SOME CONTRIBUTIONS TO MAINTENANCE AND ACCELERATED DEGRADATION TEST UNDER COMPLEX FAILURE PROCESS

CHEN LIANGPENG
(*B.Eng.,Tsinghua University*)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

CLPG

CHEN LIANGPENG

16 January 2014

Acknowledgements

First and foremost I offer my sincerest gratitude to my main supervisor, Dr. Boray HUANG, for his guidance, patience, and encouragement throughout my Ph.D study. Without him, the thesis would not have been possible. I am also deeply indebted to my co-supervisor, Professor Loon Ching TANG. His wisdom and experience has illuminated my path to enter the world of reliability and solve practical problems. I am also grateful to my another co-supervisor, Professor Min XIE, who always provides timely help and great care to his students. I also would like to thank Dr. Zhi-Sheng YE, who offers continuous help, discussion, and encouragement for my research.

My sincere thanks also go to other ISE faculty members and office staff, who has helped me in one way or another. I am very grateful to my fellow graduate students in ISE for their friendship and company.

Last but not least, I would like to dedicate this dissertation to my family - my parents in China and my wife, Guangpu SUN. Your endless love has always been my strongest motivation to complete this dissertation.

Summary

This thesis investigates several practical issues in maintenance and accelerated degradation testing (ADT), which are two important techniques implemented in the product/system's life cycle reliability engineering. First off, the statistical analysis of repairable systems provides useful tools to characterize and predict the system failure behaviour. In view of the widely observed bathtub type failure rate and intensity during the system lifetime, we propose a flexible superposed piecewise constant intensity model, which also takes into consideration the possible substantial changes/shifts due to rectifications/reliability growth at failures or other time epochs. Next, we broaden the context to consider repairable production systems, and derive an optimal bivariate maintenance policy to achieve the cost efficiency. Utilizing the modern monitoring technology, the condition-based maintenance is facilitated in recent years, we propose a competing risk model to incorporate both soft failure due to natural degradation and traumatic failure due to random shocks. We then analyse the system reliability and obtain a periodic inspection schedule with degradation-threshold based preventive maintenance. While maintenance is normally performed when the product is deployed to the field use, ADT is carried out in design and verification phase before the mass production. Note that the underlying degradation of some devices in practice cannot be well described by the existing models in ADT literature, we propose the implementation of an inverse Gaussian process. Optimal testing plans are derived to achieve good statistical precision in estimating the product's important reliability index, such as the life percentile. Finally, we pay attention to the practical ADT planning considering the estimation bias incurred due to the heterogeneity of field conditions.

Keywords: reliability, life cycle, maintenance, degradation, stochastic process, testing.

Contents

List of Figures

List of Tables

Abbreviations

PCI	**P**iecewise **C**onstant **I**ntensity
NHPP	**N**on **H**omogeneous **P**oisson **P**rocess
MLE	**M**aximum **L**ikelihood **E**stimate
CDF	**C**umulative **D**istribution **F**unction
PDF	**P**robability **D**ensity **F**unction
LHS	**L**eft **H**and **S**ide
RHS	**R**ight **H**and **S**ide
IFR	**I**ncreasing **F**ailure **R**ate
EM	**E**xpectation **M**aximization
IG	**I**nverse **G**aussian

Symbols

Chapter 3

X_j	Random variable, system's lifetime between the $j-1$-th and j-th failure, fix/change
W	The length of replacement cycle
$C(N/T)$	The long run average cost with replacement schedule N/T
c_r	Cost of repair, fix/change
c_p	System replacement cost
$\alpha_1, \alpha_2, \delta_1, \delta_2$	Parameters of the PCI model
Θ	Parameter space for system replacement
B	Bootstrap sample size
Z	Random effect
N	Number of failures till a system replacement
T	Time to replacement

Chapter 4

d	demand rate
p	production rate
k	setup cost per production run
h	holding cost per unit per unit time

c_r	maintenance cost after system failure
c_p	maintenance cost before system failure
c_d	rework cost per defective unit
r	cost of each replacement
B	random variable,related to the rate of defective item production
Z_i	time (or age) to breakdown after the start of the i-th production run
α	parameter of imperfect maintenance effect
$AVC(\{T_i\}, N)/AVC(T, N)$	expected long run average cost under $(\{T_i\}, N)/(T, N)$policy
$CL(\{T_i\}, N)/CL(T, N)$	expected cycle length under $(\{T_i\}, N)/(T, N)$policy
$DE(\{T_i\}, N)/DE(T, N)$	expected total rework cost of defective units in a production cycle under $(\{T_i\}, N)/(T, N)$policy
$HC(\{T_i\}, N)/HC(T, N)$	expected maintenance cost in a production cycle under $(\{T_i\}, N)/(T, N)$policy
$SC(\{T_i\}, N)/SC(T, N)$	total production setup cost in a production cycle under $(\{T_i\}, N)/(T, N)$policy
$TC(\{T_i\}, N)/TC(T, N)$	total expected cost in a production cycle under $(\{T_i\}, N)/(T, N)$policy

T_i/T	time to perform maintenance after the start of i-th production run. In the case of (T, N) policy,$T_i = T, i = 1, 2, ..., N$.
N	number of production runs per production cycle
Chapter 6	
$Y(t)$	product degradation path
T_D	product lifetime
D	failure critical level
$\mathcal{IG}(a, b)$	the inverse Gaussian distribution
$\boldsymbol{\theta_l}$	parameter space of model l
ρ_q	product failure quantile
N	sample size
P	number of stress levels
m	measurement times
τ	measurement frequency
$\mathbf{I}(\boldsymbol{\theta}_l)$	Fisher information matrix under parameter space $\boldsymbol{\theta_l}$
$\psi(\cdot)$	digamma function
$\Phi(\cdot)$	standard normal CDF
$\phi(\cdot)$	standard normal PDF
$s_i, i = 1, \cdots, P$	stress levels
$n_i, i = 1, \cdots, P$	number of units allocated to s_i
Chapter 7	
$Y(t)$	product degradation path
T	product lifetime
D	failure critical level
$g(\xi, \eta)$	inthe Gamma distribution with shape parameter η and scale parameter ξ
$\boldsymbol{\theta}$	parameter space

θ_1, θ_2	parameter of the actual acceleration relationship
a, b	parameter of the transformed acceleration relationship
μ_s', σ_s'	distribution parameter of the actual stress level
μ_s, σ_s	distribution parameter of the transformed stress level
μ_ξ, σ_ξ	indistribution parameter of the acceleration relationship with Gamma process
$\gamma(s, x)$	lower incomplete Gamma function
t_w	warranty length
s'	actual stress level
s	transformed stress level
c_0	warranty cost per repair
$N_{wi}, i = 1, 2$	expected number of warranty repairs for two policies
$E(C_{wi}), i = 1, 2$	expected warranty cost for two policies
N	test sample size
P	number of stress levels
m	measurement times
$\triangle t$	measurement frequency
$\mathbf{I}(\boldsymbol{\theta})$	Fisher information matrix
$\psi(\cdot)$	digamma function
$\psi_1(\cdot)$	trigamma function

Chapter 1

INTRODUCTION

1.1 Background

Facing the intensive competition and customers' expectation, manufacturers today are under great pressures to improve the product's reliability during its life cycle. For example, in the design and development phase, maximum reliability needs to be built into the product. To verify the success of a completed design, a small number of prototypes are collected and tested in the verification and validation phase. Once a product is deployed into the field, a warranty usually accompanies with the product in which appropriate maintenance operations are performed to mitigate the product degradation and consequent failures in the field operation. On the contrary, ignoring degradation and failures results in significant loss. For example, Attardi et al. (2005) shows that the lack of reliability tests prior to product delivery will incur a substantial amount of failures in the automobile industry. In the electrical industry in US, there is 150 billion in economic losses due to outages related to reliability issues (Rouse and Kelly, 2011). To this end, many reliability techniques are involved and inherently related in the above phases, such as failure mode and effect analysis (FMEA), accelerated test, maintenance and warranty analysis, etc.

This thesis investigates maintenance and accelerated degradation tests (ADT) under complex failure processes. In reliability theory, the lifetime distribution model is adopted by most of the literature, due to the tradition and convenience, as well as its description for items' ageing characteristics and fitness to the data. However, the ageing nature of the lifetime model restricts its capability of interpreting either more essential or complicated situations encountered in recent research and practice. Some of these situations are summarized as follows:

- The reliability requirement for products is increasingly high to meet a sequence of specified performances. For example, an electronic product may be viewed as a complex system that consists of many components. To maintain high reliability for the entire system, it generally requires that the individual components have extremely high reliability (Lu and Meeker, 1993). Therefore, censored data may be collected with very few failures. Consequently, it is generally difficult to implement the lifetime model for inference.

- Products possess their own specific failure-generating mechanisms, many of which can be traced to an underlying degradation process. The unique failure mechanism, however, is unable to be captured by the lifetime model. On the contrary, a degradation model based on the product's physics, if appropriately chosen and readily observed by modern monitoring technology, will be more timely and informative (Meeker et al., 1998).

- At the system level, especially for repairable systems, component failures usually result in minor rectifications of the system, and system may experience a series of failures before it is completely overhauled. Typically, it is inadequate to use the lifetime model to characterize the system failure behaviour (recurrent events) during its lifetime.

To some extent, the above difficulties encountered can be readily resolved by using stochastic process models. Stochastic processes emerge as a new class of failure models (Singpurwalla, 1995) and receive increasing attention in various areas of reliability engineering. In the following, we briefly introduce the implementation of stochastic process models in the framework of both maintenance and accelerated degradation tests. Issues of current research are also addressed, based on which the motivations of this thesis are highlighted.

1.2 System maintenance modelling and optimization

System maintenance is very common in practice. Usually the system is repairable, meaning that when failing to perform its functions satisfactorily, it can be restored to fully satisfactory performance by some method other than replacement of the entire system. Consequently, recurrent events of failures/rectifications are observed during system lifetime. As these events occur randomly and inherently related to each other on some level, stochastic processes appear suitable models to characterize the *failure process* and determine future maintenance actions. Various models are proposed to accommodate different types of systems as well as their failure time data. For example, the renewal process, where the system is *as good as new* after a repair, and non-homogeneous Poisson process (NHPP), where the system always receives a minimal repair at failures, are two widely-used models. Statistical analysis of failure data and maintenance planning can be found in several studies, such as Jaturonnatee et al. (2006), Lawless et al. (2012), Pulcini (2014) and Rigdon and Basu (2000), to name a view.

Another class of models implement the stochastic process models to depict the underlying process of system degradation toward failures. As a result, maintenance is then performed to alleviate the degradation and prevent failures. Usually system degradation signals in one way or another, which can be measured directly or indirectly. With the advance of modern measurement technology and sensors, the monitoring of system real-time health becomes feasible within an increasing number of areas. For instance, the health condition of helicopter drive train system can be monitored by collecting the vibration signals from the hanger bearings, testing the electrical insulation degradation serves as a diagnostic tool of electric motor's health condition. With the degradation information on hand, stochastic process models are commonly chosen to characterize system degradation due to their flexibility to account for the correlation of time-dependent degradation measurement. More precise estimates of system reliability and better maintenance planning are then obtained. Optimal maintenance policies under different stochastic degradation models have been discussed by Liao et al. (2006), Dieulle et al. (2003a), Ye et al. (2012), Si et al. (2014), etc. An overview of the application of gamma process in maintenance can be found in Van Noortwijk (2009).

Although many models have been proposed on maintenance using stochastic process approach, a number of deficiencies still exist. First, for modelling the failure process of repairable systems, most of the existing models only describe either reliability improvement or reliability degradation, notwithstanding the fact of the well-known bathtub type failure intensity. In addition, possible substantial changes at failures, leading to reliability growth in practice, are seldom considered. Last but not least, most models assume single failure mode while this failure mode is subject to the degradation of system exceeding certain critical threshold. Nevertheless, most systems can fail due to a variety of failure modes or competing risks. Therefore, an integrated framework is desired for characterizing their joint effects.

1.3 Accelerated degradation test

An important program initiated in practice to obtain the reliability metric of developed products is to perform the tests in the earlier stage. It is extensively conducted in both design and production phases on materials, components and systems. However, the time duration allowed for testing is usually much shorter than the expected operating lives of products. In line with the modern quality philosophy for producing high-reliability products, most of products are designed to operate without failures for years, decades, or longer. Therefore, testing under normal use condition is costly and unrealistic. Accelerated tests are hence motivated to obtain timely information in which test units are exposed to harsh conditions. Degradation is accelerated and more failures occur. Reliability estimates under normal use level can be obtained by extrapolation from high level through some physically meaningful statistical model that links the stress with unit reliability.

In some cases, failures of sample units are rather frequently observed during test. However, for highly reliable products, failures seldom occur even under elevated stresses. To overcome this situation, degradation tests are facilitated, which measure some characteristic of interest in testing units. This characteristic represents unit's degradation gradually, and failure occurs when the degradation is not acceptable, i.e., exceeds some threshold. For example, the carbon film resistors may exhibit a shift on the resistance and fail when the shift is too large. In accelerated degradation test, stochastic processes are widely employed to model unit's degradation. For example, Wiener process and gamma process models are implemented in various ADT studies (Tang et al., 2004; Tseng et al., 2009; Lim and Yum, 2011).

Despite the wide applications of wiener process and gamma process models, there are circumstances where neither of the two models is appropriate. Some complementation to the family of degradation models are essential, and their applications in ADT need to be explored. For example, as a limiting process of the compound Poisson process, the inverse Gaussian processes are physically suitable to characterize the gradual growth of degradation, such as wear, crack, etc. Therefore, it is important to invoke the inverse Gaussian process for ADT study when it describes the degradation data well. Moreover, the objective of current ADT plans are focused on the life quantiles. Other objectives of practical importance should also be investigated, such as the fraction failings within a warranty period which is directly related to the field return of products, and the total warranty cost. Besides, the product's operating condition in the field is heterogeneous, such as the usage rate of products varies for different customers, the operating environment is different, etc. This may bias the estimation and thus should be considered if possible.

1.4 Research objective and structure

As indicated in the above comprehensive review, the gaps of current research in maintenance and accelerated degradation test under complex processes can be summarized as follows:

- The characterization of failure process of repairable systems usually implement the stochastic point processes where the failure intensity is naturally denoted by the arrival rate. However, the existing models depict either monotone increasing or decreasing failure intensity, ignoring the commonly observed bathtub type. Moreover, the continuity of point process implicitly assumes minimal repair or renewal of system whenever failure occurs, which is unable to account for maintenance degree in between or substantial changes (e.g. reliability growth) at failures or other time epochs.

- Stochastic process are widely used to describe system/component degradation. Complex systems usually fail due to a variety of sources. Few studies address the maintenance of degraded systems subject to multiple failure modes.

- The selection of an appropriate candidate model is the essential step in planning the accelerated degradation tests. Models include linear degradation path and stochastic processes are both advocated. Within the stochastic process category, only weiner process and gamma process are used, and it is found that some dataset of products in practice are not adequately fitted by these two models. Therefore, the accelerated degradation test of these products may be planned based on a new stochastic process.

- The objective of current research on ADT is focused on the product life quantile or D-optimality. However, since the main objective of ADT is to predict the product reliability in the field, the heterogeneous field condition needs to be considered.

This thesis intends to propose some practical models to resolve the above problems. Specifically, the objectives of this research are to:

- Develop a versatile model which takes into consideration the bathtub failure intensity of repairable system, as well as the substantial changes at failures or other epochs.
- Develop a bivariate optimal maintenance policy in manufacturing systems with lot sizing production and inventory.
- Develop an optimal periodic-inspection model for system maintenance with multiple dependent failure modes.
- Investigate the planning of ADT under a newly developed stochastic process model, i.e. the inverse Gaussian process.
- Develop an optimal ADT plan which incorporates the heterogeneity of field use.

The results of this study provide some new perspectives for maintenance and ADT under complex processes, which is helpful for the reliability decision making during the product's life cycle. Moreover, this research may be helpful in increasing:

- The understanding of the failure process of repairable system during its lifetime.
- The accuracy of maintenance planning for complex systems subject to various issues, e.g. production and inventory, multiple failure modes, etc.

- The diversity of model candidates and objectives in planning ADT.

The structure of this thesis is sketched in Figure 1.1. In Chapter 2, a detailed and comprehensive review of maintenance and ADT under complex processes is presented. Chapter 3 proposes a piecewise constant intensity model. Chapter 4 develops a bivariate maintenance policy for manufacturing system. Chapter 5 considers the maintenance planning for systems with multiple dependent failure modes. Chapter 6 investigate the inverse Gaussian process in ADT planning with and without random effects. Chapter 7 proposes an optimal ADT plan which considers the heterogeneity when product is deployed to the field. The conclusion for the whole thesis is given in Chapter 8, along with remarks and further research topics.

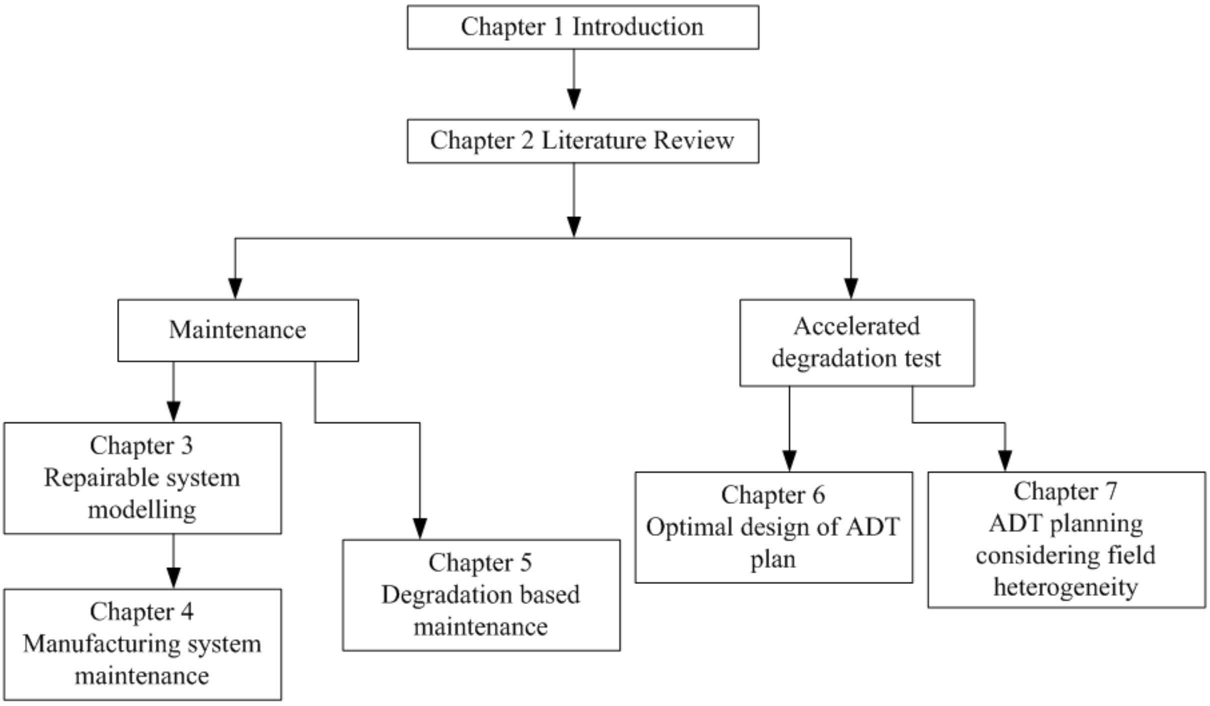


FIGURE 1.1: The structure of the thesis.

