorter features are selected. Although relating the physical characteristics of the system to the feature length in this way, it does not give a direct estimation of the feature length, but does show the general vicinity of where one should look for the feature lengths in order to obtain meaningful results from the regression. Once the general vicinity of the feature length has been determined Cross-Validation can be used to determine the optimal feature length. Figure 17.6 shows the average cross validation error as a function of feature length. The optimal feature length in this case was 250. The number of components retained in using SBTR was determined by k-fold cross validation [11].

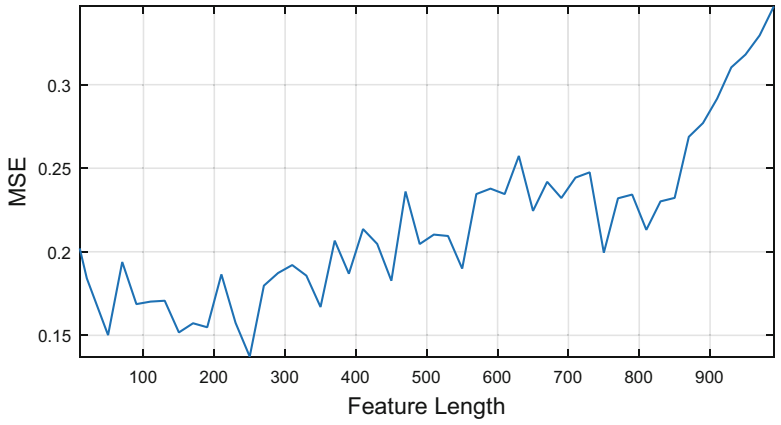

Fig. 17.6 Average Cross-Validation error as a function of feature length for the data generated by the nonlinear quarter car model

17.4.4 Visualising Model Performance

For a linear system represented by some function, f, there exists a corresponding matrix, A, which represents the mapping f [28]. Following the definition of linear maps it can be shown that the eigenvalues and eigenvectors of the input space will be the same as that of the output space: The eigenvalues and eigenvectors of a space are invariant under linear transformations. This however is no longer true for nonlinear mappings, and both the eigenvectors and their corresponding eigenvalues are allowed to change. This therefore gives a way to visualise the effect of the nonlinear mapping on the structure of the space on which it is operating.

When working with non-rectangular matrices the way to find the eigenvalues and their eigenvectors is by SVD. The left and right singular vectors are the principal components of the property and sample spaces respectively [27]. Thus in order to track the change in the singular values, which can be interpreted as the change in the participation of each principal component in reconstructing a given response, we define:

$$\Delta S = \Sigma_x - \Sigma_y. \tag{17.7}$$

Here the diagonal matrices Σ_x and Σ_y are the diagonal matrices obtained from the SVD of the input space and output space respectively. These diagonal matrices contain the singular values of the input and output spaces respectively. If the rank of the one of the spaces exceeds that of the other the number of singular values will not be the same. In this case ΔS is computed the corresponding singular values of each space, the remaining singular values are then unused.

The nonlinear quarter car model discussed in Sect. 17.3.2 was used with a high degree of nonlinearity brought about by a third order stiffening spring and quadratic relationship for the damping. One hundred road profiles, with 1000 data points each, were generated as discussed in Sect. 17.3. The principal component participation of the system and the corresponding modes constructed by NARX, HW and SBTR are shown in Fig. 17.7.

Secondly the principal components of the output property space YY^T, which are also called the property eigenvectors, are compared among the fitted models. The properties of the actual response is thus compared to the predicted properties. The first to third property eigenvectors are shown in Figs. 17.8, 17.9, and 17.10.

The change in mode participation gives information related to the severity and characteristics of the nonlinearity of the system. The difference between the output space property eigenvectors gives an indication as to how well the model is able to capture the response characteristics, and this in turn is related to how well the model will generalise. When comparing the MSE of the property eigenvectors and that of the testing MSE a direct correlation is seen. It is clearly seen from Figs. 17.8, 17.9, and 17.10 that SBTR performs better in capturing the properties of the response in the output space.

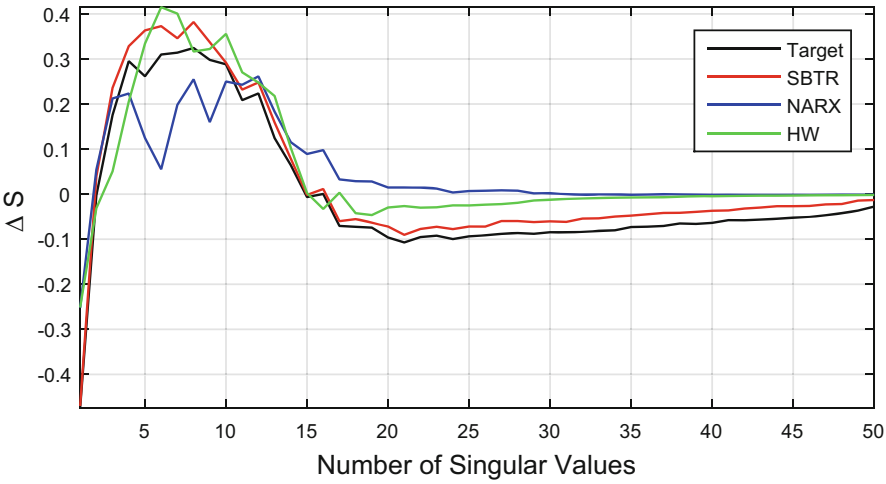

Fig. 17.7 The difference between the singular values of the input and output space, this gives an indication of the change in mode participation due to the nonlinearity in the system

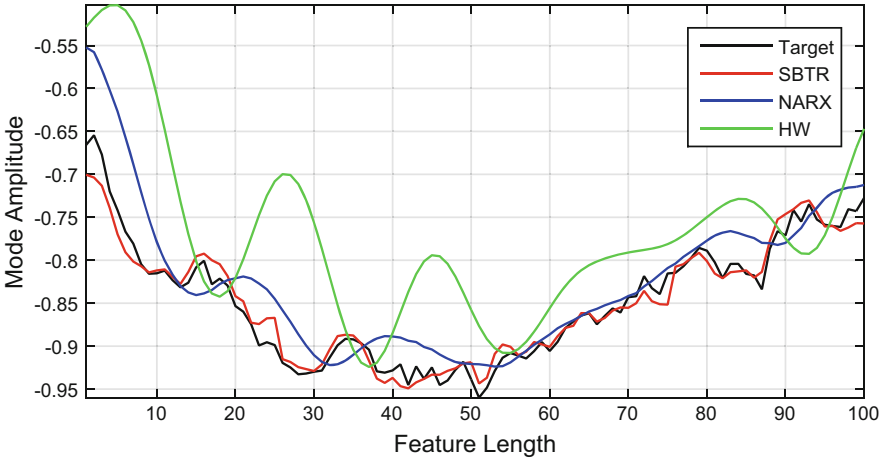

Fig. 17.8 First Property mode plotted for the System and NARX, SBTR and HW models

17.5 System Identification for Nonlinear Systems

After the quarter vehicle road simulator has been setup as depicted in Fig. 17.3, constructing a plant model of the system commences. When conducting ADD testing, one is presented with a twofold dilemma: on the one hand data is needed in order to construct an accurate plant model of the system, and on the other hand generating data in order to identify the system causes fatigue on the system thus compromising the integrity of the destructive degradation test objective.