Chapter 2

LITERATURE REVIEW

In this chapter, we provide a detailed and comprehensive review of relevant current development on maintenance and accelerated degradation tests in line with the emphasis pointed out in Chapter 1. The first and second section focus on maintenance and ADT literature separately. Then the last section covers the joint studies on maintenance and reliability tests.

2.1 Maintenance modelling and optimization

The vast literature on maintenance study can date back to several decades ago (Barlow and Proschan, 1965). However, driven by the development of modern industries and management, new techniques, methodologies, approaches are continuously brought out by researchers and employed in a variety of applications (Murthy and Kobbacy, 2008). This section mainly concentrates on recent research from the perspective of stochastic process approach. Within this category, the first stream of research involves the characterization and statistical analysis of failure processes of repairable systems, while the second stream of research traces the underlying deterioration of systems and components with the help of modern monitoring technology.

2.1.1 Repairable systems

The manner that degradation and failures occur is often uncertain. The analysis of inter-failure time is of interest to the reliability community (Tang and Olorunniwo, 1989), as the knowledge of system's inter-failure time behaviour is helpful to understand the reliability growth of systems, which facilitates subsequent maintenances and other reliability programs. The concept and modelling of repairable systems using stochastic processes was first proposed in Ascher (1968). A treatment of book length can be found in Ascher and Feingold (1984) and Rigdon and Basu (2000).

Renewal process (RP) and non homogeneous Poisson process (NHPP) are two common approaches studied in literature. RP models the *as good as new* maintenance at each failure, while NHPP assumes minimal repair or *as bad as old* maintenance. A widely adopted NHPP is the power law process

(PLP) where the intensity function is $\lambda(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}$. For example, Rigdon (1998) presented a statistical analysis of the failure data of repairable systems using PLP. Gilardoni and Colosimo (2007) determined the estimate of optimal preventive maintenance time for repairable systems when the underlying process is PLP. Gaudoin et al. (2003) introduced a goodness-of-fit test for the PLP based on the Duane plot. Besides PLP, another class of NHPP is the log linear process (LLP) first proposed in Cox (1955), where the intensity function is $\lambda(t) = \exp(\alpha_0 + \alpha_1 t)$. Lee (1980) compared the adequacy of PLP and LLP using optimal conditional tests. Coetzee (1997) systematically addressed the application of two NHPP models in analysis of maintenance failure data.

Some recent research extends and generalizes RP and NHPP in a number of ways. Most of them are motivated by modelling the intermediate degree of maintenance between *as good as new* and *as bad as old*. The modulated power law process (MPLP) is one type of generalization of RP, where the inter-failure cumulative intensity are gamma distribution (Muralidharan, 2002). A further extension of the MPLP assumes that the distribution is arbitrary instead of gamma, which is called the trend renewal process (TRP). A fully characterization of TRP can be found in Lindqvist et al. (2003). Recently Yang et al. (2012) used the TRP to analyse systems with multiple failure modes. Other classes of models that deal with situations in-between perfect and minimal repair include the Brown-Proschan model (Brown and

Proschan, 1983), the Kijima's virtual age model (Kijima, 1989) and the arithmetic reduction of age (ARA) and arithmetic reduction of intensity (ARI) model (Doyen and Gaudoin, 2004) wherein a bulk of extensions can be found respectively in the literature.

In parallel, a piecewise exponential model (PEXP) generalizes RP/HPP by allowing the inter-failure times to be independent but not identically distributed random variable. The PEXP was first proposed in Sen and Bhattacharyya (1993) and Sen (1998) in view of the reliability growth in product development phase. Since the reliability growth or other substantial changes also occur with maintenance actions in both development and production phase, Rigdon and Basu (2000) incorporated PEXP into reliability and maintenance studies. A recent study by Arab et al. (2012) studied PEXP with two Bayes approaches: empirical Bayes approach and hierarchical Bayes approach. A class of geometric process (Lam, 2007) or quasi-renewal process (Wang and Pham, 1996) models were also addressed in literature which assume a geometric sequence of inter-failure times.

In addition to the above generalizations, some models describe the failure intensity focusing on its limiting behaviour. As indicated by Drenick limit theorem, the failure intensity of the repairable system will approach a constant after a sufficiently long time when the system consists of a mixture of parts with randomized mix of ages. Pulcini (2001) proposed a bounded intensity process for the repairable system and further analysed the reliability under various operating conditions in Pulcini (2008).

Most of the models above result in a monotonic trend of failure intensity, ignoring the fact of bathtub behaviour that commonly observed in systems and components. Moreover, maintenance utilizing these stochastic process models needs to be properly scheduled when other issues are incorporated such as production, inventory, etc.

2.1.2 Condition based maintenance

Another perspective of handling the manner of degradation and failure of system/component is to directly characterize its degradation process before failures. Condition based maintenance can thus be adaptively and effectively planned. The underlying degradation varies from system to system, and the captured degradation is usually more informative and provide more precise estimate of system reliability (Meeker and Escobar, 1998). According to the stochastic process models adopted, the literature on condition based maintenance of degrading systems can be classified into two types.

The first type of models utilize Markov processes and partition the system degradation into several discrete states (from *as good as new* to an absorbing state of *failure*). Maintenance actions are taken including minimal repair, replacement etc. An optimal policy is derived to minimize the relevant costs in which a series of decision rules are obtained and associated with each degradation state. The procedure is typically regarded as Markov decision process (MDP) and its special cases and variations. For example, Chen and Trivedi (2005) first proposed a semi-Markov decision process (SMDP) model and determined the state threshold for minimal and major maintenance. Maillart (2006) analytically derived the optimal maintenance policy when system is not directly observable at every time epoch, using a partially observed Markov decision process (POMDP). Makis and Jardine (1992) embedded MDP into the covariate of a proportional hazard model (PHM) and obtained the optimal replacement threshold. The advantage of using MDP is that the resulting optimal maintenance policy analytically corresponds to the system's deterioration, which is insightful and easy to implement. However, the obstacle is the classification of system degradation state, which is usually arbitrary and difficult to justify.

The second type of models describe the continuous evolution of system degradation using stochastic processes such as the Lévy process. Compared with MDP, the class of Lévy processes may provide a closer characterization of system degradation mechanism. For example, many degradation such as corrosion, fatigue crack growth and physical wear can be viewed as accumulations of additive and irreversible damage caused by a series of external shocks. The arrival of shocks may be approximated as Poisson process, each causes random and tiny wear. Then the gamma process, for instance, is a physically meaningful model since the gamma process is essentially the limit of a compound Poisson process with infinite jump rate and proportional infinitesimal jump size (Lawless and Crowder, 2004). Studies on maintenance policies with gamma process can be found in Dieulle et al. (2003a) and Liao et al. (2006). An overview of the application of gamma process in maintenance can be found in Van Noortwijk (2009). Wiener process was considered in Liu et al. (2012). Crowder and Lawless (2007) studied both gamma process and wiener process and proposed a predictive maintenance scheme. A more general situation with non-negative stationary and statistically independent degradation increments are treated in Grall et al. (2002), Deloux et al. (2009) and Lu et al. (2007), etc. Some physics-based models are proposed for specific problems (Wang, 2000; Peng et al., 2009). Despite the wide implementation of Lévy process models, most of them focus on the soft failure where system degradation crosses a critical threshold. Few of the existing studies considered multiple failure modes for a system, even fewer considered the dependence between different failure modes. Actually complex systems usually fail in a variety of modes, and one mode of failure may exacerbate another potential mode. Maintenance policy need to be investigated under these circumstances.

2.2 Accelerated degradation test planning

The utilization of degradation data for reliability estimates can date back to Nelson (1990) and Lu and Meeker (1993). Later Tang and Chang (1995) studied power supply system reliability using collected ADT data. Meeker et al. (1998) systematically addressed the advantage of degradation data and proposed several physically useful acceleration models together with statistical inference procedures. These studies, however, did not consider the planning and execution of an ADT. Although ADT is efficient in testing reliability, it is usually expensive to conduct. In addition, even within a same cost budget, an ADT still needs to be appropriately settled to achieve statistical efficiency and precision in terms of reliability estimate. The key planning variables in an ADT include test sample size, test duration, measurement frequency, number of measurements, test stress level, etc.

Based on the degradation models used, the literature on ADT planning can be classified into two categories: 1. degradation path (DP) models; 2. stochastic process (SP) models. The DP model assumes some specific function of time with random coefficients and an error term, while the SP model normally select candidates from the Lévy processes. A pioneer study of designing an ADT experiment in the DP category was given by Boulanger and Escobar (1994) where the selection of stress levels and sample size were determined. Later the study was extended by Yu and Tseng (1998) and

Tseng and Yu (1997) who incorporated a termination rule for stopping the test (test duration). Tseng and Wen (2000) considered the change of stress level during test and proposed a step stress ADT (SSADT) model. In view of the destructive measurement in some applications, recently Shi et al. (2009) obtained both statistically optimum plan and compromise plan under the framework of accelerated destructive degradation test (ADDT).

The merit of DP model is the analytical tractability in deriving reliability objective of interest and planning the ADT. However, the simple form adopted in DP model limits its ability to incorporate the time-dependent correlation in degradation measurement. Naturally, this limitation can be readily resolved by stochastic process models such as Lévy process. Tang et al. (2004) was among the first to plan the ADT using stochastic processes. The wiener process was adopted and a cost-effective SSADT plan was obtained. Later Liao and Tseng (2006) also planned an ADT using wiener process models and determined the measurement frequency, sample size and number of measurements. Instead of SSADT, Peng and Tseng (2010) proposed a test plan in which the stress level progressively increases. Tseng et al. (2009) obtained a similar plan with Liao and Tseng (2006) using gamma process. Recently Tsai et al. (2012) incorporate the random effect into gamma process and derived an optimal cost-effective test plan.

It is noted that only two classes of Lévy process are used in ADT planning, i.e. wiener process and gamma process. Obviously they are unable to deal with all types of product degradation. In fact, the GaAs laser data cannot be fit well by either of the two processes (Wang and Xu, 2010). Complementary models are necessary to handle these circumstances in ADT. Moreover, the objectives in most of the existing studies focus on the life quantile or D-optimality. To meet the primary goal of predicting the fraction failings in the field in most ADTs, other objectives need to be explored for management.

2.3 Joint maintenance and reliability test

In the literature, most studies address maintenance and reliability test separately, as these two reliability programs are usually implemented at different stages within product life cycle. Maintenance is performed after product is put into operation in the field, while most reliability tests are done before production. However, it should be noted that apart from the burn-in effect, the inherent reliability built in the product that undergoes maintenance and reliability test is essentially the same. The reliability information obtained from a properly planned reliability test is valuable and timely for manufacturers to predict the maintenance costs in the field in the long run or within a specific period (warranty). For example, Liao (2009) designed an accelerated life test (ALT) plan to predict the costs under mandatory maintenance regulations in some industries (e.g.airline).

By assuming minimal repair at each failure and random use stress across customer population, Yang (2010) obtained an optimal compromise ALT plan to minimize the asymptotic variance of the warranty costs. On the other hand, in contrast to the reliability test which is usually conducted in a static laboratory environment, practical operation of product is greatly influenced by dynamic environment such as geographic location, customer usage, etc. This gap may introduce bias in estimation using the in-lab experiment results. Some recent studies addressed this concern and proposed some generic models (Meeker et al., 2009; Hong and Meeker, 2010, 2013). However, no test plan is determined in these studies.

The above papers used lifetime models to describe the reliability metrics, which may not be appropriate for highly reliable product. The implementation of ADT is thus motivated to predict the field performance for a range of products. Currently no study is found to study joint maintenance and ADT planning. Moreover, proposition of new ADT models that overcome the gaps between lab and field is essential.

Chapter 3

A PIECEWISE CONSTANT INTENSITY MODEL AND RELATED OPTIMAL MAINTENANCE PLANNING

3.1 Introduction

As reviewed in Chapter 2, there has been considerable interests in the study of repairable system reliability under regular maintenance. The time epochs of failure events during the system lifetime give rise to the failure process which is the subject of interest in system reliability analysis (Ascher and Feingold, 1984; Rigdon and Basu, 2000). Here the time epochs refer to the points of time at which failures are observed.

Among the failure processes proposed in literature, the renewal process and Poisson process are the two most commonly used models. When a renewal process is adopted, it is assumed that the rectification performed when system fails always results in an *as good as new* system and thus the system failure process is repeated identically and periodically. However, a renewal process is unable to model the reliability growth or reliability decay often observed in repairable systems. For the Poisson process, a widely studied model is the power law process (PLP) and its variations which is a particular form of the non-homogeneous Poisson process (NHPP) (Gaudoin et al., 2003; Pulcini, 2001). The rectification under a PLP is assumed to be minimal and the system is *as bad as old* after maintenance. In addition, compromise models between renewal process and NHPP are addressed, such as modulated PLP (Muralidharan, 2002), modulated gamma process (Berman, 1981), trend renewal process (Lindqvist et al., 2003; Yang et al., 2012), etc. Maintenance planning under these processes are also covered Gilardoni and Colosimo (2007); Fuqing and Kumar (2012). While these models are typically assumed for repairable systems analysis, they are not appropriate in many applications as some substantial rectifications at system failures or other time epochs cannot be taken into consideration explicitly. As pointed out in Meeker and Escobar (1998)(p.394), “the term ’system repair’ should describe a general event of interest”, and “may be an adjustment”. In Thompson (1988)(p.54), the author argued that “some provision needs to be present for altering the process of failures when modifications or corrective actions are applied to the system”. With respect to the PLP model, Sen Sen (1998) criticised that the continuity assumption “fixed in advance all times, fails to portray the effect of fixes and design changes applied to the system”.

In practice, rectifications and other maintenance actions may trigger level shifts in the failure intensity function. For example, the test-analyse-and-fix cycle during product development often leads to reliability growth (Fries and Sen, 1996; Sen, 1998). Although the root cause for the change point of level is usually unknown, it is typically adequate to assume a constant value of failure intensity over a period of time between each changes. This is because time intervals between consecutive rectifications/adjustments are relatively short compared to system lifetime. Barring major change in the underlying operation conditions, the failure intensity is not expected to exhibit significant changes. To characterize the system failure process in these circumstances, Rigdon and Basu (2000) first introduced a piecewise exponential model (PEXP) for reliability and maintenance studies. The model assumes an independent exponential inter-failure time with parameter $\lambda_j = \frac{\mu}{\delta} j^{1-\delta}$, where λ_j is the failure intensity, $\mu, \delta > 0$ are parameters. Later Hamada et al. (2008) incorporated the Bayesian methods into the PEXP model. Recently, Arab et al. Arab et al. (2012) implemented PEXP in multiple repairable systems with two Bayesian approaches, parametric empirical Bayes approach and hierarchical Bayes approach.

The PEXP model can only describe a monotonically decreasing or increasing trend in system failure intensity. As a result, the PEXP cannot model the system behaviour where the failure intensity does not exhibit a monotonic trend. For example, some system failure intensity process often displays the well-known bathtub curve (Guida and Pulcini, 2009; Mun and Bae, 2011), comprising three successive distinct periods; namely, a decreasing failure period, following a random failure period of time with a constant failure intensity and an increasing failure period. The PEXP can only model system behaviour within one of the three periods in a bathtub curve. In view of this deficiency, we propose a general piecewise constant intensity (PCI) model in this paper, which flexibly models the level shift of the failure intensity after each rectification. This also permits a more objective study of the related optimal maintenance plan as maintenance actions could be taken to delay the on-set of increasing failure intensity even as the failure intensity function is decreasing. On the other hand, if system deterioration is indeed evident from the failure data, a cost-effective maintenance plan can also be derived so as to minimize system whole life cycle costs.

In the following sections, we first introduce the model equation and describe its parameters. We then introduce the statistical inference procedures for the PCI model in Section 3.3. This is followed by a study of the maintenance planning problem for systems with PCI failure intensity in Section 3.4. Two examples are used to illustrate the proposed model in section 3.5. Finally, Section 3.6 concludes the chapter.

3.2 Model formulation

The piecewise constant intensity (PCI) model characterizes the system's failure process by depicting the lifetime between failures. Specifically, if X_j is system lifetime between the $(j-1)$-th and j-th failure, the PCI states that X_js are independent exponential random variables with parameter

$$\lambda_j = \frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}, j = 1, \cdots, N. \tag{3.1}$$

where $\alpha_1, \alpha_2, \delta_1, \delta_2 > 0$ are unknown parameters.

Remark: From the definition of PCI in (3.1), several special cases follow, and their possible realizations are depicted in Figure **??**.