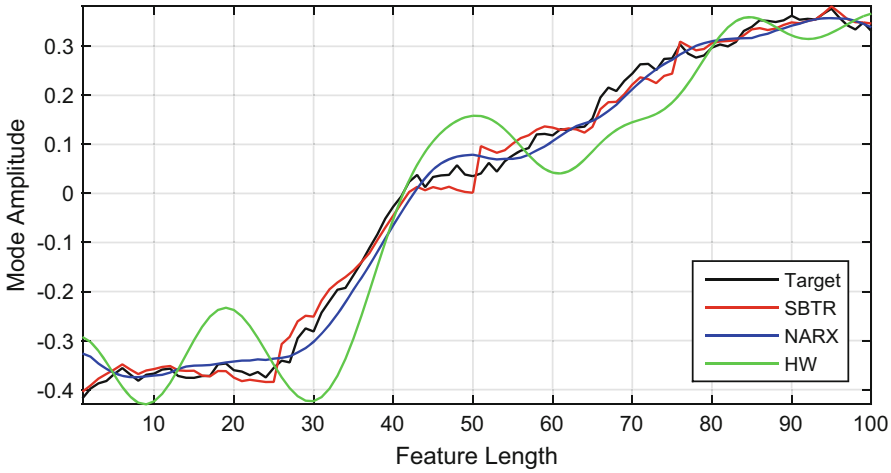

Fig. 17.9 Second Property mode plotted for the System and NARX, SBTR and HW models

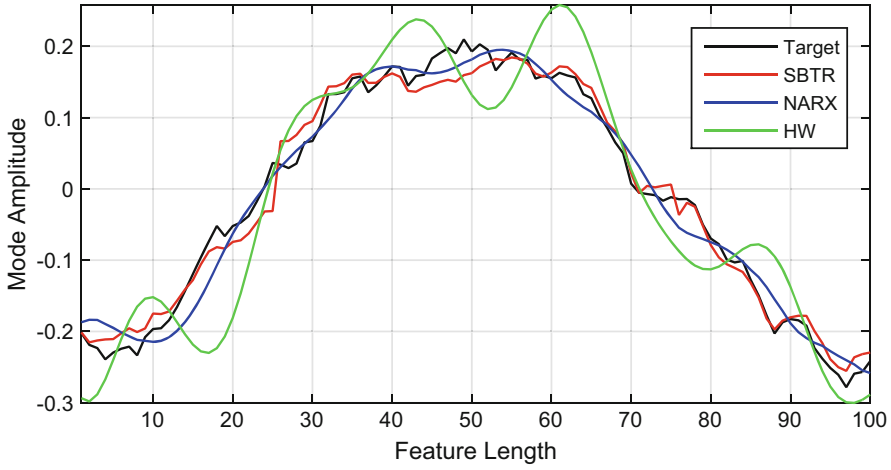

Fig. 17.10 Third Property mode plotted for the System and NARX, SBTR and HW models

Generating data for system identification is inevitable. Current strategies for data generation have been adopted from linear system identification [8]: this is based on generating a Power Spectral Density (PSD) distribution of the system that encapsulates the frequency content of the responses. This approach goes hand in hand with linear control where the frequency space is used to design controllers [25]. These linear methods are well established and are powerful tools for analysis. Therefore a reasonable first attempt to the problem is to ascertain if these methods can be adapted to a nonlinear setting [13].

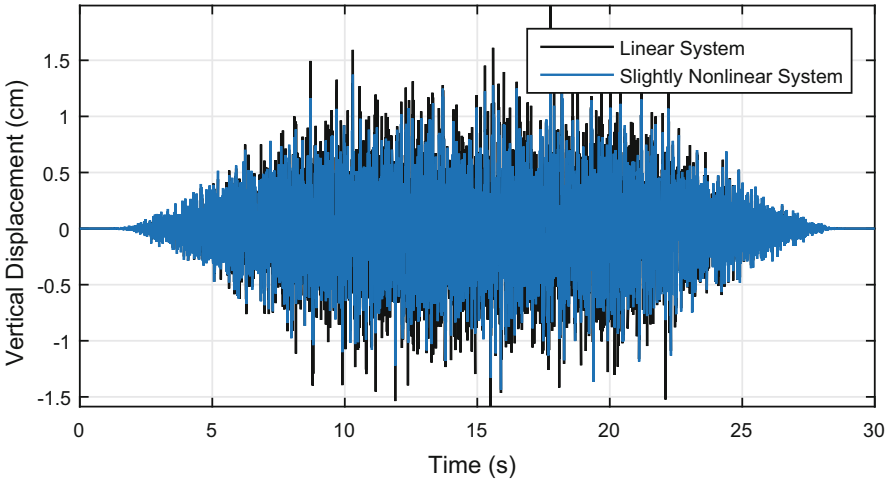

Fig. 17.11 Linear and slightly nonlinear systems' responses from PSD generated input data

To this end we consider a slightly nonlinear quarter car model configuration by appropriate choices for the stiffness and damping parameters. Figure 17.11 presents a plot of the responses of the slightly nonlinear system and a linear system from PSD generated input data.

Data for the system identification process was generated by randomly sampling the PSD to generate a random pulse in the frequency domain. Using an inverse Fast Fourier Transform (iFFT) the time series signal is derived. Each pulse has a duration of 30 s and is sampled at 1 kHz, thus each pulse consists of 30,000 data points. Six pulses were generated, five were then used to construct a plant model using HW, NARX and SBTR, and finally the last pulse was used for predicting and evaluating the model performance.

The HW model had a wavelet network with one unit for the input mapping and a piecewise linear output map consisting of 10 linear output maps. For the NARX model a sigmoid network with 10 hidden nodes was found to perform the best. The NARX model used 4 inputs and 2 outputs as the regressors, the delay between input and output was one data point. The feature length selected, using cross-validation, for the SBTR model was 100. Figure 17.12 shows the estimation made by each of the models for the entire pulse while Fig. 17.12 shows the estimation made between 12 and 13 s.

The time to compute the estimation was two orders of magnitude faster for SBTR compared to NARX and HW. The accuracy of the estimations were seen to be roughly the same, while NARX gives the best estimation when considering the application of destructive degradation. For the slightly nonlinear system the adoption of a linear approach to data generation was thus successful. The main benefit gained from the alternative approach to system identification discussed in this work is the computational time required to compute the estimation (Fig. 17.13).