(1) if $\alpha_1 = 0$, or $\alpha_2 = 0$, then the PCI reduces to the PEXP as in Arab et al. (2012);

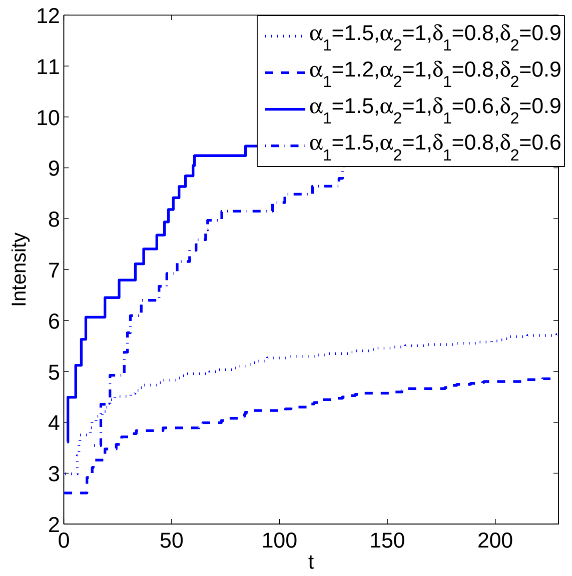
(2) if $\delta_1 = 1, \delta_2 = 1$, then the PCI is the HPP;

(3) if $\delta_1, \delta_2 > 1$, the PCI describes the reliability improvement;

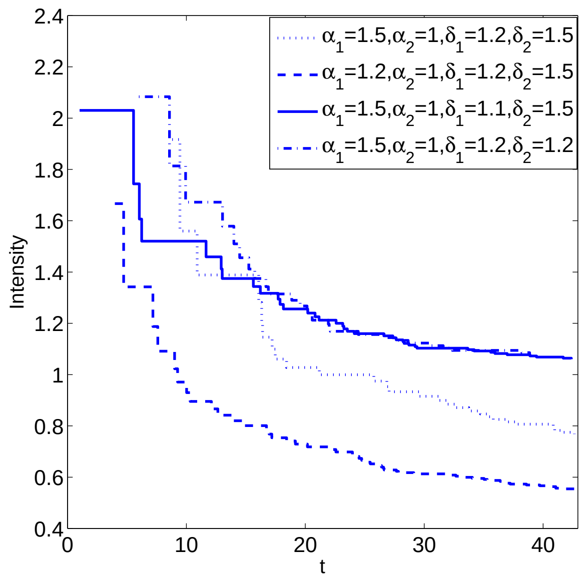
(4) if $0 < \delta_1, \delta_2 < 1$, the PCI describes the reliability deterioration;

(5) if $\delta_1 > 1, 0 < \delta_2 < 1$, or $0 < \delta_1 < 1, \delta_2 > 1$, the PCI can model a system with a more complicated failure pattern, e.g. the bathtub type failure intensity.

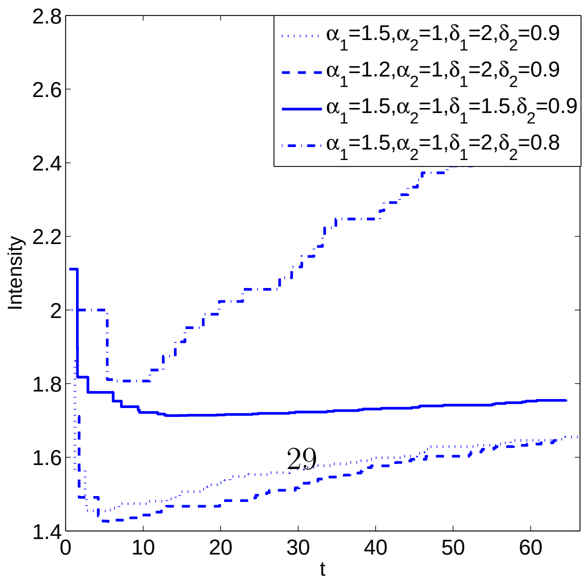
The expression in (3.1) is suitable for single system or several identical systems. More generally, however, system individual differs, either from the manufacturing stage of the systems (e.g., the variation of raw materials), or their non-identical working conditions. To extend the model for a collection of heterogeneous systems, one common approach in repairable system studies is to assume different independent parameters for each system. While this approach is intuitive, it will incur a large parameter space when the number of systems under observation increases. One solution of this issue is to incorporate the system-to-system random effect (Cook and Lawless, 2007). Specifically, we denote $z_i, i = 1, \cdots$, the unobservable random effect on system i, where z_is are taken to be i.i.d. with some distribution function $G(z)$. For system i, given z_i, the time between failures $(j-1)$ and j is exponentially distributed with parameter $\lambda'_j = z_i \lambda_j$. Note that this approach


α1=1.5,α2=1,δ1=0.8,δ2=0.9
α1=1.2,α2=1,δ1=0.8,δ2=0.9
α1=1.5,α2=1,δ1=0.6,δ2=0.9
α1=1.5,α2=1,δ1=0.8,δ2=0.6
Intensity
t

(a)


α1=1.5,α2=1,δ1=1.2,δ2=1.5
α1=1.2,α2=1,δ1=1.2,δ2=1.5
α1=1.5,α2=1,δ1=1.1,δ2=1.5
α1=1.5,α2=1,δ1=1.2,δ2=1.2
Intensity
t

(b)


α1=1.5,α2=1,δ1=2,δ2=0.9
α1=1.2,α2=1,δ1=2,δ2=0.9
α1=1.5,α2=1,δ1=1.5,δ2=0.9
α1=1.5,α2=1,δ1=2,δ2=0.8
Intensity
t

(c)

is similar to the parametric empirical Bayes model in (Rigdon and Basu, 2000).

3.3 Statistical inference

3.3.1 Identical systems

In the following, we give the likelihood function from which the parameters in the PCI model can be estimated from observed data. Suppose m detected systems operate. Denote $0 < o_{i1} < \cdots < o_{i,n_i}, i \geq 1$ the n_i-th failure time of the i-th system, and $o_{i,C}$ the time at which the i-th system is truncated. The likelihood function relative to the data is given by

$$\begin{aligned}\mathcal{L}(\Theta|\mathcal{D}) \; &= \prod_{i=1}^{m} \Bigg\{ \exp\left[-\left(\tfrac{\alpha_1}{\delta_1} n_i^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} n_i^{1-\delta_2} \right) (o_{i,C} - o_{i,n_i}) \right] \\ & \prod_{j=1}^{n_i} \left(\tfrac{\alpha_1}{\delta_1} j^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} j^{1-\delta_2} \right) \exp\left[-\left(\tfrac{\alpha_1}{\delta_1} j^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} j^{1-\delta_2} \right) (o_{i,j} - o_{i,j-1}) \right] \Bigg\},\end{aligned} \tag{3.2}$$

and the log-likelihood function is

$$\begin{aligned}l(\Theta|\mathcal{D}) \; &= \sum_{i=1}^{m} -\left(\tfrac{\alpha_1}{\delta_1} n_i^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} n_i^{1-\delta_2} \right) (o_{i,C} - o_{i,n_i}) \\ & + \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left[\log\left(\tfrac{\alpha_1}{\delta_1} j^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} j^{1-\delta_2} \right) - \left(\tfrac{\alpha_1}{\delta_1} j^{1-\delta_1} + \tfrac{\alpha_2}{\delta_2} j^{1-\delta_2} \right) (o_{i,j} - o_{i,j-1}) \right].\end{aligned} \tag{3.3}$$

For the failure truncated case, suppose a total of N_i failures are observed for system i, denote $0 < o_{i1} < \cdots < o_{i,N_i}$, all the failure epochs for the i-th system, we have

$$\mathcal{L}(\mathbf{\Theta}|\mathcal{D}) = \prod_{i=1}^{m}\prod_{j=1}^{N_i}\left(\frac{\alpha_1}{\delta_1}j^{1-\delta_1} + \frac{\alpha_2}{\delta_2}j^{1-\delta_2}\right)\exp\left[-\left(\frac{\alpha_1}{\delta_1}j^{1-\delta_1} + \frac{\alpha_2}{\delta_2}j^{1-\delta_2}\right)(o_{i,j} - o_{i,j-1})\right], \tag{3.4}$$

and

$$l(\mathbf{\Theta}|\mathcal{D}) = \sum_{i=1}^{m}\sum_{j=1}^{N_i}\left[\log\left(\frac{\alpha}{\delta_1}j^{1-\delta_1} + \frac{\alpha_2}{\delta_2}j^{1-\delta_2}\right) - \left(\frac{\alpha_1}{\delta_1}j^{1-\delta_1} + \frac{\alpha_2}{\delta_2}j^{1-\delta_2}\right)(o_{i,j} - o_{i,j-1})\right]. \tag{3.5}$$

3.3.2 Non-identical systems

Suppose the observation on system i is terminated when totally n_i failures are observed, $z_i, i = 1, \cdots, m$ follows a gamma distribution with mean 1 and variance θ. So the shape and scale parameter in the gamma distribution is θ^{-1} and θ respectively. Note that when $\theta \to 0$, the z_i distribution degenerate at a single value 1, and the model reverts back to the case of identical systems. Denote $0 < o_{i1} < \cdots < o_{i,n_i}$ all the failure epochs of system i. The parameter space is now $\mathbf{\Theta} = (\alpha_1, \alpha_2, \delta_1, \delta_2, \theta)$, the marginal likelihood function of system i is

$$\begin{aligned}\mathcal{L}_i(\mathbf{\Theta}|\mathcal{D}) &= \int_0^\infty \left(\prod_{j=1}^{n_i}\lambda_j z_i \exp[-\lambda_j z_i(o_{i,j} - o_{i,j-1})]\right)\frac{z_i^{\theta^{-1}-1}}{\Gamma(\theta^{-1})\theta^{\theta^{-1}}}\exp(-z_i/\theta)\,dz_i \\ &= \frac{\prod_{j=1}^{n_i}\lambda_j\Gamma(\theta^{-1}+n_i)}{\Gamma(\theta^{-1})}\frac{(\theta^{-1} + \sum_{j=1}^{n_i}\lambda_j(o_{i,j} - o_{i,j-1}))^{-\theta^{-1}-n_i}}{\theta^{\theta^{-1}}} \end{aligned} \tag{3.6}$$

The total likelihood is $\mathcal{L}(\mathbf{\Theta}|\mathcal{D}) = \prod_{i=1}^{m} \mathcal{L}_i(\mathbf{\Theta}|\mathcal{D})$. And the total log likelihood is given by

$$\begin{aligned} l(\mathbf{\Theta}|\mathcal{D}) &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} \ln\lambda_j + \sum_{i=1}^{m}\left[\ln\Gamma(\theta^{-1}+n_i) - (\theta^{-1}+n_i)\ln\left(\theta^{-1} + \sum_{j=1}^{n_i}\lambda_j(o_{i,j}-o_{i,j-1})\right)\right] \\ & \quad -m\ln\Gamma(\theta^{-1}) - m\theta^{-1}\ln\theta \end{aligned} \tag{3.7}$$

The maximum likelihood estimates of parameters can be readily obtained using some numerical direct search methods. For the heterogeneous case, the EM algorithm can be implemented by treating the random-effect as missing data. With the gamma distribution for z_i and given the observed data $\mathcal{D}$, the overall complete data log-likelihood l_c can be expressed as

$$l_c = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\ln\lambda_j - \lambda_j z_i(o_{i,j}-o_{i,j-1})) + \sum_{i=1}^{m}\left[(n_i+\theta^{-1}-1)\ln z_i - \frac{z_i}{\theta} - \frac{\ln\theta}{\theta} - \ln\Gamma(\theta^{-1})\right] \tag{3.8}$$

In the expectation step, we need to compute $E(z_i|\mathcal{D};\mathbf{\Theta}^{(k-1)})$ and $E(\ln z_i|\mathcal{D};\mathbf{\Theta}^{(k-1)})$. Note that conditional on $\mathcal{D}$, the distribution of z_i is still gamma, with parameters $\theta^{-1}+n_i$, and $\theta^{-1}+\sum_{j=1}^{n_i}\lambda_j(o_{i,j}-o_{i,j-1})$.

Therefore, we have

$$E(z_i|\mathcal{D};\mathbf{\Theta}^{(k-1)}) = \frac{(\theta^{-1})^{(k-1)}+n_i}{(\theta^{-1})^{(k-1)}+\sum_{j=1}^{n_i}\lambda_j^{(k-1)}(o_{i,j}-o_{i,j-1})}$$

and

$$E(\ln z_i|\mathcal{D};\mathbf{\Theta}^{(k-1)}) = \psi\left((\theta^{-1})^{(k-1)} + n_i\right) - \ln\left((\theta^{-1})^{(k-1)} + \sum_{j=1}^{n_i}\lambda_j^{(k-1)}(o_{i,j} - o_{i,j-1})\right),$$

It can be easily shown that the maximization step is similar to the procedure without random effect. It is noted that there are four and six parameters in the model of identical systems and heterogeneous systems respectively. We might expect that when the sample size is small, the information about the model parameters is limited, especially when the number of parameters increases. Therefore, a simulation study is carried out to estimate the parameter and its variance under different sample sizes and failure times. It is found that when the sample size is small, the estimate deviates from its true value with a large variance. The precision of estimation is improved when the simulated sample size increases. For estimation purpose of both models, a moderate sample size can be chosen. The coverage probability is often used to assess the effect of sample size on the estimation precision. We present in the following the coverage probability of α_1 and δ_1 when m and n_i vary.

3.3.3 Confidence interval

The confidence interval of the parameters can be constructed using the bootstrap method, which can provide better approximate results than the

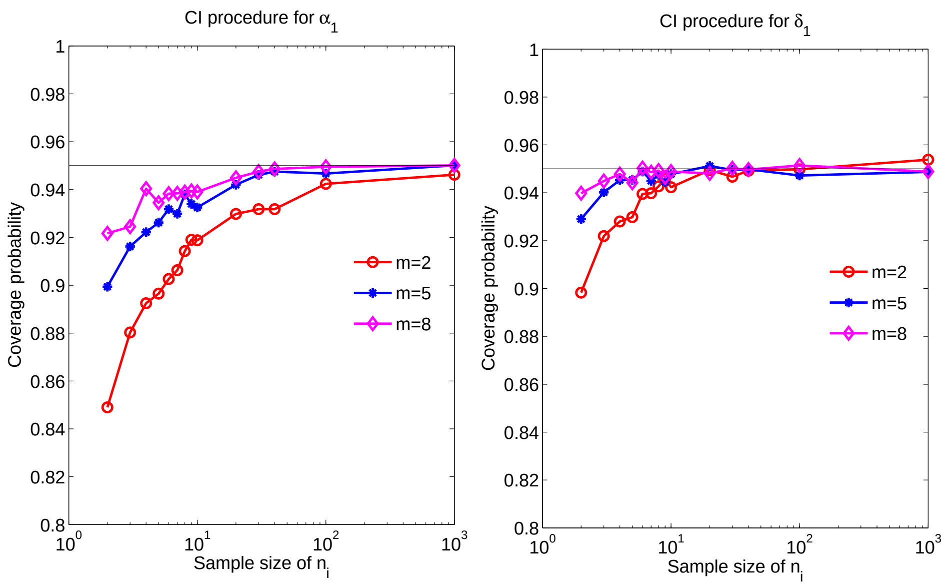


FIGURE 3.2: Coverage probability of asymptotic CI procedure with varying m and n_i.

traditional normal-approximation method, especially when the sample size is not large. We consider the case of identical systems, the heterogeneous systems case can be dealt with in a similar vein. In particular, when ML estimates of parameters are obtained from above, i.e. $\hat{\mathbf{\Theta}} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta}_1, \hat{\delta}_2)$, a parametric bootstrap percentile procedure can be developed as below (Efron and Tibshirani, 1994).

1. Generate the pseudorandom bootstrap sample $\mathbf{Q}^* = [Q_{11}, Q_{12}, \cdots, Q_{1,n_1}, Q_{21}, \cdots, Q_{m,n_m}]$, where $Q_{i,j}, i = 1, \cdots, m, j = 1, \cdots, n_i$ is the exponential time interval with rate parameter $\hat{\lambda}_{i,j}$, and $\hat{\lambda}_{i,j}$ is the plug-in estimate of $\lambda_{i,j}$, i.e. $\hat{\lambda}_{i,j} = \frac{\hat{\alpha}_1}{\hat{\delta}_1} j^{1-\hat{\delta}_1} + \frac{\hat{\alpha}_2}{\hat{\delta}_2} j^{1-\hat{\delta}_2}$;

2. Obtain the ML estimates of parameters $\hat{\Theta}^* = (\hat{\alpha}_1^*, \hat{\alpha}_2^*, \hat{\delta}_1^*, \hat{\delta}_2^*)$ using $\mathbf{Q}^*$;

3. Repeat step 1 to 2 B times to get $\hat{\mathbf{\Theta}}_1^*, \cdots, \hat{\mathbf{\Theta}}_B^*$;

4. Sort the elements in $\hat{\mathbf{\Theta}}_1^*, \cdots, \hat{\mathbf{\Theta}}_B^*$ in ascending order and construct the confidence interval using the percentile.

On the other hand, a nonparametric bootstrap procedure can also be used. This procedure replaces step 1 of the parametric bootstrap with

1. Generate a bootstrap sample $\mathbf{Q}^* = [Q_{11}, Q_{12}, \ldots, Q_{1,n_1}, Q_{21}, \ldots, Q_{m,n_m}]$, where $Q_{i,j}, i = 1, \cdots, m, j = 1, \ldots, n_i$ is obtained by randomly drawing, with replacement, from $\mathbf{o}_{.j}$, and $\mathbf{o}_{.j} = [o_{1j}, o_{2j}, \ldots, o_{mj}]$ is the vector consisting of intervals of j of all systems.

In the above procedures, it is assumed that the number of failures for each system $n_i, i = 1, \cdots, m$ and the number of systems m are non-random. To incorporate the randomness of n_i and m, we can adopt a nested bootstrap by first non-parametrically resampling with replacement from $n_1, \cdots, n_m$ and $1, \cdots, m$, then implement step 1 to 2.

3.3.4 Goodness-of-fit and model selection

The graphical check of goodness-of-fit can be done by assessing the fitness of the expected mean cumulative function (EMCF) estimated from the PCI model, to the mean cumulative function (MCF) of observed data. More specifically, the nonparametric estimate of MCF is evaluated according to Nelson (1988)(p.397). The EMCF is plotted by increasing the number of failures by 1 when the elapsed time since last failure equals the expected inter-failure time ($E(X_j)$) under the PCI model. Alternatively, we may fit the intensity plot of PCI directly to its nonparameteric estimate from the data (Rigdon and Basu, 2000) (p.99). In addition, as indicated in Sen (1998), the residual plot based on the exponential scores can be drawn. Specifically, the residuals are $e_i = \lambda_i^{-1}(o_i - o_{i-1}), i = 1, ..., n$, where o_i is the time of the i-th failure and n is the sample size.

For model selection, the most straightforward way is to use the Akaike Information Criterion (AIC) defined as $AIC = 2k - 2l(\hat{\mathbf{\Theta}})$, where $k = |\mathbf{\Theta}|$ is the number of parameters. Based on this criterion, the model with the minimum AIC value is selected. On the other hand, the root mean squared prediction residual (RMSPR) can also be used. The residual is estimated by leave-one-out cross-validation. More specifically, the system $i, i = 1, \cdots, m$ is routinely set aside, and the parameter $\mathbf{\Theta}_{-i}$ is estimated using the remaining $m-1$ systems. Then the prediction error of the failure time of system i is computed as

$$e_{ij} = o_{ij} - [o_{ij}|o_{ik}; k < j, \hat{\Theta}_{-i}], j = 1, 2, \cdots, n_i, \tag{3.9}$$

where $[o_{ij}|o_{ik}; k < j, \hat{\Theta}_{-i}]$ is the predicted inter-failure time of system i between failure $j-1$ and j conditional on all observations before o_{ij} as well as on $\hat{\Theta}_{-i}$. Then we can compute $RMSPR = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij}^2 / \sum_{i=1}^{m} n_i}$.

3.4 Maintenance planning

For repairable systems, various maintenance policies have been proposed to deal with deterioration as well as failures. Among them, system replacement/overhaul that restores the system to *as-good-as-new* status is the most common in literature and practice (Barlow et al., 1996). However, the replacement is usually expensive, especially for large systems. Therefore, based on the system failure process observed, planning for cost-effective system replacement is essential. In this section, replacement decisions are made according to the optimal control limit policies based on either events or system age when the PCI model is used to describe the system failure process.

3.4.1 Event Based Policy

A policy is called an “event based policy” if system replacement is performed when the number of events exceeds some control N since the beginning of system operation or the last replacement. The events include but are not limited to system failures, associated fixes and design changes, etc. Optimal replacement policy is determined so that the average cost in the long run is minimized, which, according to the renewal reward theorem, is equal to the average cost within a replacement cycle. Denote W the duration of each replacement cycle, $C(N)$ the average cost during W. As in the PCI model, X_j denotes the exponentially distributed random time between event $j-1$ and j. Then

$$C(N) = \frac{(N-1)c_r + c_p}{E(W)}, \tag{3.10}$$

c_r is the cost of failures and associated fixes/design changes, c_p is the system replacement cost. It is assumed that $c_r < c_p$, which usually represents the practical situation. Note that the maintenance time is usually negligible compared to the operating time, we have $W = \sum_{j=1}^{N} X_j$.

The optimization procedure is carried out through marginal analysis, i.e. incremental cost, and the optimum N^* is defined as $N^* = \min\{n|C(n+1) > C(n)\}$. Specifically,

$$\begin{aligned} C(N+1) - C(N) &= \frac{Nc_r + c_p}{\sum_{j=1}^{N+1} \frac{1}{\frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}}} - \frac{(N-1)c_r + c_p}{\sum_{j=1}^{N} \frac{1}{\frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}}} \\ &= \frac{c_r \sum_{j=1}^{N} \frac{1}{\frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}} - ((N-1)c_r + c_p) \frac{1}{\frac{\alpha_1}{\delta_1} (N+1)^{1-\delta_1} + \frac{\alpha_2}{\delta_2} (N+1)^{1-\delta_2}}}{\sum_{j=1}^{N+1} \frac{1}{\frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}} \sum_{j=1}^{N} \frac{1}{\frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}}}, \end{aligned} \tag{3.11}$$

Next, we investigate the range of parameters of which the optimum replacement exists.

Proposition 3.1. *If* $\delta_1, \delta_2 \geq 1,\ N^* = \infty$*; otherwise,*

$$\begin{aligned} N^* &= \min\Big\{N \big| \tfrac{c_r}{\frac{\alpha_1}{\delta_1}+\frac{\alpha_2}{\delta_2}} - \tfrac{c_p}{\frac{\alpha_1}{\delta_1}2^{1-\delta_1}+\frac{\alpha_2}{\delta_2}2^{1-\delta_2}} + \textstyle\sum_{j=1}^{N-1}(jc_r+c_p)\Big(\tfrac{1}{\frac{\alpha_1}{\delta_1}(j+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(j+1)^{1-\delta_2}} \\ &\quad - \tfrac{1}{\frac{\alpha_1}{\delta_1}(j+2)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(j+2)^{1-\delta_2}}\Big) > 0\Big\}. \end{aligned}$$

Proof: Denote $P(N)$ the numerator of $C(N+1)-C(N)$. When $\delta_1, \delta_2 > 1$, we have

$$\begin{aligned} P(N) &= c_r \textstyle\sum_{j=1}^{N} \tfrac{1}{\frac{\alpha_1}{\delta_1}j^{1-\delta_1}+\frac{\alpha_2}{\delta_2}j^{1-\delta_2}} - ((N-1)c_r+c_p)\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} \\ &\leq c_r \textstyle\sum_{j=1}^{N} \tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} - ((N-1)c_r+c_p)\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} \\ &= (Nc_r - ((N-1)c_r+c_p))\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} \\ &\leq 0 \end{aligned} \tag{3.12}$$

So $N^* = \infty$.

Otherwise, when $\delta_1 \geq 1$ or $\delta_2 \geq 2$, it is obtained that

$$\begin{aligned} P(N+1)-P(N) &= c_r \tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} + ((N-1)c_r+c_p)\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} \\ &\quad -(Nc_r+c_p)\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+2)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+2)^{1-\delta_2}} \\ &= (Nc_r+c_p)\left(\tfrac{1}{\frac{\alpha_1}{\delta_1}(N+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+1)^{1-\delta_2}} - \tfrac{1}{\frac{\alpha_1}{\delta_1}(N+2)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(N+2)^{1-\delta_2}}\right) \end{aligned} \tag{3.13}$$

When $N = 1,\ P(1) = \frac{c_r}{\frac{\alpha_1}{\delta_1}+\frac{\alpha_2}{\delta_2}} - \frac{c_p}{\frac{\alpha_1}{\delta_1}2^{1-\delta_1}+\frac{\alpha_2}{\delta_2}2^{1-\delta_2}}$;

When $N > 1$, $P(N) = P(1) + \sum_{j=1}^{N-1}(P(j+1) - P(j))$,

Therefore,

$$
\begin{aligned}
N^* &= \min\{N|C(N+1) - C(N) > 0\} \\
&= \min\{N|P(N) > 0\} \\
&= \min\Big\{N|\tfrac{c_r}{\frac{\alpha_1}{\delta_1}+\frac{\alpha_2}{\delta_2}} - \tfrac{c_p}{\frac{\alpha_1}{\delta_1}2^{1-\delta_1}+\frac{\alpha_2}{\delta_2}2^{1-\delta_2}} + \textstyle\sum_{j=1}^{N-1}(jc_r + c_p)\Big(\tfrac{1}{\frac{\alpha_1}{\delta_1}(j+1)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(j+1)^{1-\delta_2}} \\
&- \tfrac{1}{\frac{\alpha_1}{\delta_1}(j+2)^{1-\delta_1}+\frac{\alpha_2}{\delta_2}(j+2)^{1-\delta_2}}\Big) > 0\Big\}.
\end{aligned}
\tag{3.14}
$$

From the above statements, the argument of Proposition 3.1 is readily obtained. □

3.4.2 Age Based Policy

In the age based policy, system replacement is performed when the elapsed time since the last replacement exceeds some threshold T. Consequently, the replacement cycle is $W = T$. The average maintenance cost within replacement cycle, $C(T)$ is

$$
C(T) = \frac{\sum_{i=1}^{\infty}((i-1)c_r + c_p)P(\sum_{j=1}^{i} X_j < T \leq \sum_{j=1}^{i+1} X_j)}{T}. \tag{3.15}
$$

To minimize $C(T)$, the following lemma is introduced here to facilitate the subsequent analysis.

Lemma 3.2. Ross (1996).*Let* $(X_j), j = 0, 1, \cdots, n, n \geq 2$ *be independent exponential random variables with pairwise distinct respective parameters* λ_j. *Then the density of their sum is*

$$f_{X_1+X_2+\cdots+X_n}(x) = \left[\prod_{j=1}^{n} \lambda_j\right] \sum_{k=1}^{n} \frac{\exp(-\lambda_k x)}{\prod_{\substack{l \neq k \\ l=1}}^{n} (\lambda_l - \lambda_k)}, x > 0 \tag{3.16}$$

The proof can be found in Ross (1996) and thus omitted here. In the PCI model, $\lambda_j = \frac{\alpha_1}{\delta_1} j^{1-\delta_1} + \frac{\alpha_2}{\delta_2} j^{1-\delta_2}$.

Note that

$$C(T) = \frac{\sum_{i=1}^{\infty} ((i-1)c_r + c_p) \left(P(\sum_{j=1}^{i+1} X_j \geq T) - P(\sum_{j=1}^{i} X_j \geq T) \right)}{T} \tag{3.17}$$

where

$$\begin{aligned} P(\textstyle\sum_{j=1}^{i} X_j \geq T) &= \int_T^{\infty} f_{X_1+X_2+\cdots+X_i}(x)\, dx = \int_T^{\infty} \left[\prod_{j=1}^{i} \lambda_j\right] \sum_{k=1}^{i} \frac{\exp(-\lambda_k x)}{\prod_{\substack{l \neq k \\ l=1}}^{i} (\lambda_l - \lambda_k)}\, dx \\ &= \left[\prod_{j=1}^{i} \lambda_j\right] \sum_{k=1}^{i} \frac{1}{\prod_{\substack{l \neq k \\ l=1}}^{i} (\lambda_l - \lambda_k)} \frac{1}{\lambda_k} \exp(-\lambda_k T) \end{aligned} \tag{3.18}$$

Letting $D_i(T) = P(\sum_{j=1}^{i+1} X_j \geq T) - P(\sum_{j=1}^{i} X_j \geq T)$ yields

$$\begin{aligned}
D_i(T) &= \left[\prod_{j=1}^{i+1}\lambda_j\right]\sum_{k=1}^{i+1}\frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i+1}(\lambda_l-\lambda_k)}\frac{1}{\lambda_k}\exp(-\lambda_k T) - \left[\prod_{j=1}^{i}\lambda_j\right]\sum_{k=1}^{i}\frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i}(\lambda_l-\lambda_k)}\frac{1}{\lambda_k}\exp(-\lambda_k T)\\
&= \sum_{k=1}^{i}\left[\prod_{j=1}^{i}\lambda_j\right]\frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i}(\lambda_l-\lambda_k)}\frac{1}{\lambda_k}\exp(-\lambda_k T)(\lambda_{i+1}\frac{1}{\lambda_{i+1}-\lambda_k}-1)\\
&\quad + \left[\prod_{j=1}^{i+1}\lambda_j\right]\frac{1}{\prod\limits_{l=1}^{i}(\lambda_l-\lambda_{i+1})}\frac{1}{\lambda_{i+1}}\exp(-\lambda_{i+1}T)\\
&= \sum_{k=1}^{i}\left[\prod_{j=1}^{i}\lambda_j\right]\frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i}(\lambda_l-\lambda_k)}\frac{1}{\lambda_k}\exp(-\lambda_k T)(\frac{\lambda_k}{\lambda_{i+1}-\lambda_k})\\
&\quad + \left[\prod_{j=1}^{i+1}\lambda_j\right]\frac{1}{\prod\limits_{l=1}^{i}(\lambda_l-\lambda_{i+1})}\frac{1}{\lambda_{i+1}}\exp(-\lambda_{i+1}T)
\end{aligned} \tag{3.19}$$

The optimum is sought by taking the first order derivative of $C(T)$ with respect to T, that is

$$\frac{\partial C(T)}{\partial T} = \frac{\sum_{i=1}^{\infty}((i-1)c_r + c_p)\,(D_i'(T)T - D_i(T))}{T^2} \tag{3.20}$$

where $D_i'(T) = dD_i(T)/dT$, and

$$\begin{aligned}
&D_i'(T)T - D_i(T) = \\
&-\Big\{ \textstyle\sum_{k=1}^{i} \Big[\prod_{j=1}^{i} \lambda_j\Big] \frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i} (\lambda_l - \lambda_k)} \exp(-\lambda_k T)(\frac{\lambda_k}{\lambda_{i+1}-\lambda_k}) + \Big[\prod_{j=1}^{i+1} \lambda_j\Big] \frac{1}{\prod\limits_{l=1}^{i} (\lambda_l - \lambda_{i+1})} \exp(-\lambda_{i+1} T) \Big\} T \\
&-\Big\{ \textstyle\sum_{k=1}^{i} \Big[\prod_{j=1}^{i} \lambda_j\Big] \frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i} (\lambda_l - \lambda_k)} \frac{1}{\lambda_k} \exp(-\lambda_k T)(\frac{\lambda_k}{\lambda_{i+1}-\lambda_k}) + \Big[\prod_{j=1}^{i+1} \lambda_j\Big] \frac{1}{\prod\limits_{l=1}^{i} (\lambda_l - \lambda_{i+1})} \frac{1}{\lambda_{i+1}} \exp(-\lambda_{i+1} T) \Big\} \\
&= - \textstyle\sum_{k=1}^{i} \Big[\prod_{j=1}^{i} \lambda_j\Big] \frac{1}{\prod\limits_{\substack{l\neq k\\ l=1}}^{i+1} (\lambda_l - \lambda_k)} \exp(-\lambda_k T)(\lambda_k T + 1) \\
&- \Big[\prod_{j=1}^{i} \lambda_j\Big] \frac{1}{\prod\limits_{l=1}^{i} (\lambda_l - \lambda_{i+1})} \exp(-\lambda_{i+1} T) \Big(T + \frac{1}{\lambda_{i+1}}\Big)
\end{aligned} \tag{3.21}$$

It is seen that when $T \to 0$, (3.21) approaches to some negative constant, so $\frac{\partial C(T)}{\partial T} \to -\infty$; when $T \to \infty$, $\frac{\partial C(T)}{\partial T} \to 0$. Therefore, there exist two possible scenarios: in scenario 1, there is at least a local optimum T^*, which satisfies $\frac{\partial C(T)}{\partial T}|_{T=T^*} = 0$; in scenario 2, there is no finite T^*, i.e. $T^* \to \infty$.

In Section 3.3, parameters are estimated using data consisting of failure times. However, in some circumstances the maintenance epochs may also be recorded in the dataset (Blischke and Murthy, 2003). Suppose only replacements are considered. Then for the system k within the fleet, denote $0 < t_{k1} < \cdots < t_{km}, m \geq 1$ the replacement epochs, and $t_{k,i-1} < o_{ki1} < \cdots < o_{ki,n_{ki}} < t_{k,i}, 0 \leq i \leq m$ the failure times between replacement $(i-1)$

and i. The likelihood function for the k-th system is

$$\mathcal{L}_k((\Theta)|\mathcal{D}) = \prod_{i=1}^{m} \left\{ \exp\left(-\lambda_{n_{ki}+1}(t_{ki} - o_{ki,n_{ki}})\right) \prod_{j=1}^{n_i} \left[\lambda_j \exp\left(-\lambda_j(o_{ki,j} - o_{ki,j-1})\right)\right] \right\} \tag{3.22}$$

and the log-likelihood function is

$$l_k(\boldsymbol{\Theta}|\mathcal{D}) = \sum_{i=1}^{m} \left(-\lambda_{n_{ki}+1}(t_{ki} - o_{i,n_{ki}})\right) + \sum_{i=1}^{m} \sum_{j=1}^{n_{ki}} \left[\log\lambda_j - \lambda_j(o_{ki,j} - o_{ki,j-1})\right] \tag{3.23}$$

The total log-likelihood function is thus $l(\boldsymbol{\Theta}|\mathcal{D}) = \sum_k l_k(\boldsymbol{\Theta}|\mathcal{D})$. The estimates of the optimal replacement plans N^* and T^*, $\hat{N}^*$ and $\hat{T}^*$ can be obtained by plugging in the estimate of the parameters combined with the decision rule in Proposition 3.1 and the results in Section 4.2. Their confidence intervals are constructed again using the bootstrap method.

3.5 Numerical example

3.5.1 The load-haul-dump machine data

In this section, the failure data of the load-haul-dump machine system data is used to demonstrate the proposed model. The data listed in Table 3.1 was originally described by Kumar and Klefsjö (1992) and reported in Hamada et al. (2008) and Arab et al. (2012). We show the adequacy of the PCI model and perform the system replacement.

TABLE 3.1: Inter-failure time data for the LHD machine.

LHD1	327,125,7,6,107,277,54,332,510,110,10,9,85,27,59,16,8,34,21 152,158,44,18
LHD3	637,40,197,36,54,53,97,63,216,118,125,25,4,101,184,167,81,46 18,32,219,405,20,248,140
LHD9	278,261,990,191,107,32,51,10,132,176,247,165,454,142,39,249 212,204,182,116,30,24,32,38,10,311,61
LHD11	353,96,49,211,82,175,79,117,26,4,5,60,39,35,258,97,59,3,37 8,245,79,49,31,259,283,150,24
LHD17	401,36,18,159,341,171,24,350,72,303,34,45,324,2,70,57,103 11,5,3,144,80,53,84,218,122
LHD20	231,20,361,260,176,16,101,293,5,119,9,80,112,10,162,90,176 360,90,15,315,32,266

We first check the goodness-of-fit of the model using the graphical approach discussed in Section 3.3.4. The residual plot of each system is presented in Figure 3.3 and no significant departure is found. Alternatively, using the MCF plot, the nonparametric estimate of MCF is compared with both the PCI model and the PEXP model in (Arab et al., 2012). The results are presented in Figure 3.4 and it is found that the three curves tally reasonably well, which indicates the adequacy of the PCI model.