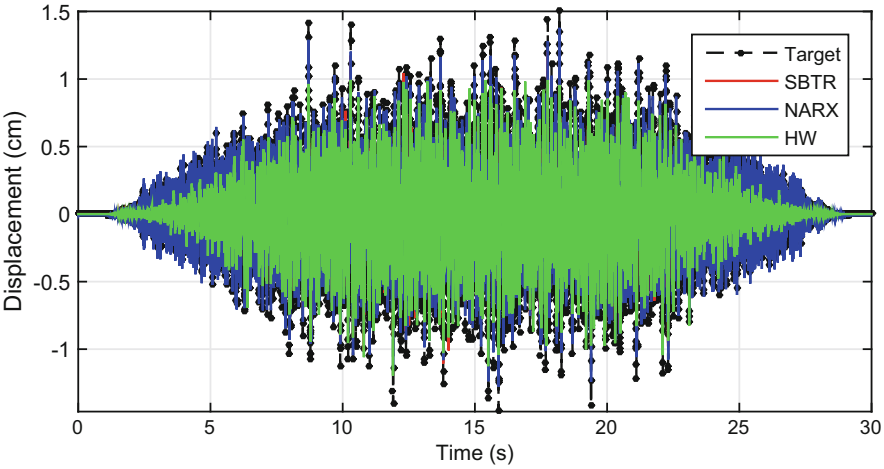

Fig. 17.12 NARX, SBTR and HW estimation of the entire input signal for a slightly nonlinear quarter car. The data used to construct the models were generated from a PSD. The mean square error for each of the models were: 0.0199, 0.0193, 0.0186 for NARX, SBTR and HW respectively. The runtime for each of the models were 30.5, 0.282, 45.21 seconds for NARX, SBTR and HW respectively

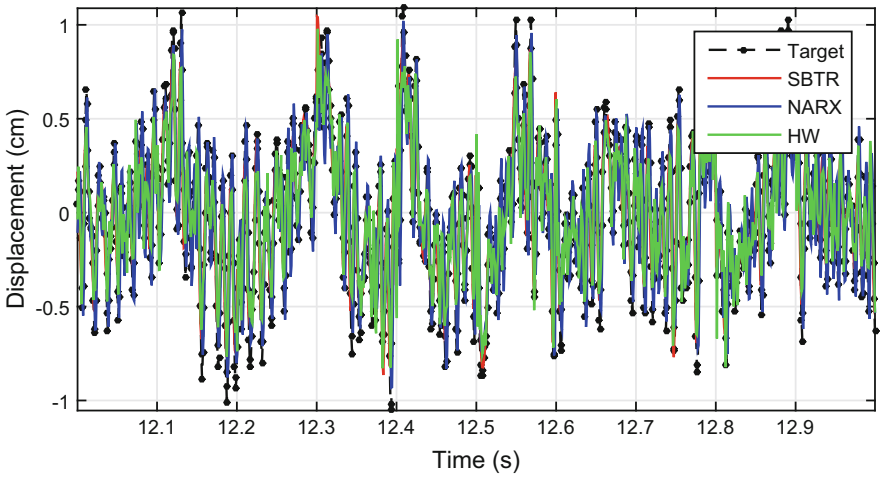

Fig. 17.13 NARX, SBTR and HW estimation of the input signal of the slightly nonlinear system between 12 and 13 s driven by PSD generated data

A similar conclusion was made when comparing linear system identification methods to the proposed alternative approach for linear systems [5].

The main point of interest though is not linear nor slightly nonlinear system, since high fatigue incidents usually occur when looking at real world system subjected to large displacements or forces. At this point we turn to the quarter vehicle road simulator which is a severely nonlinear system [3]. Following the approach outlined discussed in Sect. 17.4.4 the nonlinearity of the quarter vehicle road simulator is visualised in Fig. 17.14.

The results that follow are presented similarly to the results of the slightly nonlinear system. Figure 17.15 shows the estimated inputs for the entire time domain and Fig. 17.16 shows the estimated input between 12 and 13 s.

Figure 17.17 shows the change in mode participation due to the nonlinearity in the system, and the change in mode participation as traced by each of the algorithms are also presented.

In Figs. 17.18, 17.19, and 17.20 the first to third property modes are shown. It is quite clear from these figures that both NARX and HW are unable to capture the underlying characteristics of the system.

In the case of a severely nonlinear system the proposed technique, SBTR, is seen to be both computationally more efficient and more accurate. The issue when working with linear to moderately nonlinear systems is that collinearity in the data reduces the amount of information available in the data. Collinearity is a particularly prominent issue when the amplitude varies significantly through the time series data, as is the case with the pulse data generated from the PSD. In a nonlinear system more information is available since the system's response varies nonlinearly with amplitude [13]. Although more information is available in the data more data is also required to construct an adequate nonlinear model due to the higher complexity of the underlying structure. As a whole more information is available about the system but much less information is available for the various amplitudes at which the system is operating. Due to the amplitude changing the characteristics of the

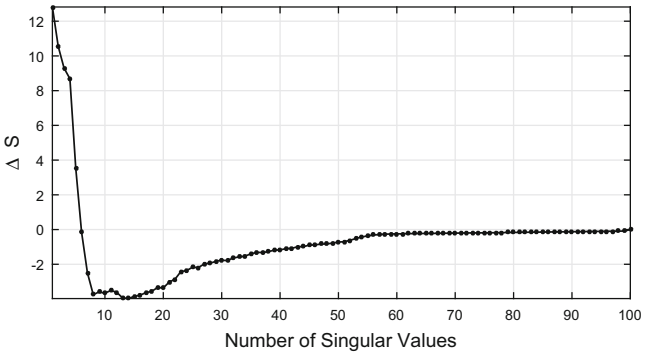

Fig. 17.14 A plot of the difference between the input space singular values and the output space singular values. The SVD was computed using feature lengths of 100. The plot clearly shows that the quarter vehicle road simulator is a severely nonlinear system

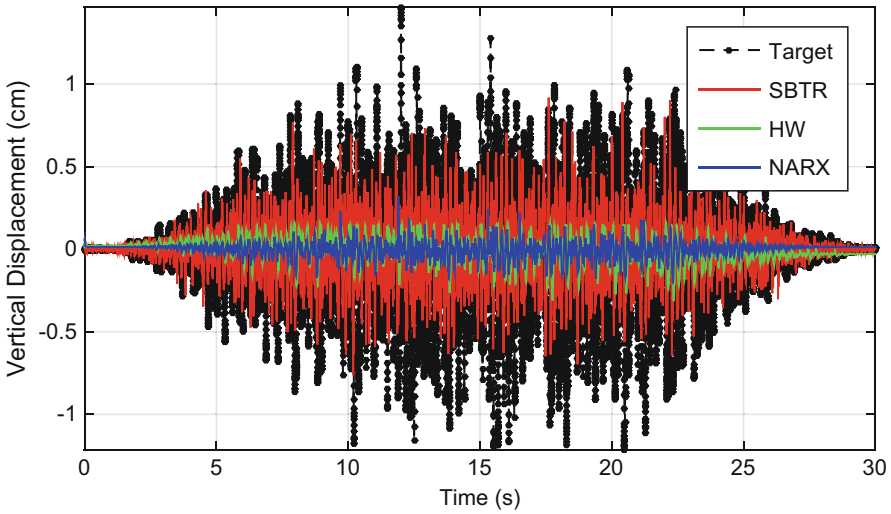

Fig. 17.15 NARX, SBTR and HW estimation of the entire input signal for the quarter vehicle road simulator. The data used to construct the models were generated from a PSD. The mean square error for each of the models were: 0.0687, 0.0320, 0.0664 for NARX, SBTR and HW respectively. The runtime for each of the models were 561.2, 1.052, 67.1, seconds for NARX, SBTR and HW respectively