Next we check the PCI model using the procedure in Section 3.3.4. The nonparametric estimate of mean cumulative function (MCF) of failures which naturally describes the system's failure process based on the data. The algorithm in Meeker and Escobar (1998) (pp.397) is used to compute the MCF estimate. The results are presented in Figure 3.4 along with the plot using the PEXP model (Arab et al., 2012). It is found that the three

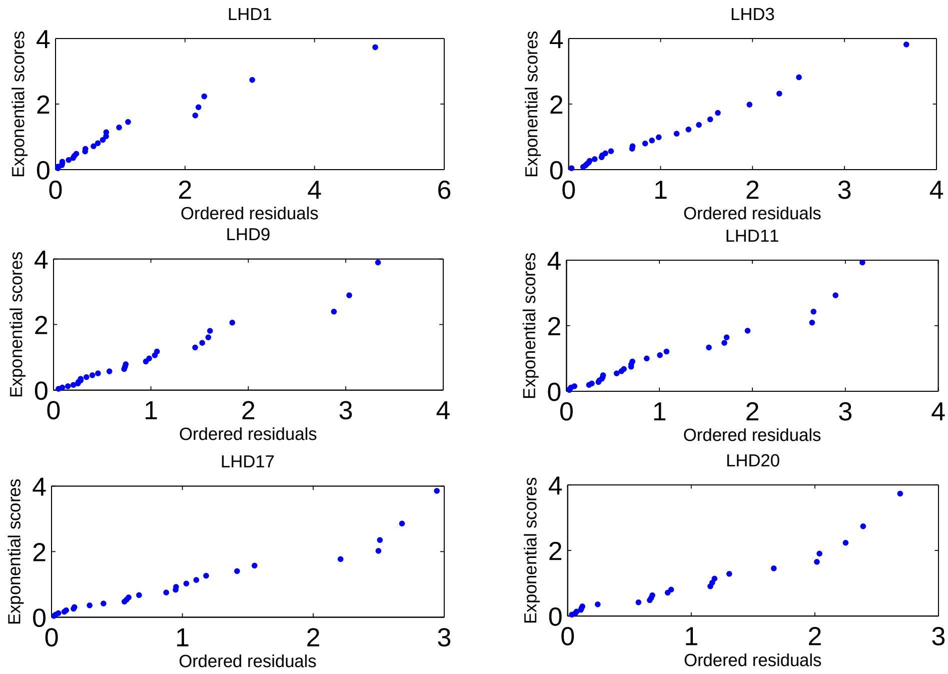


FIGURE 3.3: Residual plot for the LHD machine data.

curves tally reasonably well.

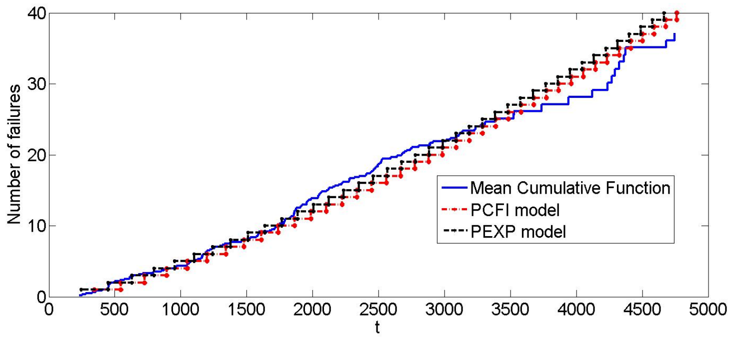


FIGURE 3.4: The nonparametric MCF, the parametric PCI model and PEXP model based on the LHD machine data.

If the systems are assumed identical, the MLE of parameters can be obtained from Section 3.1 as $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta}_1, \hat{\delta}_2) = (0.0012, 0.0017, 0.716, 0.734)$, with a log-likelihood value of -889.7. On the other hand, if the possible heterogeneity across systems is considered, the method introduced in Section 3.2 is used to estimate the parameters. The results are $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta}_1, \hat{\delta}_2, \hat{\theta}) = (0.0014, 0.0017, 0.704, 0.714, 1.1 \times 10^{-6})$, and the log-likelihood is -889.1. Note that the small value of $\hat{\theta}$ corresponds to a small variance of the random effect, which indicates the insignificance of the heterogeneity in the LHD systems. Therefore, the consideration of heterogeneity will not be carried over in the subsequent analysis.

It is noted that the LHD machines exhibit reliability deterioration since $\hat{\delta}_1, \hat{\delta}_2 < 1$, consequently it is important to plan a maintenance such as an overhaul or replacement of the whole system. To facilitate a cost effective maintenance policy, the unit cost of maintenance, c_p and c_r can be estimated and obtained in practice. Here for illustrative purpose, the value configuration is adopted as $c_p = \$10,000, c_r = \$500$. Then by Proposition 3.1, system replacement is planned at the $N^* = 48$-th failure epoch, and the long run average cost is $C(N^*) = 6.26$.

The 95% bootstrap confidence interval for $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta}_1, \hat{\delta}_2$ is constructed with the bootstrap sample size $B = 1000$. Based on this, the confidence interval for the optimal maintenance policy N^* and $C(N^*)$ is also obtained. The results are presented in Table 3.2.

Figure 3.5 shows the long run average cost $C(N)$ when the maintenance

TABLE 3.2: 95% confidence interval for the parameters and optimal maintenance decisions.

Parameter	α_1	α_2	δ_1	δ_2	N^*	$C(N^*)$
interval	(0.0008,0.0024)	(0.001,0.0032)	(0.56 0.88)	(0.57 0.89)	(30,143)	(5.5,6.58)

epoch N varies. The 95% confidence interval of $C(N)$ is also plotted. It is seen that the optimal cost is quite flat when N varies, which is in accordance with the relatively loose 95% confidence interval of N^*.

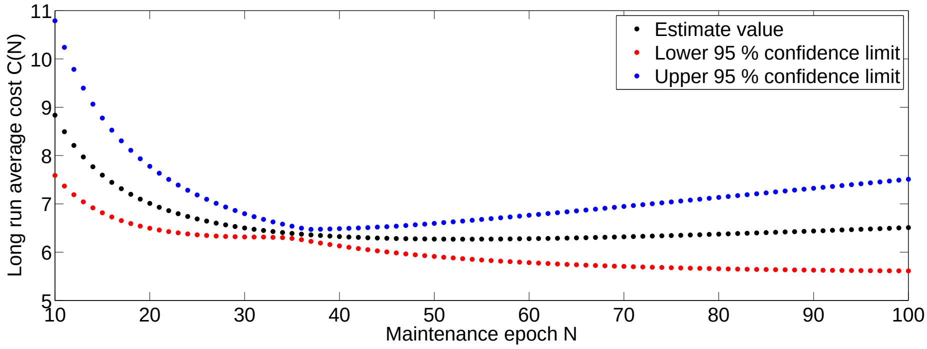


FIGURE 3.5: Long run average cost versus various maintenance epochs.

The effects of cost parameters c_r, c_p on the optimal maintenance policy are depicted in Figure 3.6 by varying c_r from 200 to 1000 and c_p from 2000 to 20,000. It is found that the values of optimal maintenance policy is largely dependent on the cost of each repair/fix/change and replacement. Therefore, it is critical to forecast and estimate the cost in practice before maintenance decisions are made.

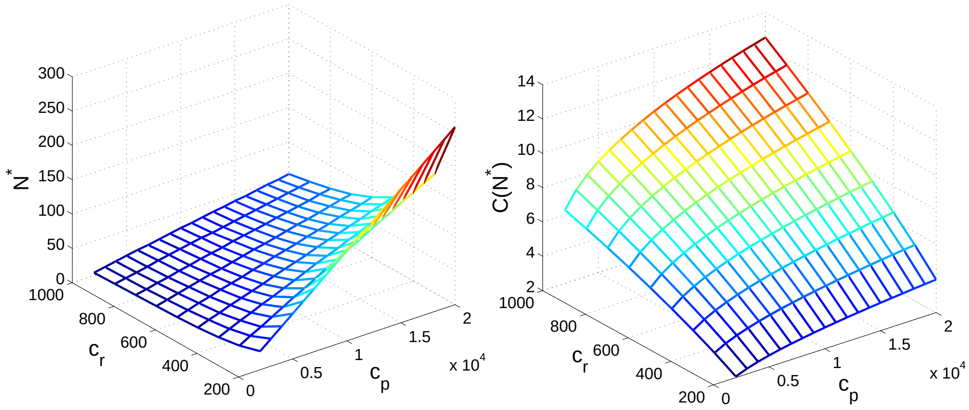


FIGURE 3.6: N^* and $C(N^*)$ versus combinations of c_r and c_p.

3.5.2 The rear dump truck data

To illustrate the superiority of PCI model over the PEXP model, we consider the data of a 180 ton rear dump truck given by Coetzee (1996). The dataset consists of 128 failure times as listed below.

TABLE 3.3: Failure time data for the real dump truck.

78	158	331	381	523	620	664	1805	1817	2068	3253
4489	4725	4961	5138	5200	5278	5711	6400	6444	6677	7999
8001	8489	9000	9086	10262	10817	11062	11082	11086	11122	11534
12031	12339	12733	13265	13508	13673	13780	14443	14501	14656	14906
14983	15004	15062	15072	15136	15206	15247	15700	15714	15972	16186
16284	16329	16425	16605	16723	16731	16797	16859	17090	17305	17484
17510	17511	17536	17621	17703	17809	17968	17984	18175	18443	18458
18667	18669	18701	18723	18822	18860	18922	18935	18945	18960	18961
18979	19013	19032	19034	19169	19184	19201	19416	19455	19525	19595
19601	19613	19643	19671	19713	19785	19801	19937	19990	20432	20433
20434	20698	21460	21543	21584	21602	21645	21706	21762	21867	21912
21914	21937	21938	21939	21951	21954	21982				

We use PCI model and PEXP model to fit this data and the estimated parameters and AIC values are presented in Table 3.4. Besides, we plot the MCF, and EMCF under both PCI and PEXP in Figure 3.7. Both the AIC value and the plot show that the PCI model fits the MCF better than the PEXP model. Moreover, the RMSPR for PCI and PEXP is 114.3 and 165.9 respectively, indicating a superior performance of PCI for predictive ability.

TABLE 3.4: Estimated parameters of the PCI and PEXP model in fitting the rear dump truck data

Model	Estimated Parameters	AIC
PCI	$\hat{\alpha}_1 = 8.86 \times 10^{-7}$, $\hat{\alpha}_2 = 0.46$, $\hat{\delta}_1 = 6.85 \times 10^{-4}$, $\hat{\delta}_2 = 2.45$	923.7
PEXP	$\hat{\mu} = 5 \times 10^{-4}$, $\hat{\delta} = 0.52$	936.2

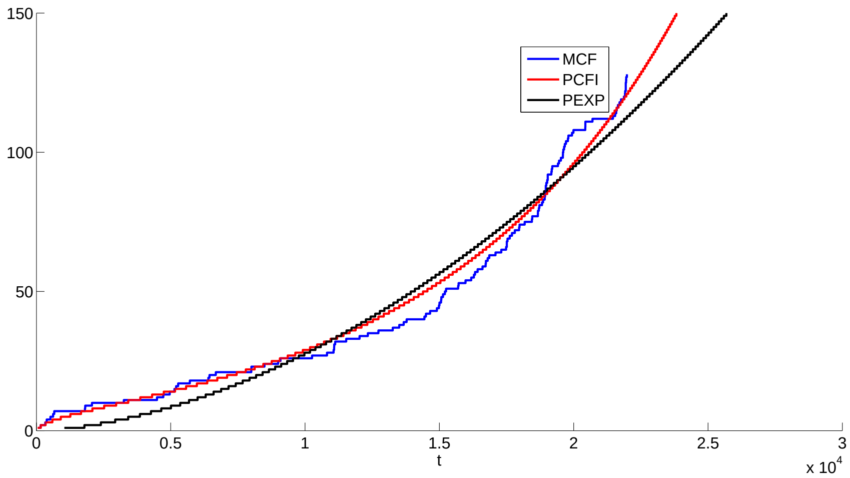


FIGURE 3.7: The nonparametric MCF, the parametric PCI model and PEXP model based on the real dump truck data.

3.6 Conclusion

A new versatile model for describing the failure process of repairable systems is proposed in this chapter. It has been shown that the PCI model is a useful alternative approach to characterize failure processes of repairable systems, especially when the failure intensity function is non-monotone, which is commonly seen in practice, such as infant mortality which incurs a significant proportion of early failures. It is also a more realistic model compared to PLP and renewal processes in the event that there are substantial rectifications/changes at failures or other time epochs, which usually appears in reliability growth programs especially in military industries. This chapter studies some properties of the PCI model and presents the statistical estimation of the model parameters. A goodness-of-fit test and a bootstrap confidence bound for the failure intensity function are also derived. The associate optimal maintenance plans are also investigated and it is found that optimal plan exists for bath-tub shape failure intensity function. The flexibility of the PCI model is illustrated using real dataset example.

Chapter 4

MAINTENANCE IN AN UNRELIABLE PRODUCTION SYSTEM WITH IMPERFECT PRODUCTION

4.1 Introduction

In the mass production environment, reliability of production systems is a critical issue. As indicated in Hopp and Spearman (2008), unscheduled breakdowns are the single largest source of production variability in many systems, generating problems of high inventory and late delivery. On the other hand, manufacturers are facing increasing pressure to improve the product quality. In the production stage, defective products may result in additional rework costs. They can also disrupt the production process and cause a delay, especially in a lean manufacturing environment. In the post-sale stage, low quality products will cause high expenditure in warranty or even the lost good-will of customers. The issues of reliability and quality in production system have posed challenges to modern industrial engineers.

One commonly used mitigation approach for system failure(breakdown) is maintenance. Among the numerous maintenance optimization methods in the literature, preventive maintenance(PM) and corrective maintenance(CM) are two mostly implemented policies(Wang, 2002). The pioneer study of Groenevelt et al. (1992) investigated CM actions when machine failures occur in lot-sizing production system. When the production system is under monitoring and its failure/maintenance history is observed, PM is facilitated based on the statistical analysis of failure time data. Zhou et al. (2007) proposed a joint age-based PM and reconfiguration decision for recoverable systems. El-Ferik (2008) adopted an age-based imperfect PM policy where the optimal PM schedule and number of production runs between two system overhauls/replacements are obtained. Liao et al. (2009) used a periodical imperfect PM which is self-improving with a learning effect.

Parallel to the research on maintenance due to production system failure, another stream of research focus on rectifications of system's defective outputs. This is sometimes called *imperfect production* in the literature (Cheng, 1991; Salameh and Jaber, 2000). To characterize the imperfect production, one approach treats it as random shocks during the production. When the shock occurs, the production system shifts to an out-of-control state in which the subsequent production will incur a proportion of defective outputs. This approach is adopted in Tsao et al. (2013), Sana (2010) for example. In these papers, maintenance is also performed to restore the system to in-control state. Both perfect maintenance (Lee and Rosenblatt, 1987; Tseng, 1996) and imperfect maintenance policies (Tseng et al., 1998; Ben-Daya, 2002; Sheu and Chen, 2004; Wang et al., 2009) were investigated. The second approach stems from the quality control concept and assumes that a random proportion of system outputs are defective all the time. This approach can be found in Wee et al. (2013) and Eroglu and Ozdemir (2007). However, this stream of research does not consider the system failure.

While the above research addresses reliability and quality issues separately, some attempts have been made to jointly consider both two issues. This is legitimate since the two issues are not isolated. In fact, as pointed out by

Cassady et al. (2000), there is a close relationship between system maintenance and product quality. In practice, production systems degrade over time, there may be malfunction of its components when the degradation is large. This malfunction is likely to cause the output products to deviate from the standard quality and eventually become defective. Following the preliminary investigation of Cassady et al. (2000), combination studies of stochastic process control (SPC) and maintenance are discussed by several researchers (Yeung et al., 2007; Wang, 2012). On the other hand, some studies view the quality issue as the consequence of system deterioration. As a result, they focus on maintenance planning to cope with system failure while the cost of imperfect production is included in decision. For example, Boone et al. (2000) determined the optimal production run where both imperfect production and machine breakdowns are taken into account. Chakraborty et al. (2009) further incorporated the inspection schedule into maintenance policy.

The purpose of this study is to propose an integrated maintenance framework that jointly considers reliability and quality issues in production system. Different from previous studies, we use a flexible quasi-renewal process(Wang and Pham, 2006) to characterize the system failure process, in view of the common multi-stage production with imperfect maintenance. This process is widely studied in reliability literature(Lam, 2009; Samatl-Pa and Taner, 2009). It is able to describe different system behaviours including system deterioration, improvement or renewal. Besides, as indicated in Yeh and Chan (1998), Chan et al. (2004) and Braun et al. (2005), it is satisfactory in modelling a variety of systems. Moreover, Yeh (1992) and Braun et al. (2005) showed that commonly used life distributions(Weibull, normal or gamma distributions) can be naturally embodied in the process, which greatly enhances its flexibility. In terms of product quality aspect, we use the second approach of modelling imperfect production. We assume that there is a random proportion of defective output in each production run, and this random proportion propagates through each production runs by a fixed parameter. In particular, for the i-th production run, the proportion of defective products is $\beta_i B$, where β_i is the parameter and B is a random variable representing the proportion of defective outputs. The proposed model is easy to implement and the statistical inference procedure is facilitated. In addition, the dependence of system reliability and product quality can be easily imposed when necessary. One approach is to assign a joint distribution between β_i and the parameter of quasi-renewal process. In addition, analytical results on the optimal maintenance decisions are available based on this modelling.

The rest of the chapter is organized as follows. In Section 2, a detailed description of the model and relevant notations are presented. Section 3 derives detailed components of the model. The analytical derivations can be found in Section 4. Section 5 provides numerical studies and Section 6 concludes the paper.

4.2 Model formulation

A production cycle is defined as the time interval between two replacement/overhaul of the whole production system. There are several production runs in each production cycle. Since production ceases at maintenance, a production run is the time interval between two PMs/CMs. In particular, in the i-th production run, $i = 1, 2, \cdots$, the PM time is scheduled at T_i which is a decision variable. Therefore, if the system survives till T_i, a PM is performed with a cost c_p; if system fails before T_i, a CM is carried out which incurs a cost c_r. It is assumed that $c_r > c_p$. After either PM or CM, if the product inventory has depleted, the system starts the next production run immediately; otherwise,the system stays idle until the inventory decreases to 0. During all production runs, the production rate is p and customer demand rate is d, where $p > d$ and p, d are both constants.