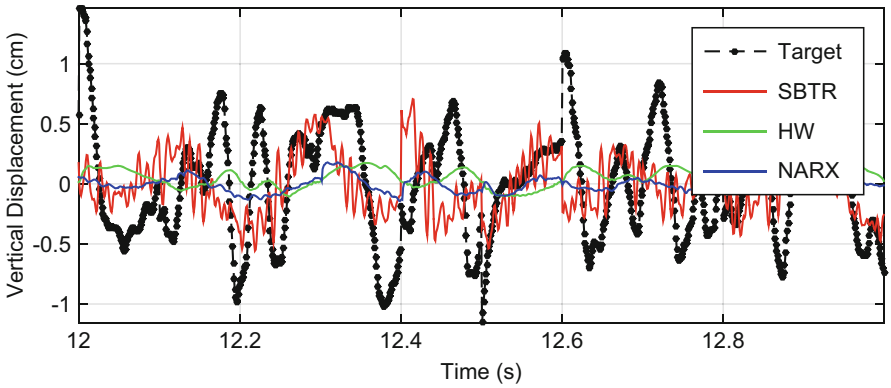

Fig. 17.16 NARX, SBTR and HW estimation of the input signal of the system between 12 and 13 s for the quarter vehicle road simulator. The data used for this estimation was generated by PSD data

system's response the feature space is much larger than that of a linear system. For this reason the data in the feature space is sparse and fitting nonlinear models to it does not yield sufficiently accurate results. This is remedied in the current work by means of more sophisticated control strategies that use the crude plant model and attempts to trace the desired target signal. To this end various machine learning strategies have been incorporated into control systems [7, 8, 10, 26]. The success

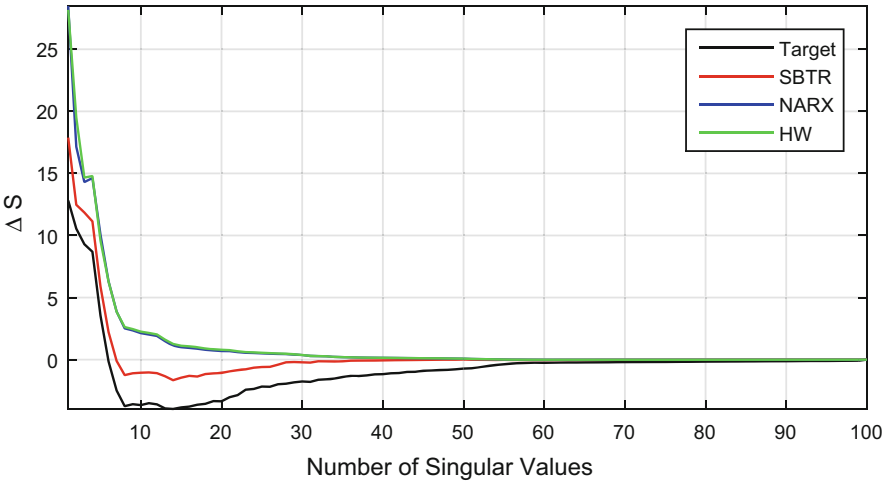

Fig. 17.17 Difference between the input and output space singular values for the data generated from the quarter vehicle road simulator

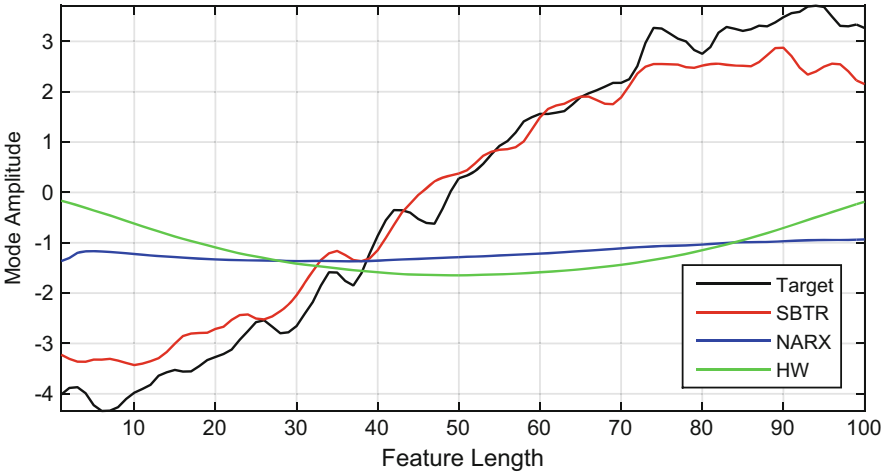

Fig. 17.18 The first property mode for the quarter vehicle road simulator and the estimated properties by NARX, SBTR and HW models respectively

of these methods varies according to the application. One advantage of the problem at hand, that differs from the cited work, is that we are presented with the desired signal that is to be reconstructed. Therefore an alternative means of data generation is sought that will generate relevant data: by relevant data we mean data that falls in the operating conditions of the target signal.

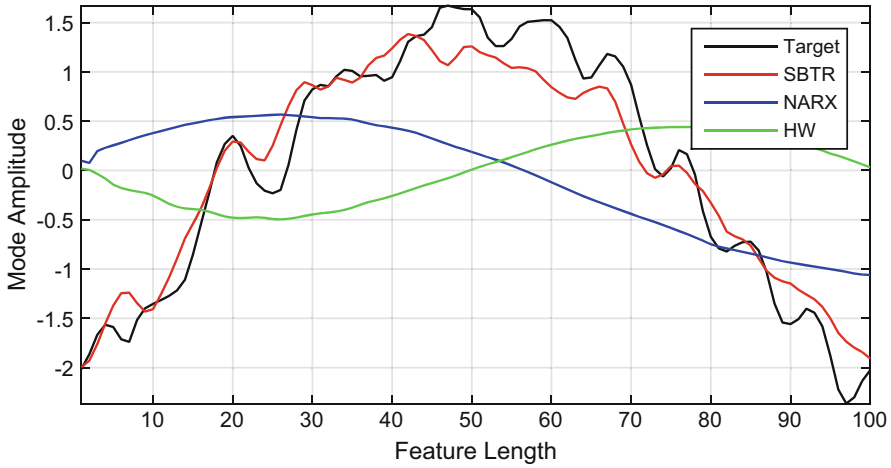

Fig. 17.19 The second property mode for the quarter vehicle road simulator and the estimated properties by NARX, SBTR and HW models respectively

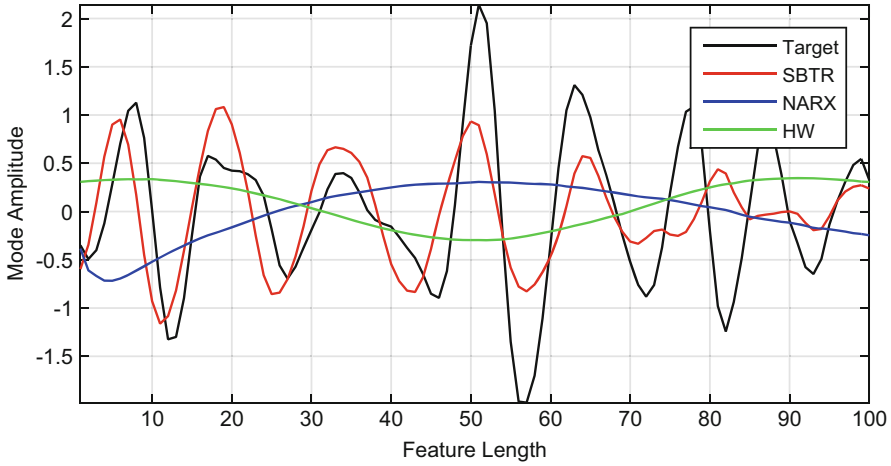

Fig. 17.20 The third property mode for the quarter vehicle road simulator and the estimated properties by NARX, SBTR and HW models respectively

17.6 Data Generation Strategies

In this section we look at three distinct strategies for generating the data that can be used to perform system identification for the quarter vehicle road simulator. These strategies however were developed using the severely nonlinear quarter car model discussed in Sect. 17.3.2. The only *a priory* information available is the response that is to be reconstructed of the system. Therefore the data generation strategies proceed

as follows: an initial signal is generated that sweeps a large range of amplitudes in order to determine the amplitude of the input signal that generated the target response. Once this is known the identified subspace within the feature space is to be populated with data in order to construct a plant model of the system for that particular subspace. Three distinct strategies for populating the aforementioned subspace are presented and evaluated.