Denote Z_i the random variable that represents the system lifetime in the i-th production run, following the definition of the quasi-renewal process(Wang and Pham, 2006), it is assumed that $\alpha^{i-1} Z_i$ are i.i.d random variables, where $\alpha \geq 0$ is the parameter. Meanwhile, the system is also subject to imperfect production. As mentioned earlier, this means that a proportion $\beta_i B$ of the production outputs are defective in the i-th production run, where β_i is a parameter affected by maintenance and B is a random variable. It is assumed that all the defective units are fixed or reworked immediately at a cost c_d per unit and put back to the inventory. In addition to the maintenance and imperfect production costs, a production set-up cost k, and the inventory holding cost h per unit (including all good and reworked units) per unit time are also included. After N production runs which is a decision variable, system is replaced and the production cycle ends. Therefore, two decisions are considered in this model: maintenance time in each production run $T_i, i = 1, 2, ..., N$ and replacement schedule N. Our purpose is to minimize the long-run average cost. Clearly there is a trade-off between T_i and N. Therefore, a well planned maintenance policy needs to be investigated. In next section, we derive the related costs and formulate the model.

4.3 Cost functions and the optimization problem formulation

According to the renewal reward theorem (Ross, 1983), the long-run average cost is readily represented by the average cost within a replacement cycle, as given in the equation below.

$$AVC(\{T_i\}, N) = \frac{TC(\{T_i\}, N)}{CL(\{T_i\}, N)}$$

$$= \frac{HC(\{T_i\}, N) + SC(\{T_i\}, N) + MC(\{T_i\}, N) + DE(\{T_i\}, N)}{CL(\{T_i\}, N)}, \tag{4.1}$$

where $AVC(\{T_i\}, N)$ is the long-run average cost, $TC(\{T_i\}, N)$ is the total costs, $CL(\{T_i\}, N)$ is the cycle length, $HC(\{T_i\}, N)$ is the inventory holding cost, $SC(\{T_i\}, N)$ is the production set-up cost, $MC(\{T_i\}, N)$ is the maintenance costs, $DE(\{T_i\}, N)$ is the rework costs of defective products. The expressions for various costs and the cycle length are derived in the following.

(1) *Expected holding cost*

The holding cost is simply given by

$$HC(\{T_i\}, N) = \sum_{i=1}^{N} HC_i,$$

where HC_i is the holding cost in the i-th production run. Given T_i, the production run may end up with regular maintenance if system is still working at that time, or with a system breakdown before T_i. Thus for an production lot sizing model with constant production and demand rate,

HC_i is obtained as,

$$HC_i = \frac{hp(p-d)}{2d}\left(\int_0^{T_i} t^2 f_{Z_i}(t)\,dt + \int_{T_i}^{\infty} T_i^2 f_{Z_i}(t)\,dt\right), \tag{4.2}$$

where $f_{Z_i}(t)$ is the PDF of the time to break down in the i-th production run and is given by

$$f_{Z_i}(t) = \frac{dF_{Z_i}(t)}{dt} = \frac{dF_{Z_i}(\alpha^{i-1}t)}{dt} = \frac{\alpha^{i-1}dF_{Z_i}(\alpha^{i-1}t)}{d\alpha^{i-1}t} = \alpha^{i-1}f_{Z_i}(\alpha^{i-1}t), \tag{4.3}$$

Similarly, the CDF of Z_i, $F_{Z_i}(t)$ can be obtained as follows,

$$F_{Z_i}(t) = P(Z_i < t) = P(\alpha^{i-1}Z_i < \alpha^{i-1}t) = P(Z_1 < \alpha^{i-1}t) = F_{Z_1}(\alpha^{i-1}t), \tag{4.4}$$

(2) *Expected cost of imperfect production*

The cost of imperfect production is given by

$$DE(\{T_i\}, N) = \sum_{i=1}^{N} DE_i$$

$$= c_d \sum_{i=1}^{N} E\left[\beta_i B \cdot \text{number of items produced in the } i\text{-th production run}\right], \tag{4.5}$$

where $E[B]$ is the expectation of the random variable B.

where DE_i is imperfect production cost in the i-th production run. It is obvious that DE_i is proportional to the production time of the i-th production run. We then obtain

$$DE\left(\{T_i\}, N\right) = \sum_{i=1}^{N} DE_i = pc_d E\left[B\right] \sum_{i=1}^{N} \left[\beta_i \left(\int_0^{T_i} t f_{Z_i}(t)\, dt + \int_{T_i}^{\infty} T_i f_{Z_i}(t)\, dt\right)\right]. \tag{4.6}$$

(3) *Expected cost of maintenance*

Maintenance cost is associated with each production run conditional on the operating time of system in that production run. Besides, a replacement is incurred after N production runs. Thus,

$$MC\left(\{T_i\}, N\right) = \sum_{i=1}^{N} \left[c_r \int_0^{T_i} f_{Z_i}(t)\, dt + c_p \int_{T_i}^{\infty} f_{Z_i}(t)\, dt\right] + r. \tag{4.7}$$

(4) *Expected set-up cost for each cycle*

Each time a production run starts, it incurs a setup cost k, so the total expected setup cost is

$$SC\left(\{T_i\}, N\right) = Nk. \tag{4.8}$$

(5) *Expected cycle length*

For each production run, since maintenance time is negligible, a production cycle ends when inventory depletes to 0. For a production run with production time t, the production run length is pt/d. Thus,

$$CL\left(\{T_i\}, N\right) = \sum_{i=1}^{N} CL_i = \sum_{i=1}^{N} \left[\int_0^{T_i} \frac{pt}{d} f_{Z_i}(t)\, dt + \int_{T_i}^{\infty} \frac{pT_i}{d} f_{Z_i}(t)\, dt \right]. \tag{4.9}$$

The above derivations can be substituted into (4.1) to obtain the long-run average cost.

4.4 Model analysis and optimality conditions

In this section, we study some properties and conditions of the optimal maintenance policy under general system lifetime distributions. To begin with, we seek the existence of T_i and N, which is given in the following lemma.

Lemma 4.1. *$\forall N \in \mathbf{N}$, where $\mathbf{N}$ is the set of natural numbers, there exists $T_i > 0, i = 1, 2, ..., N$ which at least lead to a local minimum of $AVC(\{T_i\}, N)$.*

On the other hand, given $T_i > 0, i = 1, 2, ..., N$, denoted by N^* that minimizes $AVC(\{T_i\}, N)$, *when* $\alpha > 1$*,* N^* *exists and is finite;*

when $\alpha < 1$*, if* $\beta_i, i = 1, 2, ..., N$ *is nondecreasing,* N^* *exists and is finite.*

The proof is given in Appendix A.

Although for each production run, it is ideal to assign a specific scheduled maintenance time T_i, in practice, however, it may not be efficient and appropriate due to the possible difficulty of implementation and insignificance in cost reduction. Therefore, to simplify the $(\{T_i\}, N)$ maintenance policy, one way is to assign an identical maintenance time T for every production run. Then the decision variables become (T, N) and it is called the (T, N) maintenance policy in this paper. As a special case of the $(\{T_i\}, N)$ maintenance policy, the (T, N) maintenance policy inherits the properties of $(\{T_i\}, N)$ policy. Besides, additional results are obtained as follows.

From the above derivations (e.g.(4.6)), it is seen that N is implicitly expressed in the cost function, i.e. in the summation function. Hence it is not easy, if not impossible, to determine the global optimum of N. In this study we focus on the possible local optimum of N and define it as

$$N^* = min\{N : AVC(T, N + 1) > AVC(T, N), N \in \mathbf{N}\}. \qquad (4.10)$$

It is shown later in the numerical example that under an appreciable range of N, $AVC(T, N)$ is convex to N,in which case N^* defined in (4.10) becomes a global optimum. Some important properties of the model are listed in the proposition below. Note that these properties are considered in the general $(\{T_i\}, N)$ policy.

Proposition 4.1. *For any given $N \in \mathbf{N}$, if $p > d$ and $F_{Z_1}(t)$ is IFR,*

when $\alpha > 1$, if $\beta_i, i = 1, 2, ..., N$ is nondecreasing,then the sequence of optimal maintenance times $T_i^*, i = 1, 2, ..., N$ is in nonincreasing order,i.e.$T_1^* \geq T_2^* \geq \cdots \geq T_N^*$;

when $\alpha < 1$, if $\beta_i, i = 1, 2, ..., N$ is nonincreasing,then the sequence of optimal maintenance times $T_i^*, i = 1, 2, ..., N$ is in nondecreasing order,i.e.$T_1^* \leq T_2^* \leq \cdots \leq T_N^*$.

The proof is given in Appendix A.

Proposition 4.2. *In the long run, $AVC(\{T_i\}, N)$ increases with α increasing.*

The proof is given in Appendix A.

4.5 Numerical example

Consider a manufacturing system with production rate 180 units per year, and the annual demand rate is flat with the value of 90 units. The imperfect production proportion B follows the uniform distribution over the interval [0,0.2]. Moreover, the system lifetime is Weibull with CDF $F_{Z_1}(t) = 1 - exp(-\theta_2 t^{\theta_1})$, and $\theta_1 = 2, \theta_2 = 1, \beta_i = 1,$ for $i = 1, 2, ..., N$. Other parameters are

$\alpha = 1.2$

$r =$ \$50 per replacement

$c_r =$ \$12

$c_p =$ \$2

$c_d =$ \$0.4 per unit

$k =$ \$20 per setup

$h =$ \$0.3 per unit per unit time

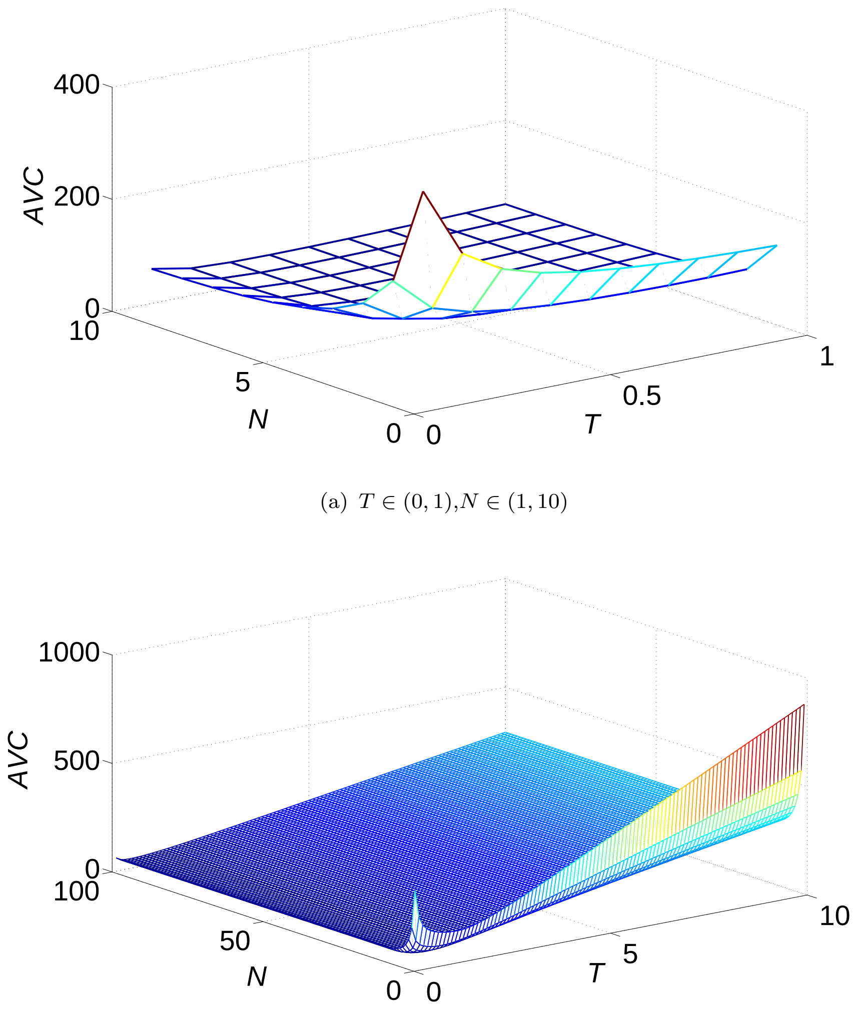
Under the parameter settings, for $(\{T_i\}, N)$ policy, $T_i^*, i = 1, 2, ..., N$ and N^* are obtained by solving the equations of first-order derivative under definition (4.10); for (T, N) policy, T^* and N^* are determined by implementing some direct search methods provided in commercial software. The result of optimal $(\{T_i\}, N)$ and (T, N) maintenance policy are listed in Table 4.1. It is seen that N^* under $(\{T_i\}, N)$ policy is larger than that under (T, N) policy, but $AVC(\{T_i\}, N)$ is smaller than $AVC(T, N)$, which implies that $(\{T_i\}, N)$ policy is superior to (T, N) policy since the latter is a special case of the former. It is also noted that $T_N^* < T^* < T_1^*$, which indicates that T^* is a moderate tradeoff of T_i^*s.

TABLE 4.1: Optimal $(\{T_i\}, N)$ and (T, N) policy

policy	$(\{T_i\}, N)$policy	(T, N) policy
N^*	5	4
optimal maintenance time	$T_1^* = 1.19$,$T_2^* = 1.06$,$T_3^* = 0.92$,$T_4^* = 0.77$,$T_5^* = 0.63$	$T^* = 1.07$
AVC^*	47.51	47.63

As mentioned in the definition (4.10), $AVC(T, N)$ tends to be convex to both T and N. This is investigated within different ranges of (T, N) in Figure 4.1. Although it is difficult to prove or search for all T and N, in practice, definition (4.10) is sufficient to support a useful optimum. Nevertheless, note that this observation is preserved under the condition in Lemma 4.1. It is noted that α and $\beta_i, i = 1, 2, ..., N$ characterize the effect of maintenance on machine deterioration and production quality respectively. In the example, we assume that $\alpha > 1$ and β_is are identical. In practice, however, various (α, β_i) are possible and the combination may even change with different production runs during a production cycle. Therefore, an extensive search of (α, β_i) combinations and its effect on optimal $AVC(T, N)$ is studied and presented in Table 4.2. We investigate 3 scenarios of β_i sequence,i.e. 1. identical β_is, 2. linear increasing β_is, 3. linear decreasing β_i. In some combinations, the optimal AVC may appear when $N \to \infty$, these AVC values are denoted $\times$ here.

Figure 4.2 presents the trends of T^* and N^* when α and β gradually change. It is shown that generally N^* first increases with α increasing and decreases



(a) $T \in (0,1)$,$N \in (1,10)$

(b) $T \in (0,10)$,$N \in (1,100)$

FIGURE 4.1: AVC with varying T and N

after α exceeds certain value(around 1), while for T^*, the opposite trend applies, it first decreases when α increases and increases after α is large enough($\geq$1). This observation shows that when $\alpha < 1$, system tends to favour longer scheduled maintenance time instead of more production runs, the explanation are two folds, the first is that system is already in a healthier status and the second is that increasing production runs will incur higher

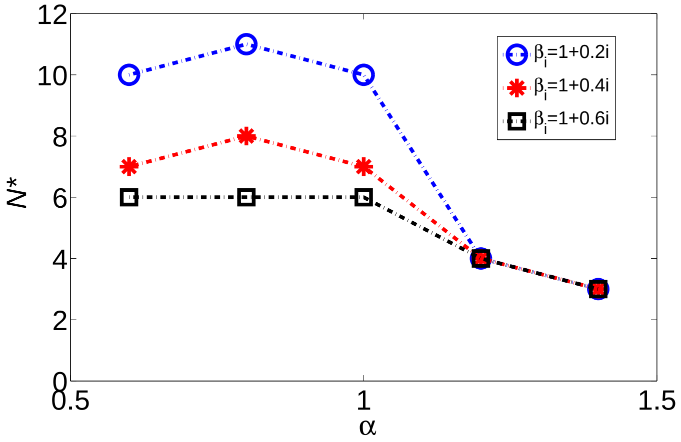
TABLE 4.2: Optimal $AVC(T, N)$ with various (α, β_i) combinations.

β_i	α				
	0.6	0.8	1	1.2	1.4
0.6	×	×	×	46.19	50.96
1	×	×	×	43.63	52.4
1.5	×	×	×	49.43	54.2
$1 + 0.2i$	35.59	36.4	41.22	49.3	53.7
$1 - 0.05i$	×	×	×	47.21	52.08
$1 - 0.2i$	×	×	×	45.77	51.1
$1 + 0.4i$	39.04	40.19	44.51	50.98	55
$1 + 0.6i$	41.8	43.16	47.12	52.65	56.3

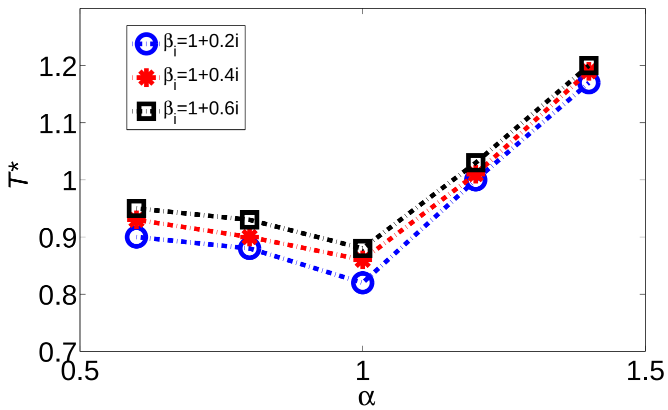
imperfect production costs. More specifically, imperfect production cost is the major concern in the total costs; when α is around 1, it is recommended to perform maintenance more frequently because system is not improved a lot after maintenance. In other words, system breakdown penalty is comparable to imperfect production cost; when $\alpha > 1$, it is better to stay in a production run rather than maintenance frequently because system is inferior to before after maintenance and the cost of breakdown is the major concern.

On the other hand, N^* tends to decrease when β_i increases and T^* increases with β_i increasing. The intuition is that when imperfect production risk becomes high with production run proceeding, it is better to reduce the number of production runs and maintain infrequently.

We next investigate the (T, N) policy when Z_i follow various parametric lifetime distributions. Specifically, three distributions are studied:



(a)



(b)

FIGURE 4.2: Optimal T and N with varying α.

- Weibull(λ, k), $f_{Z_1}(t) = \frac{k}{\lambda}(\frac{t}{\lambda})^{k-1} exp\left[-(\frac{t}{\lambda})^k)\right]$;
- Normal(μ, σ), $f_{Z_1}(t) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-(t-\mu)^2/2\sigma^2\right]$;
- Gamma(k, θ), $f_{Z_1}(t) = \frac{1}{\Gamma(k)\theta^k} t^{k-1} exp(-\frac{t}{\theta})$, where Γ is the gamma function.

The results are presented in Table 4.3~4.5. It is seen that firstly, optimal $AVC(T,N)$ decreases when the mean of any distribution increases, for normal distribution, it is clearly observed that optimal $AVC(T,N)$ increases when variance increases, implying a more unstable system will result in higher cost; secondly, when both mean and variance of the distributions are close, the optimal $AVC(T,N)$ are also close (which is labelled in italic font in the table). This shows that the model is robust to different distributions and can apply in different systems in practice.

TABLE 4.3: Optimal $AVC(T,N)$ when Z_1 ~Weibull(λ, k).

λ	k							
	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6
1	60.35	57.52	54.96	52.99	*51.45*	50.31	49.44	48.76
2	54.87	49.44	45.72	43.22	41.43	40.15	39.18	38.44
3	52.2	46.02	42.21	39.76	38.12	36.97	36.13	35.49

TABLE 4.4: Optimal $AVC(T,N)$ when Z_1 ~Normal(μ, σ).

σ	μ							
	1	2	3	4	5	6	7	8
0.2	42.24	33.84	31.99	31.22	30.77	30.6	30.6	30.6
0.5	44.66	34.38	32.21	31.37	30.9	30.62	30.6	30.6
1	*47.7*	36.3	32.9	31.68	31.09	30.73	30.62	30.6

TABLE 4.5: Optimal $AVC(T,N)$ when Z_1 ~Gamma(k, θ).

θ	k							
	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6
1	176.44	97.41	71.5	58.82	*51.45*	46.72	43.47	41.09
2	116.16	68.53	53.19	45.71	41.43	38.74	36.84	35.54
3	96.02	58.97	47.04	41.39	38.12	36.1	34.71	33.74

4.6 Conclusion

In this chapter the problem of production lot sizing is generalized by incorporating quality and reliability issues and multi-period production runs. An optimal maintenance policy is derived considering the scheduled preventive maintenance time and the number of production runs. Analytical results are obtained to explore the bounds of the optimal maintenance policy.

Our model complements the existing literature concerning quality and reliability issues in production lot sizing problems. It helps to make decisions under a more realistic environment and to answer questions such as how to coordinate all the relevant issues and balance various costs related to production and maintenance. Further research may extend the model by covering several other aspects. For example, in our model, it is implicitly assumed that the system is under continuous monitoring, other inspection policies can be used (Dieulle et al., 2003b), i.e. periodic inspection, sequential inspection, etc. Besides, external shocks which often appear and cause production interruption in real world also need to be investigated.

Chapter 5

CONDITION BASED MAINTENANCE FOR SYSTEMS UNDER DEPENDENT COMPETING FAILURES

5.1 Introduction

Most systems are inclined to failure in one of several ways and sometimes in more than one way at a time. Failure mechanisms may include wear, corrosion, shock loads, fatigue, etc. In general, it is beneficial to distinguish between different failure modes. For instance, wear or degradation modelling has attracted much attention since early 1990s (Lu and Meeker, 1993; Singpurwalla, 1995). Shocks are also important mechanisms accounting for product failures, they are interpreted as some intermittent, instantaneous and potentially harmful event (Finkelstein, 2007).

In practice, maintenance is carried out to restore systems to reduce failures and maintain high availability. As reviewed in Chapter 2, there is extensive study on maintenance scheduling when systems are subject to degradation or random shocks. For maintenance of systems that experience degradations, Grall et al. (2002) studied a single unit deteriorating system based on a Gamma process. The preventive replacement threshold was determined and inspections were scheduled to minimize maintenance cost function. Yang et al. (2008) developed a cost-effective maintenance schedule in manufacturing systems where there is degradation in equipment performance based on genetic algorithm. For maintenance of systems suffer random shocks, Lam and Zhang (2004) studied the maintenance problem of a repairable system subject to shocks under a geometric process framework. An optimal replacement N policy is adopted to minimize the average cost rate. Chien et al. (2006) assumed two types of failures and developed a hybrid age-based and replacement N maintenance policy.

Nevertheless, to better study complex systems whose failures are often the result of either intended operation themselves or external sources and schedule maintenance activities to enhance their reliability, it is desirable to model the system under a joint failure mechanism scheme. In fact, there are already several papers addressing this issue. Klutke and Yang (2002) derived an availability model for an inspected system subject to continuous smooth degradation and shocks that also cause additional degradation damage. Kharoufeh et al. (2006) derived the system lifetime distribution and the limiting average availability for a similar failure process. Huang and Askin (2003) worked on independent multiple catastrophic and degradation failure processes and Wang and Zhang (2005) studied two random shock processes using the extreme shock model and the δ-shock model. Li and Pham (2005) analyzed the reliability for a multi-state degraded system and developed an inspection-maintenance schedule. Ye et al. (2011) studied a distribution-based system under natural and traumatic failures and presented statistical inference methods to illustrate the model adequacy.

However, little research is devoted to consider the failure modes when they are dependent. Recently, Peng et al. (2010) proposed a multiple dependent competing failure model where the external random shocks contribute to the internal degradation. Huynh et al. (2011) modelled the dependency of failure modes by assuming that the arrival rate of shocks are functions of the degradation level. Liu (2012) considered the accelerated life test when there are dependent competing risks.

It is widely observed that when systems are more deteriorated, they are more vulnerable to external shocks. Therefore, an alternative approach of characterizing the dependency between failure modes is similar with the hazard potential in Singpurwalla (2006), by interpreting the potential of degradation as some unknown resource with which the system is endowed at the time of its inception. The destructive probability of a shock depends on the remaining degradation quota until a failure occurs. From this perspective, in this chapter, we study the systems under competing dependent failure modes and propose another innovative approach to model the dependency between shocks and degradation. The rest of this chapter is organized as follows. Section 5.2 lists the assumptions and analyses system reliability. Section 5.3 presents the maintenance modelling. In Section 6.4, a numerical example is given to illustrate the proposed model and sensitivity analysis is also given. Finally Section 6.5 concludes the article.

5.2 Assumptions and system reliability analysis

5.2.1 Assumptions

We assume the following model in this paper:

1. The system is subject to two dependent failure modes, i.e. natural failure caused by degradation and traumatic failure caused by shocks. The degradation is modeled by a stochastic process D(t) and a natural failure occurs when the degradation characteristic exceeds the failure threshold D_f.

2. Random shocks arrive according to a Poisson process $N(t), t \geq 0$,with intensity λ. The shock arrival process is s-independent of the degradation process. A traumatic failure occurs with a probability dependent the degradation level. More specifically, a shock at time t either breaks down the system with probability $p(t) = \exp(\alpha(D(t) - D_f))$, or has no impact on the system, where $\alpha > 0$ is a regression parameter.

3. The system is periodically inspected. The periodic inspection is assumed to be instantaneous, perfect and non-destructive. There is a cost c_i incurred for every inspection. If the system fails, it will remain idle until the next scheduled inspection.

4. The system is non-repairable. If it is detected to have failed, it will be replaced instantly with a new one and the replacement time is negligible.

5.2.2 System reliability analysis

The natural degradation due to continuous wear may follow one of the various degradation paths models, such as a linear degradation path with random coefficients Christer and Wang (1992); Lu and Meeker (1993); Peng et al. (2010), Gamma process (Liao et al., 2006) or randomized logistic degradation path (Li and Pham, 2005), etc. For illustration, a linear degradation path is used in this paper, i.e. $D(t) = \mathrm{B}t$, where B is a random variable. Denote by T_d the failure time of the natural degradation, then

$$P(T_d < t) = P(D(t) > D_f) = P(\mathrm{B}t > D_f) = 1 - F_{\mathrm{B}}\left(\frac{D_f}{t}\right), \tag{5.1}$$

where $F_{\mathrm{B}}(t)$ is the cumulative density function of B.

Let T_s be traumatic failure time. Then the hazard rate of traumatic failure given $\mathrm{B} = \beta$ is

$$h_s(t|\mathrm{B} = \beta) = \lambda \exp(\alpha(\beta t - D_f)). \tag{5.2}$$

Denote by $T = \min\{T_d, T_s\}$, the system failure time. The survival function of T is given by

$$R_T(t) = P(T > t) = P(T_d > t, T_s > t) = P(T_s > t|T_d > t)P(T_d > t), \tag{5.3}$$

where $P(T_d > t)$ is derived in (5.1), and $P(T_d > t, T_s > t)$ is derived as follows.

$$\begin{aligned} P(T_d > t, T_s > t) &= E[\exp(-\textstyle\int_0^t h_s(u|\mathrm{B}=\beta)\,du)|\mathrm{B} < \frac{D_f}{t}] \\ &= \textstyle\int_{-\infty}^{D_f/t} \exp(-\int_0^t \lambda \exp(\alpha(\beta u - D_f))\,du) f_{\mathrm{B}}(\beta)\,d\beta. \end{aligned} \tag{5.4}$$

Therefore, the system survival function is

$$R_T(t) = \int_{-\infty}^{D_f/t} \exp(-\int_0^t \lambda \exp(\alpha(\beta u - D_f))\,du) f_{\mathrm{B}}\,d\beta \cdot F_{\mathrm{B}}\left(\frac{D_f}{t}\right). \tag{5.5}$$

5.3 Maintenance modelling and optimization

5.3.1 Maintenance modelling

According to the renewal theory, the long run average cost is equal to the expected cost in a cycle divided by the expected cycle length. Therefore,

$$\lim_{t\to\infty} \frac{C(t)}{t} = \frac{\text{Expected total cost in a cycle}}{\text{Expected cycle length}} = \frac{E[TC]}{E[CL]}. \tag{5.6}$$

The total costs incurred in the model include inspection cost, maintenance cost and system downtime cost. Then the expected total cost of a renewal cycle is given as

$$E[TC] = c_i E[I] + c_d E[\xi] + c_R, \tag{5.7}$$

where c_i is the cost associated with each inspection, c_d is the downtime cost per unit time, c_R is the replacement cost, I is the number of inspections and ξ is the possible system downtime.

Let P_i be the probability that there are a total of i inspections in a cycle. There is either a replacement or doing nothing at time $i\tau$, where τ is the inspection period. Obviously, we have $(i-1)\tau < T \leq i\tau$.

Therefore,

$$P_i = P((i-1)\tau < T \leq i\tau) = F_T(i\tau) - F_T((i-1)\tau) = R_T((i-1)\tau) - R_T(i\tau), \tag{5.8}$$

where $F_T(t) = 1R_T(t)$ is the CDF of T.

Therefore,

$$E[I] = \sum_{i=1}^{\infty} iP_i = \sum_{i=1}^{\infty} i[R_T((i-1)\tau) - R_T(i\tau)]. \tag{5.9}$$

Suppose the i-th inspection reveals that the system has failed. Then the downtime is $\xi = i\tau - T$. Therefore, the expected downtime is given by

$$E[\xi] = \sum_{i=1}^{\infty} P_i \int_{(i-1)\tau}^{i\tau} (i\tau - t) dF_T(t) = \sum_{i=1}^{\infty} \int_{(i-1)\tau}^{i\tau} (t - i\tau) dR_T(t) [R_T((i-1)\tau) - R_T(i\tau)]. \tag{5.10}$$

The length of a renewal cycle $E[CL]$ can also be obtained similarly by conditioning on the number of inspections. When the i-th inspection reveal system failure, replacement is performed and the cycle ends with length $i\tau$. Therefore,

$$E[CL] = \sum_{i=1}^{\infty} P_i \cdot i\tau = \sum_{i=1}^{\infty} i\tau[R_T((i-1)\tau) - R_T(i\tau)]. \tag{5.11}$$

5.3.2 Solution procedure

Based on (5.7) $\sim$ (5.11), the average long-run maintenance cost rate as a function of τ, $AVC(\tau)$, is given as

$$AVC(\tau) = \frac{c_i \sum_{i=1}^{\infty} i[R_T((i-1)\tau)-R_T(i\tau)]+c_d \sum_{i=1}^{\infty} \int_{(i-1)\tau}^{i\tau}(t-i\tau)dR_T(t)[R_T((i-1)\tau)-R_T(i\tau)]+c_R}{\sum_{i=1}^{\infty} i\tau[R_T((i-1)\tau)-R_T(i\tau)]}. \tag{5.12}$$

The optimization function above is a complex nonlinear function which is not easy to obtain the optimum solution, if not impossible. Nelder-Mead downhill simplex method (Nelder and Mead, 1965) is a very popular direct search method to solve unconstrained optimization problems without the calculation of derivatives. We develop an iterative algorithm based on Nelder-Mead simplex method as follows.

The method generates (n+1) vertices for n dimensional space and each time gets a new test position by extrapolating the behavior of the objective function measured at each test point arranged as a simplex. The algorithm then chooses to replace one of existing vertices with the new test point through reflection, expansion, and contraction operations. Specifically, the algorithm is performed in the following steps:

Initialize: Choose $(n+1)$ vertices $\{X_1, X_2, \cdots, X_{n+1}\}$;

Step 1: Order $\{X_1, X_2, \cdots, X_{n+1}\}$ according to the function values $f(X_i), i = 1, \cdots, n+1$ increasingly, where $f(X) = AVC(\tau)$. Specifically, $f(X_1) = \min\{AVC(\tau)\}$ and $f(X_{n+1}) = \max\{AVC(\tau)\}$;

Step 2: Calculate X_o, the center of gravity of all points except X_{n+1}, i.e. $X_o = (1/n)\sum_{i=1}^{n} X_i$;

Step 3: (Reflection) Compute reflected point: $X_r = X_o + \alpha(X_o - X_{n+1})$. If the reflected point is better than the second worst, but not better than the best, i.e. $f(X_1) \leq f(X_r) < f(X_n)$, then obtain a new simplex by replacing the worst point X_{n+1} with the reflected point X_r , and go to Step 1;

Step 4: (Expansion) If $f(X_r) < f(X_1)$, then compute the expanded point $X_e = X_o + \gamma(X_o - X_{n+1})$. If the expanded point is better than the reflected point, i.e. $f(X_e) < f(X_r)$, then obtain a new simplex by replacing the worst point X_{n+1} with the expanded point X_e,and go to step 1. Else obtain a new simplex by replacing the worst point X_{n+1} with the reflected point X_r, and go to Step 1;

Step 5: (Contraction) Compute $X_c = X_{n+1} + \rho(X_o - X_{n+1})$. If $f(X_c) < f(X_{n+1})$, then replace X_{n+1} with X_c, and go to Step 1;

Step 6: (Reduction) Replace the point with $X_i = X_1 + \sigma(X_i - X_1)$ for all $i \in \{2, 3, \cdots, n+1\}$ and go to Step 1;

Step 7: Check $\{X_1, X_2, \cdots, X_{n+1}\}$, if $\sqrt{1/(n+1\sum_{i=1}^{n+1}[f(X_i - \bar{f})]^2)} < 0.01$, stop; otherwise, go to Step 1. Here the values of parameters are adopted as: $\alpha = 1, \gamma = 0.5$ and $\sigma = 0.5$.

5.4 Numerical example

Meeker and Escobar (1998)(Example 13.5, Chapter 12) studied a degradation-threshold-shock example of a laser device. The device degrades in terms of increase in operating current and the threshold D_f is 10 percent increase. They suggested that the degradation path is linear with the form

$$D(t) = \mathrm{B}t,$$

where 1/B follows Weibull(548.8,6.612). Besides degradation, shocks may also result in system failure. Therefore, similar to Ye et al. (2011), we assume that shocks arrive according to a homogeneous Poisson process with $\lambda = 6 \times 10^{-4}$.

The cost parameters are given as follows: $c_i = 1$ unit/inspection, $c_d = 5$ units/unit time and $c_R = 10$ units/replacement. Besides, the regression parameter in the destructive probability of shocks is assumed a moderate value of 0.5.

Using this parameter setting, the reliability of laser device is shown in Figure 5.1. Table 5.1 gives the search result of the proposed Nelder-Mead algorithm. It is seen that the optimal solution is $(\tau^*, AVC(\tau^*)) = (6.94, 0.426)$. Figure 5.2 depicts the average long-run maintenance cost rate as a function of the inspection time interval.

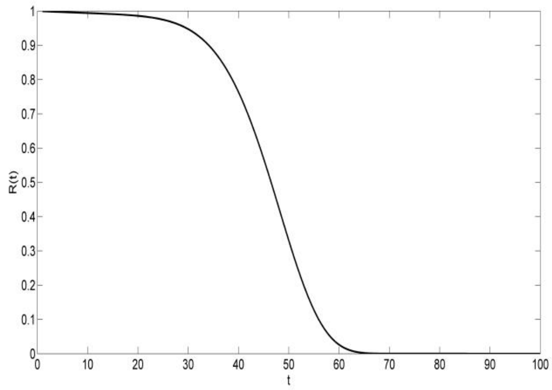


FIGURE 5.1: Plot of reliability function $R(t)$.

TABLE 5.1: Nelder-Mead algorithm result.

k	X_1	$AVC(X_1)$	X_2	$AVC(X_2)$	Search result
0	1	1.362	100	2.615	$X = 39$
1	39	1.1114	1	1.362	$X = 20$
2	20	0.768	39	1.114	$X = 10.5$
3	10.5	0.464	20	0.768	$X = 15.25$
4	10.5	0.464	15.25	0.579	$X = 5.75$
5	5.75	0.442	10.5	0.464	$X = 8.13$
6	8.13	0.431	5.75	0.442	$X = 6.94$
7	6.94	0.426	8.13	0.431	stop

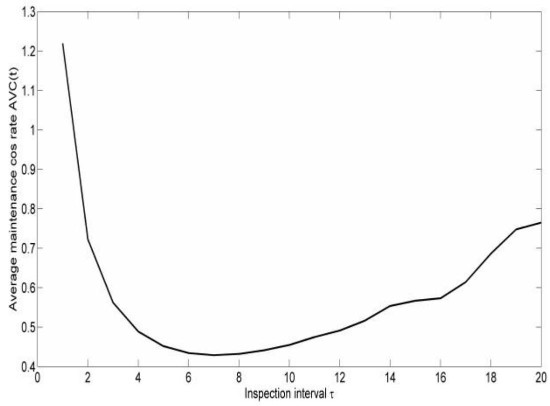


FIGURE 5.2: Plot of long-run average maintenance cost rate versus inspection interval .