The data used in this section was sampled at a higher rate than the data generated from the quarter vehicle road simulator. The data set used in this section each contained 400,000 sampled point in both the input and the output space respectively. The data generated from the quarter vehicle road simulator has 180,000 sampled points in both the input and output spaces respectively.

17.6.1 Initial Data Generation Signal

In order to construct an initial estimate of the input signal that will generate the desired response an initial signal is generated as follows: the servo hydraulic actuator used in the laboratory is able to excite 2.5 tons up to 100 Hz. Therefore we generate a signal by uniformly sampling the frequency domain from 0 to 100 Hz. A time signal can be generated from this using an iFFT. In order to produce a signal that is similar to the road profiles generated earlier we note that Eq. (17.3) weighs low frequency content higher than high frequency content. A similar approach is adopted here: low frequency content is weighted higher producing a low frequency trajectory which is perturbed by the high frequency content. Figure 17.21 presents the generated input signal the systems response.

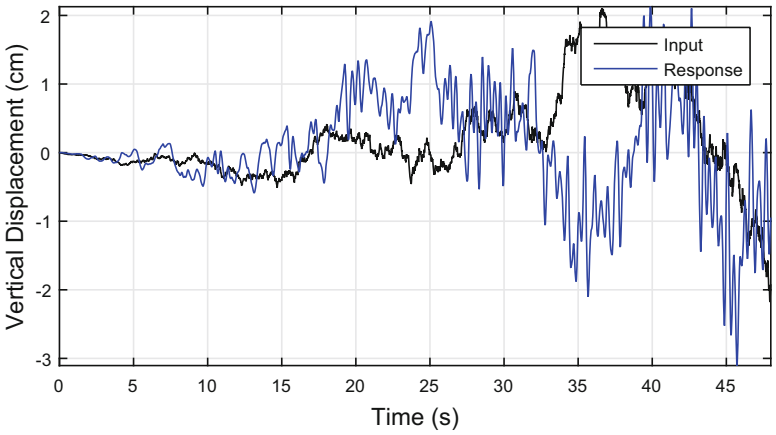

Fig. 17.21 Initial input signal generated with its corresponding response. These signals are used to compute an initial estimate of the input signal

This initial signal is divided into features which constitute the initial input and output feature spaces. This initial data is used with SBTR as described in Sect. 17.4.1.

17.6.2 Nonparametric Bootstrapping

A nonparametric bootstrapping strategy is a straight forward way to populate the input space after the initial estimation was computed. Since the amplitude of the desired input is known from the initial estimation, the approach is to generate more signals using Eq. (17.3) but scaled according to the amplitude range determined from the initial estimation. Therefore after each iteration the input and output spaces are supplemented with:

- Randomly generated signals using Eq. (17.3).
- The estimation made from the previous iteration with its corresponding response.

The convergence rate is slow as depicted in Fig. 17.22 for this crude approach. The error residuals are determined by comparing the actual desired input signal with the estimated signal and by comparing the actual desired response with the response from the estimated input signal.

The final MSE with 95% confidence bounds for the output generated from the estimated input signal was $(3.278 \pm 0.326) \times 10^{-3}$ The dataset at the final iteration

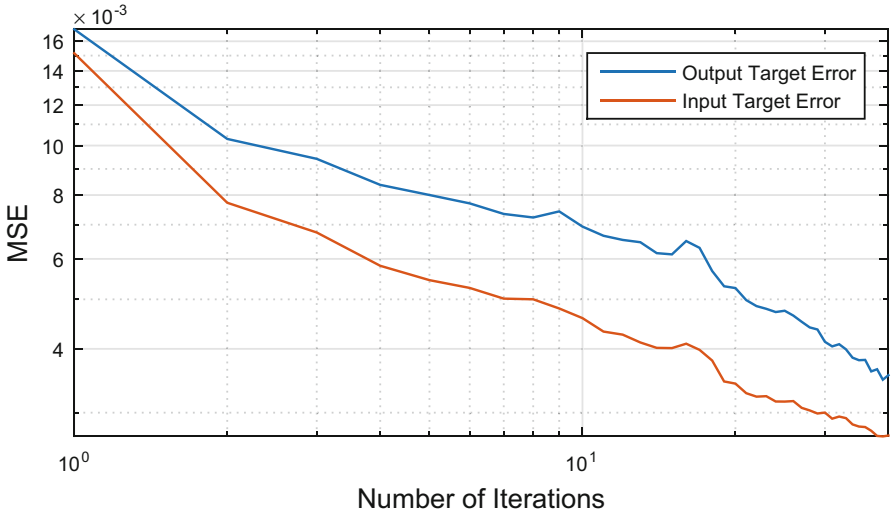

Fig. 17.22 Residual plot generated by comparing the estimated inputs and the corresponding response of the system to the target input. The data was generated using a nonparametric bootstrapping approach

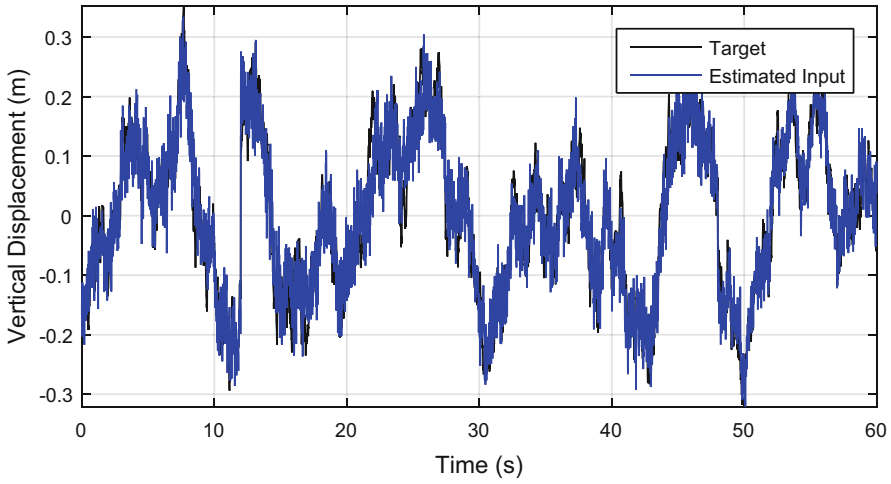

Fig. 17.23 Estimated input signal with the target input signal. Estimation was computed using data that was generated using the nonparametric bootstrap strategy

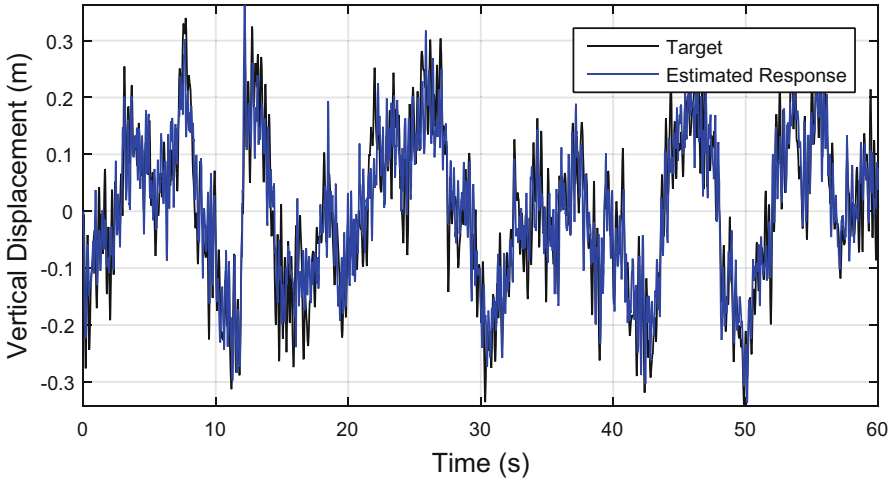

Fig. 17.24 The systems response to the estimated input signal with the desired output signal. Estimation was computed using data that was generated using the nonparametric bootstrap strategy

was 4000 by 100 (4000 is the number of observations and 100 the feature length). The data set size was the same for each of the three methods at the final iteration. The final estimation of the desired inputs are shown in Fig. 17.23 and the corresponding response of the system is shown in Fig. 17.24.

17.6.3 Parametric Bootstrapping

Parametric bootstrapping attempts to utilise the current estimation by using perturbations of the current estimates to supplement the input and output data sets. The current estimates are divided into features with feature lengths of 250. Therefore multiple estimates are presented in the feature space. A perturbation is generated by using Eq. (17.3) and scaling the generated signal to have a maximum amplitude of 5% of the maximum amplitude in the estimated signal. This random signal is then added to the current estimate in order to obtain a random perturbation of the current estimate which can then be used as a new input. Therefore at each iteration the input and output spaces are supplemented with:

- A set of randomly perturbed features based on the current estimate.
- The estimation made from the previous iteration with its corresponding response.

The MSE errors for the desired input and responses are plotted in Fig. 17.25.

The final MSE for the output generated from the estimated input signal with 95% confidence bounds were $(3.074 \pm 0.457) \times 10^{-3}$. The residual is also seen to start rising after the initial improvements made by supplementing the data. An important factor to the success of the Parametric bootstrapping approach is the number of components retrained in the SBTR algorithm. This implies that regularization is very important. This is not surprising as the initial estimates are rough and as more data becomes available the amount of information that SBTR is able to extract from the data increases. The number of components were selected at each iteration using k-fold cross validation.

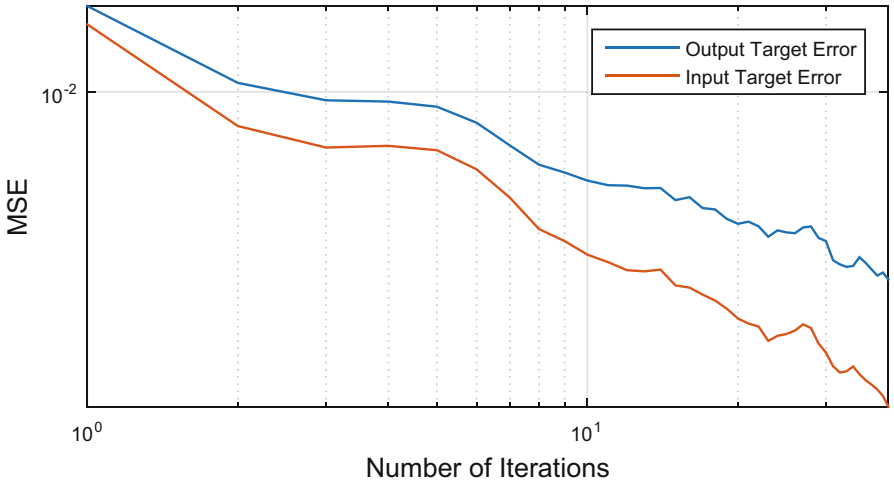

Fig. 17.25 Residual plot generated by comparing the estimated inputs and the corresponding response of the system to the target input. The data was generated using a parametric bootstrapping approach

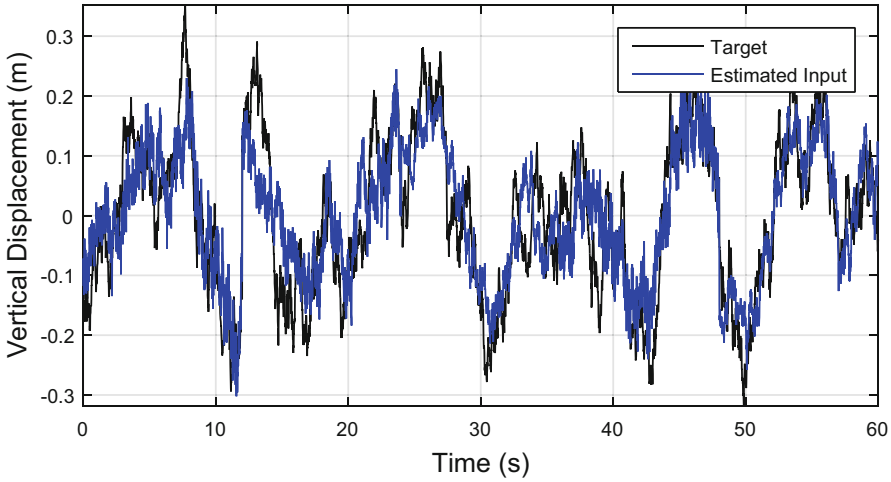

Fig. 17.26 Estimated input signal with the target input signal. Estimation was computed using data that was generated using the parametric bootstrap strategy

A further difficulty of the problem can also be seen from Figs. 17.22 and 17.25: One is able to estimate the input much more accurately than the corresponding response. This is the nature of ill-pose inverse problems [29], small changes in the inputs can generate large changes in the output. This can clearly be seen in both the nonparametric and parametric bootstrapping approaches when comparing the trace of the input target and the output target (Fig. 17.26).

Various amplitudes for the perturbations around the estimates were used and in this case a 5% of the maximum amplitude of the estimated input was seen to give the best results for perturbations between 2% to 15%. Although the parametric bootstrap approach proved only slightly more successful in estimating the input signal than the nonparametric strategy further improvements can be made by recognising the following concept: construct important centroids that represent a neighbourhood of data and sampling the feature space around these centroids (Fig. 17.27).

17.6.4 Prototype Bootstrapping with K-means Clustering

The prototype bootstrapping problem is an unsupervised learning problem where feature prototypes are constructed to represent a cluster of features. In this particular application the features describe events that explain important behaviours of the system. These prototypes would be an initial estimate of these features that are present in the target input signal. This approach can be viewed as a more sophisticated version of the parametric bootstrapping approach with one important difference: the prototypes are computed at every iteration, thus they are shifted

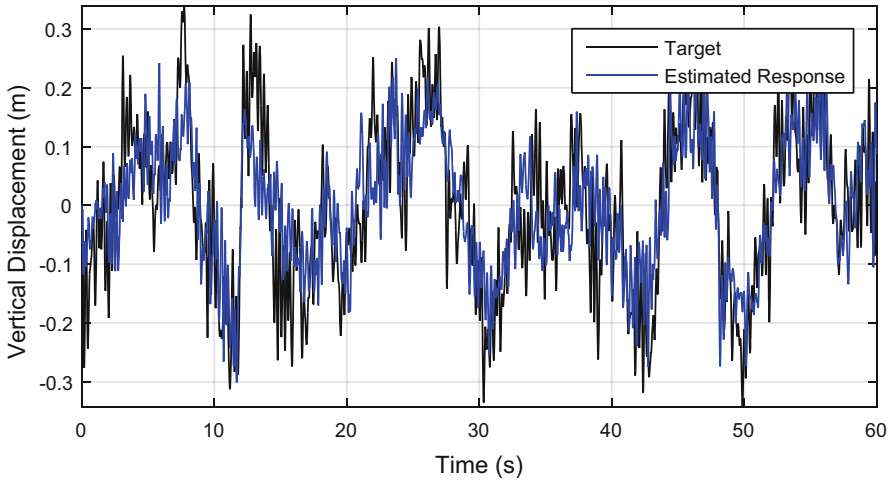

Fig. 17.27 Response of the estimated input signal with the target response signal. Estimation was computed using data that was generated using the parametric bootstrap strategy

in the feature space after each iteration. This allows for greater movement of the centres around which is sampled and thus avoids the problem faced in the parametric bootstrapping case where new data is being generated around poor centres.

In order to compute the prototypes K-means clustering is used. K-means clustering uses a distance metric to find centroids to clusters such that total cluster variance is minimised. This optimisation problem can be written as [11]:

$$\min_{C,\{m_k\}_1^K} \sum_{k=1}^{K} N_k \sum_{C(i)=k} ||x_i - m_k||^2, \tag{17.8}$$

where C is a cluster assignment, and m_k is the current mean for cluster k. The centroids of each cluster becomes a prototype and thus these centroids are perturbed in order supplement the data.

Clustering the features occurs in a much higher dimensional feature space. The feature lengths for generating the data was 1000 variables each. The reason for working in this higher dimensional space is that fewer clusters are needed to cluster the data and fewer targets exist in this space. Once the centroids have been perturbed the features are reshaped into the same feature lengths as was used for estimation in the previous case. Therefore at each iteration:

- The prototypes for the data is determined by K-means clustering in high dimensional feature space.
- Once the prototypes are determined they are perturbed. These input signals and their corresponding responses are added to the input and output data sets respectively.

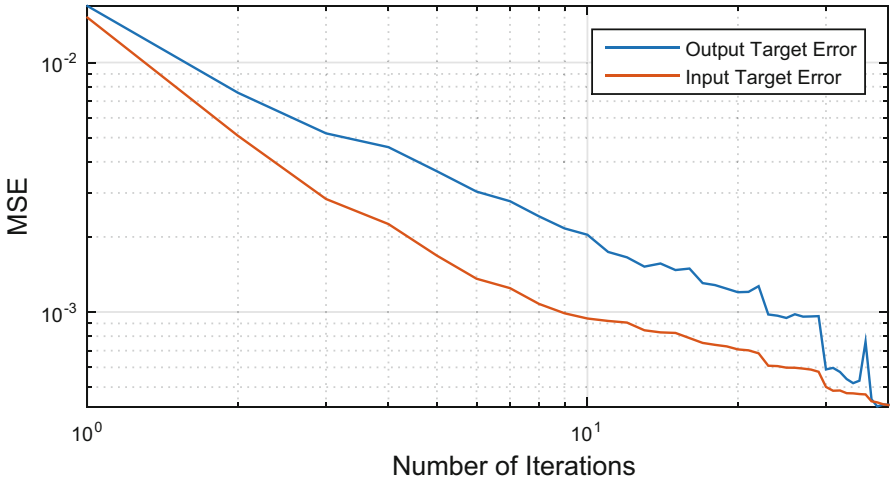

Fig. 17.28 Residual Plot generated by comparing the estimated inputs and the corresponding response of the system to the target input. The data was generated using a prototype bootstrapping approach

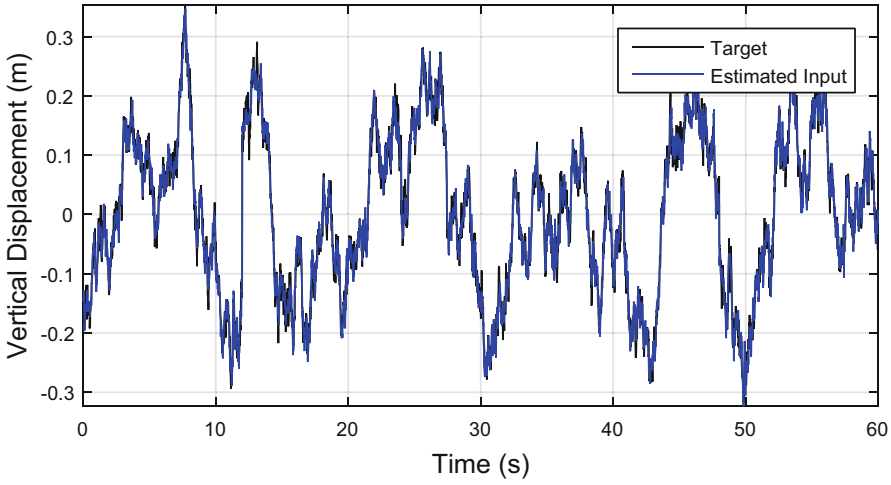

Fig. 17.29 Estimated input signal with the target input signal. Estimation was computed using data that was generated using a prototype bootstrapping approach

- The estimation made from the previous iteration with its corresponding response is added to the data set.

The residual for this approach is shown in Fig. 17.28.

Figure 17.28 shows both types of residuals are monotone decreasing. Therefore each iteration of the prototype bootstrapping approach supplements the data sets with meaningful data. Figures 17.29 and 17.30 show the estimated input and the

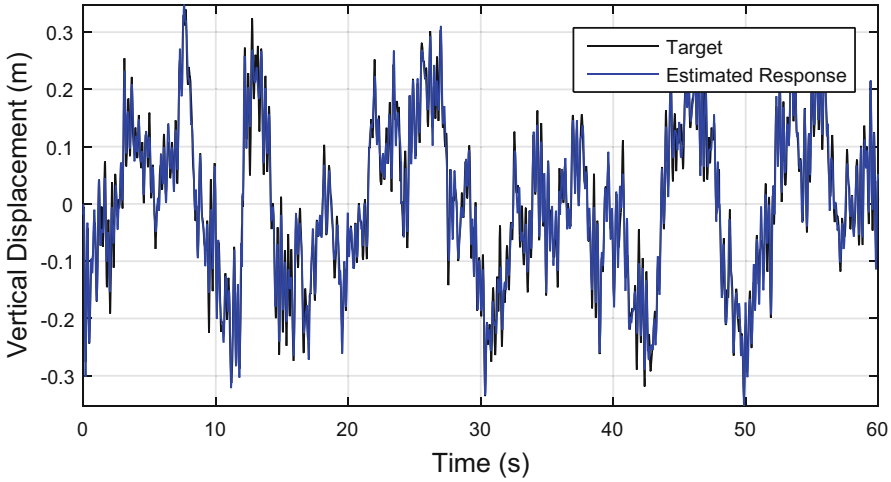

Fig. 17.30 Response of the estimated input signal with the target response signal. Estimation was computed using data that was generated using a prototype bootstrapping approach

corresponding response of the system respectively. The MSE for response generated from the system when the estimated input was used with 95% confidence bounds were found to be $(5.0785 \pm 0.8763) \times 10^{-4}$. Compared to the nonparametric and parametric bootstrapping approaches this is almost an order of magnitude better with a 84.5% and 83.4% improvement respectively.

It was shown that when using a prototype approach for data generation accurate estimations could be made with much less data. The aim of constructing a plant model for a nonlinear system is to describe the plant model accurately enough such that simple control strategies can be employed to trace the desired response. Therefore the main nonlinearity of the system is to be captured by the plant model, after which the variations around the estimated responses are small enough to be considered linear. The plant model is therefore used to obtain an estimate that is sufficiently close to the desired response such that linearisation can be applied.

17.7 Discussion and Conclusion

The integrity of ADD testing is dependent on how well the actual environmental conditions can be reconstructed in a test specimen in the laboratory. The environments of interest in this work induces high levels of stress in the system which in turn cause destructive degradation in the system. Reconstructing these events in the laboratory environment is vitally important in conducting reliable ADD testing. To this end an alternative approach to nonlinear SI is presented in this work. Comparing the proposed method, SBTR, with HW and NARX models on highly

nonlinear models it is shown that prediction accuracy is improved by 52% and 53% respectively. The reason for this is that SBTR is able to reconstruct the properties of the responses better than HW or NARX models are able to do.

For SI to be successful sufficient data is required to construct an adequate mathematical model to represent the system. The data generation techniques used for linear models can be adopted when working with slightly nonlinear systems. This however no longer holds true for highly nonlinear system. To this end three new data generation strategies were implemented, and compared. It is shown that a prototype bootstrapping strategy outperformed nonparametric and parametric bootstrapping strategies.

Further development of the prototype bootstrapping strategy is required, these include: looking at alternative ways of constructing prototypes and using an optimisation approach to perturb the prototype in the optimal way given the observed data. Finally this approach is to be implemented on the quarter vehicle road simulator in order to evaluate the performance of this strategy for real world systems.

References

1. Agostinacchio M, Ciampa D, Olita S (2014) The vibrations induced by surface irregularities in road pavements–a MATLAB® approach. Eur Transp Res Rev 6(3):267–275
2. Aguirre LA, Letellier C (2009) Modeling nonlinear dynamics and chaos: a review. Math Probl Eng. http://dx.doi.org/10.1155/2009/238960
3. Billings SA (2013) Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. Wiley, New York
4. Burden R, Faires J (2005) Numerical analysis. Cole, Belmont
5. Crous J, Kok S, Wilke DN, Heyns PS (2016) An alternative approach to system identification. In: Proceedings of the 10th South African conference on computational and applied mechanics
6. Crous J, Kok S, Wilke DN, Heyns PS (2016, Submitted to Machine Learning) A scaling-rotation based approach to multi-target regression in high dimensional spaces as an alternative to the standard projection based approaches
7. Eksteen J, Heyns P (2016) An alternative update formula for non-linear model-based iterative learning control. Inverse Prob Sci Eng 24(5):860–888
8. Eksteen JJA (2014) Advances in iterative learning control with application to structural dynamic response reconstruction. PhD thesis, University of Pretoria
9. Golub GH, Van Loan CF (2012) Matrix computations, vol. 3. JHU Press, Baltimore Maryland.
10. Han X, Xie W-F, Fu Z, Luo W (2011) Nonlinear systems identification using dynamic multi-time scale neural networks. Neurocomputing 74(17):3428–3439
11. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer, New York
12. Ikonen T, Peltokorpi O, Karhunen J (2015) Inverse ice-induced moment determination on the propeller of an ice-going vessel. Cold Reg Sci Technol 112:1–13
13. Kerschen G, Worden K, Vakakis AF, Golinval J-C (2006) Past, present and future of nonlinear system identification in structural dynamics. Mech Syst Signal Process 20(3):505–592
14. Kugler DL, Kramer LL (2003) Accelerated destructive degradation tests: data, models, and analysis. Math Stat Methods Reliab 7:319
15. Leontaritis I, Billings SA (1985) Input-output parametric models for non-linear systems part I: deterministic non-linear systems. Int J Control 41(2):303–328

16. Leontaritis I, Billings SA (1985) Input-output parametric models for non-linear systems part II: stochastic non-linear systems. Int J Control 41(2):329–344
17. MIL-STD-810G (2008) Enviromental engineering considerations and laboratory tests. Department of Defence
18. Moon FC (2004) Chaotic vibrations: an introduction for applied scientists and engineers. Wiley-Interscience, New York
19. Pearson R (2003) Selecting nonlinear model structures for computer control. J Process Control 13(1):1–26
20. Petracconi C, Ferreira S, Palma E (2010) Fatigue life simulation of a rear tow hook assembly of a passenger car. Eng Fail Anal 17(2):455–463
21. Pillonetto G, Dinuzzo F, Chen T, De Nicolao G, Ljung L (2014) Kernel methods in system identification, machine learning and function estimation: a survey. Automatica 50(3):657–682
22. Schramm D, Bardini R, Hiller M (2014) Vehicle dynamics. Springer, Berlin/Heidelberg
23. Sirca G, Adeli H (2012) System identification in structural engineering. Scientia Iranica 19(6):1355–1364
24. Sjöberg J, Zhang Q, Ljung L, Benveniste A, Delyon B, Glorennec P-Y, Hjalmarsson H, Juditsky A (1995) Nonlinear black-box modeling in system identification: a unified overview. Automatica 31(12):1691–1724
25. Slotine J-JE, Li W et al (1991) Applied nonlinear control, vol. 199. Prentice-Hall, Englewood Cliffs
26. Smolders K, Volckaert M, Swevers J (2008) Tracking control of nonlinear lumped mechanical continuous-time systems: a model-based iterative learning approach. Mech Syst Signal Process 22(8):1896–1916
27. Strang G (2007) Computational science and engineering, vol. 1. Wellesley-Cambridge Press, Wellesley
28. Strang G (2007) Linear algebra and its applications, vol. 1. Cengage Learning, Boston Massachusetts.
29. Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. SIAM, Philadelphia.

Index

D.-G. (Din) Chen et al. (eds.), *Statistical Modeling for Degradation Data*,
ICSA Book Series in Statistics, DOI 10.1007/978-981-10-5194-4