A sensitivity analysis is also performed to analyse the effect of the model parameters on the optimal solutions. The model parameters of interest include the arrival rate of random shocks, λ, the natural failure threshold, D_f and the parameter in the destructive probability of shocks, α. The effect of $\pm 50\%$ change of parameters is depicted in Table II. It is seen that AVC is more sensitive to D_f than λ and α., which is mainly due to the inspection cost. Overall the model is quite robust to the parameters.

TABLE 5.2: Sensitivity analysis of parameters λ, D_f and α within $\pm 50\%$ change.

Parameters	Variation	τ	$AVC(\tau)$	Relative bias on AVC (%)
λ	+50%	6.97	0.464	9.5
	-50%	6.84	0.395	-7.14
D_f	+50%	9.31	0.313	-26.2
	-50%	4.71	0.558	30.95
α	+50%	6.96	0.448	4.76
	-50%	6.88	0.4	-4.75

5.5 Conclusion

In this chapter, we propose a new model to characterize dependent competing failures. Specifically, we consider natural failure due to degradation and traumatic failure caused by random shocks. In view of the degradation analogous to the hazard potential which appears in reliability studies, the correlation of degradation and shocks is built. System reliability is analyzed and an optimal inspection-maintenance policy is derived to minimize the average long-run maintenance cost rate.

For future research directions, the maintenance model developed in this article can be extended to include preventive maintenance. Minimal repairs can also be included to rectify the traumatic failures. In addition, a non-linear degradation path can be constructed. Finally, availability maximization can be considered besides the cost effective model in some circumstances.

Chapter 6

ACCELERATED DEGRADATION TEST PLANNING USING THE INVERSE GAUSSIAN PROCESS

6.1 Introduction

To assess reliability of newly designed products, engineers often resort to accelerated tests in order to shorten the life of products or hasten the degradation of their performance, where products are exposed to harsh conditions, e.g., a combination of random vibration, higher temperature, voltage, and pressure. The main purpose of such accelerated testing is to obtain reliability information quickly so as to save time and money.

Nowadays, many new products are designed to be very reliable because (a) rapid advances in technology, (b) increasing consumer expectations, and (c) global competition. For example, a electronic product may be viewed as a complex system that consists of many components. To maintaining high reliability for the entire system, it generally requires that the individual components have extremely high reliability (Lu and Meeker, 1993). Traditional accelerated life test (ALT) methods are not suited for such reliable products as extremely long test duration is required to yield sufficient failures. On the other hand, we often observe that failure of a product is associated with the degradation of some quality characteristic (QC). Degradation of the product accumulates over time and causes a failure when the degradation exceeds a failure threshold. This naturally provides a linkage between product degradation and reliability (Singpurwalla, 1995; Wang, 2010; Ye et al., 2011). The degradation is most often hastened under severe stresses. Therefore, we can use accelerated degradation tests (ADTs) to quickly obtain the degradation information. In a simple constant-stress ADT experiment, a number of units are allocated to several stress levels, and the degradation levels of these units are measured, analyzed and extrapolated to the failure threshold so as to estimate the life characteristics of interest under use conditions.

ADTs are able to greatly shorten the testing duration and has attracted much attention. There are two classes of models for ADT data. The first class is called general path models proposed by (Lu and Meeker, 1993). Some developments of models in this class can be found in Meeker et al. (1998); Bae et al. (2007); Shi and Meeker (2012). On the other hand, the other class of models uses stochastic processes to capture the time-dependent structure of the degradation over time. Two popular models are the Wiener process and the Gamma process. Tseng and Wen (2000) proposed using step-stress ADT (SSADT) to assess the reliability of a light emitting diode by using empirical regression method. Optimal ADT settings were obtained by minimizing the estimated p-quantile of the product's lifetime distribution subject to a constraint on the total cost. Following this striking work, some SSADT models have been developed based on the assumptions of Wiener processes Tang et al. (2004); Liao and Tseng (2006); Ge et al. (2010, 2011); Lim (2012); Lim and Yum (2011), and Gamma process Tseng et al. (2009); Tsai et al. (2012), respectively.

Although the Wiener process and the Gamma process have received intensive applications in degradation data analysis, it is obvious that two models cannot handle all degradation problems. For instance, Wang and Xu (Wang and Xu, 2010) found that neither models fits the GaAs laser degradation data (Meeker and Escobar, 1998, Example 13.5) well. Another attractive degradation model with monotone paths is the inverse Gaussian (IG) process proposed by Wasan (1968). Recently, Wang and Xu (2010) proposed the IG process for degradation modeling and investigated semi-parametric inference for this process. After that, Ye et al. (2012) systematically investigated the IG process and showed that compared with the Gamma process, the IG process has many superb properties when dealing with covariates and random effects. Therefore, this process can be an important family for degradation analysis.

The purpose of this chapter is to investigate the planning of ADT experiments using the IG process. We first look at the simple IG process, i.e., IG process without random effects. The objective of ADT planning is to properly choose the stress levels and the number of units allocated to each stress in order to minimize the asymptotic variance of the p-quantile under use conditions. Parameter estimation for the simple IG process model is discussed. based on which the asymptotic variance of the p-quantile can be derived. Then the optimal stress levels and the allocation scheme can be obtained. Next, we also discuss ADT planning for a random-effects IG process model proposed by Wang and Xu (2010). This random effects model is called random volatility model by Ye et al. (2012). In reality, it is not difficult to observe unit-to-unit difference within a product population due to some unobserved factors, such as variations in the raw materials. Such heterogeneity is often modeled by a random-effects term. Random-effects degradation models are believed to be more realistic in modeling product degradation, and such models have found more and more applications recently, e.g., see Si et al. (2011) for an overview. Wang and Xu (2010) found that the random volatility model provides a good fit to the laser data in Meeker and Escobar (1998). Therefore, we believe that ADT planning for this random-effects model is meaningful.

The rest of the chapter is organized as follows. Section 6.2 presents the ADT settings and develops ADT planning for the simple IG process model. ADT planning for the random-volatility IG process model are discussed in Section 6.3. Section 6.4 demonstrates the developed methods by using a stress-relaxation example from Yang (2007). Section 6.5 concludes the chapter.

6.2 ADT Planning for the Simple IG process

Consider a product whose degradation is measurable. Let $\{Y(t), t \geq 0\}$ with $Y(0) = 0$ be the degradation path of a random selected unit. Failure of this product is defined to be the event that $Y(t)$ crosses a pre-specified failure threshold D, and the associated first-passage-time is denoted by T_D.

6.2.1 The IG process

We assume that the degradation follows an IG process, i.e., $Y(t)$ has independent increments and $Y(t) \sim \mathcal{IG}(\mu\Lambda(t), \lambda\Lambda^2(t))$, where $\Lambda(t)$ is a monotone increasing function with $\Lambda(0) = 0$, and $\mathcal{IG}(a, b), a, b > 0$ is the IG distribution with probability density function (PDF)

$$f_{\mathcal{IG}}(y; a, b) = \left(\frac{b}{2\pi y^3}\right)^{1/2} \exp\left[-\frac{b(y-a)^2}{2a^2 y}\right], y > 0. \tag{6.1}$$

The mean and variance of $Y(t)$ are $\mu\Lambda(t)$ and $\mu^3\Lambda(t)/\lambda$, respectively. We shall call this process the simple IG process because it assumes homogeneity among the product population. IG process with random effects will be considered in the next section. Because the path of the IG process is strictly increasing, we can obtain the cumulative distribution function (CDF) of T_D

as

$$F_{T_D}(t) = P(Y(t) > D) = \Phi\left[\sqrt{\frac{\lambda}{D}}(\Lambda(t) - \frac{D}{\mu})\right] - e^{2\lambda\Lambda(t)/\mu}\Phi\left[-\sqrt{\frac{\lambda}{D}}(\Lambda(t) + \frac{D}{\mu})\right], \tag{6.2}$$

where $\Phi(\cdot)$ is the standard normal CDF. Differentiation of $F_{T_D}(t)$ with respect to t yields

$$\begin{aligned} f_{T_D}(t) &= \sqrt{\frac{\lambda}{D}}\phi\left[\sqrt{\frac{\lambda}{D}}(\Lambda(t) - \frac{D}{\mu})\right]\Lambda'(t) - 2\frac{\lambda}{\mu}\Lambda'(t)e^{2\lambda\Lambda(t)/\mu}\Phi\left[-\sqrt{\frac{\lambda}{D}}(\Lambda(t) + \frac{D}{\mu})\right] \\ &\quad + \sqrt{\frac{\lambda}{D}}\Lambda'(t)e^{2\lambda\Lambda(t)/\mu}\phi\left[-\sqrt{\frac{\lambda}{D}}(\Lambda(t) + \frac{D}{\mu})\right], \end{aligned} \tag{6.3}$$

where $\phi(\cdot)$ is the standard normal PDF and $\Lambda'(t) = d\Lambda(t)/dt$. Based on the CDF (6.2), the p-quantile of the failure time distribution does not have a closed form. Chhikara (1988) and Ye et al. (2012) noted that when $\lambda \cdot \Lambda(t)$ is large, which is often true when t is large, $Y(t)$ is approximately normal with mean $\mu\Lambda(t)$ and variance $\mu^3\Lambda(t)/\lambda$. Then the CDF and PDF of T_D can be approximated as

$$F_{T_D}(t) \doteq 1 - \Phi\left[\frac{D - \mu\Lambda(t)}{\sqrt{\mu^3\Lambda(t)/\lambda}}\right] = \Phi\left[\sqrt{\lambda/\mu}\cdot\sqrt{\Lambda(t)} - \frac{D\sqrt{\lambda/\mu^3}}{\sqrt{\Lambda(t)}}\right]. \tag{6.4}$$

The p-quantile of T_D, based on this approximation, is

$$\xi_p = \Lambda^{-1}\left(\frac{\mu}{4\lambda}\left(z_p + \sqrt{z_p^2 + 4D\lambda/\mu^2}\right)^2\right), \tag{6.5}$$

where z_p is the standard normal p-quantile and $\Lambda^{-1}(\cdot)$ is the inverse function of $\Lambda(\cdot)$.

6.2.2 ADT Settings and Assumptions

Suppose a total number of N units are put into test. Let s_0 and s_H be the usage stress and the maximum allowable stress, respectively. To timely collect the degradation information, these units are allocated to J stress levels $s_1 < s_2 < \cdots < s_J$ with $s_0 < s_1$ and $s_J = s_H$. Suppose N_j units are allocated to the j-th level, $j = 1, 2, \cdots, J$. The stress affects the degradation of these units. We assume that the stress affects the degradation rates. More specifically, we assume $\mu = h(s)$ and λ constant over s, where $h(s)$ is a link function reflecting the effects of the stress on the degradation process. For simplicity and without loss of generality, the following assumptions are made.

(a) The measurement time intervals τ_j and the number of measurements K_j under the j-th stress level, $j = 1, 2, \cdots, J$, are pre-determined.

(b) The link function follows one of the following acceleration relations:

- *Power law relation*: $h(s) = \xi_0 \cdot s^{\alpha}$;
- *Arrhenius relation*: $h(s) = \xi_0 \cdot e^{-\alpha/s}$;
- *Exponential relation*: $h(s) = \xi_0 \cdot e^{\alpha s}$.

In many real applications, the time allowed for the test is often given by the manager, and the times at which the units are measured are often predetermined because of the working time of the experimenters. Therefore, it is reasonable to assume that τ_j and K_j are given. In case these two variables can be flexibly determined, we can treat them as two decision variables in our models developed below and optimally determined. When the assumed stress-degradation acceleration relation, i.e., $h(x)$, is correct, we can use a two-stress ADT, i.e., $J = 2$. This plan is called the minimum variance plan in the ALT literature. But the minimum variance plan does not allow us to check the validity of the assumed stress-degradation acceleration relationship. In this case, we can use a three-stress ADT, i.e., $J = 3$. Under our settings, the purpose of ADT planning is to optimally determine the stress levels (s_j) and the number of samples (N_j) for each stress level, which will be investigated in the following subsections.

6.2.3 Normalizing the Stress

Depending on the acceleration relationship of the stress on the degradation rate, we can standardize the stress levels as follows (Lim and Yum, 2011):

$$
\begin{aligned}
x_j &= \frac{\ln s_j - \ln s_0}{\ln s_H - \ln s_0} && \text{for the power law relation,} \\
&= \frac{1/s_0 - 1/s_j}{1/s_0 - 1/s_H} && \text{for the Arrhenius relation,} \\
&= \frac{s_j - s_0}{s_H - s_0} && \text{for the exponential relation.}
\end{aligned}
$$

From the above standardization, it is readily seen that $x_0 = 0, x_J = 1$, and $0 < x_j \leq 1$ for $j = 1, 2, \cdots, J$. Then $h(s)$ can be rewritten as a function of x:

$$h(x) = \exp(\alpha_0 + \alpha_1 x) \tag{6.6}$$

where

$\alpha_0 = \ln \xi_0 - \alpha/s_0, \alpha_1 = \alpha(1/s_0 - 1/s_H)$ for the Arrhenius function,

$\alpha_0 = \ln \xi_0 + \alpha \ln s_0, \alpha_1 = \alpha(\ln s_H - \ln s_0)$ for the power law function,

$\alpha_0 = \ln \xi_0 + \alpha s_0, \alpha_1 = \alpha(s_H - s_0)$ for the exponential function.

6.2.4 Statistical Inference

Following the argument in the assumptions, suppose that the i-th unit under the j-th stress level is measured at $t_{ijk} = k\tau_j$ with observations

$Y_{ij}(t_{ijk})$, $k = 0, 1, \cdots, K_j$. Let $y_{ijk} = Y_{ij}(t_{ijk}) - Y_{ij}(t_{ij,k-1})$ be the observed increments and $\Lambda_{ijk} = \Lambda(t_{ijk}) - \Lambda(t_{ij,k-1})$. Then the log-likelihood function, up to a constant, can be expressed as

$$l(\boldsymbol{\theta}) = \sum_{j=1}^{J} \sum_{i=1}^{N_j} \sum_{k=1}^{K_j} \left[\frac{\ln\lambda}{2} + \ln \Lambda_{ijk} - \frac{\lambda(y_{ijk} e^{-\alpha_0 - \alpha_1 x_j} - \Lambda_{ijk})^2}{2 y_{ijk}} \right]. \tag{6.7}$$

where $\boldsymbol{\theta}$ is the parameter vector including $\lambda, \alpha_0, \alpha_1$ and the parameters in $\Lambda(\cdot)$. This log-likelihood function can be easily maximized to obtain the maximum likelihood estimators (MLEs). We may also take the first derivative of $l(\boldsymbol{\theta})$ with respect to λ. Setting this partial derivative to zero yields the estimate of λ when α_0, α_1 and the parameters in $\Lambda(\cdot)$ are assumed fixed. Substituting this estimate into (6.7) gives the profile log-likelihood. Direct maximization of the profile log-likelihood function yields the MLEs of the remaining parameters. The MLE of σ can then be subsequently determined.

The Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ can be derived once the parametric form of $\Lambda(\cdot)$ is given. For example, if we assume a linear function for $\Lambda(\cdot)$, i.e., $\Lambda(t) = t$, then $\Lambda(\cdot)$ does not involve unknown parameters and thus $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \lambda)'$. It is readily shown that the information matrix is given by

$$\mathbf{I}(\boldsymbol{\theta}) =$$
$$\begin{pmatrix} \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \lambda e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk} & \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \lambda x_j e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk} & 0 \\ & \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \lambda x_j^2 e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk} & 0 \\ & & \frac{1}{2\lambda^2}\sum_{j=1}^{J} K_j N_j \end{pmatrix}, \tag{6.8}$$

On the other hand, if $\Lambda(t) = t^{\beta}$, then $\boldsymbol{\theta} = (\lambda, \alpha_0, \alpha_1, \beta)'$. The information matrix is

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{pmatrix} E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda^2}\right] & 0 & 0 & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\beta}\right] \\ & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0\partial\alpha_1}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0\partial\beta}\right] \\ & & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_1^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_1\partial\beta}\right] \\ & & & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta^2}\right] \end{pmatrix}. \tag{6.9}$$

Detailed expressions of the elements within the matrix can be found in the appendix.

6.2.5 Optimization Problem

An ADT experiment is characterized by the total number of test units available, the number of stress levels used in the test as well as the stress value of each level, the allocation scheme of the testing units to each stress level, the test duration, and the measurement time interval, etc. Following the assumptions, we have assumed that the number of units, the test duration and the measurement time interval are given. Therefore, the objective of the ADT planning is to determine the optimal stress levels as well as the proportion of units allocated to each level based on some optimization criterion. Usually, we are concerned with a small quantile (ξ_p) of the time to failure under normal use conditions. Therefore, our objective here is to minimize $\text{Avar}(\hat{\xi}_p)$, the asymptotic variance of $\hat{\xi}_p$. When the exact distribution (6.2) for T_D is used, the p-quantile does not bear a closed form. But according to (Tsai et al., 2012), the asymptotic variance can be obtained through the delta method as

$$\text{Avar}(\hat{\xi}_p) = \frac{\boldsymbol{\omega}'(\boldsymbol{\theta})\mathbf{I}^{-1}(\boldsymbol{\theta})\boldsymbol{\omega}(\boldsymbol{\theta})}{[f_{T_D}(\xi_p)]^2}, \tag{6.10}$$

where $\mathbf{I}^{-1}(\boldsymbol{\theta})$ is the inverse of the Fisher information matrix and $\boldsymbol{\omega}(\boldsymbol{\theta})$ is the first derivative of $F_{T_D}(t)$ with respect to $\boldsymbol{\theta}$, with t evaluated at ξ_p.

Specifically, when $\Lambda(t) = t^{\beta}$, $\boldsymbol{\omega}(\boldsymbol{\theta})$ is given by

$$\boldsymbol{\omega}(\boldsymbol{\theta}) = \left(\frac{\partial F_{T_D}(t)}{\partial \lambda}, \frac{\partial F_{T_D}(t)}{\partial \alpha_0}, \frac{\partial F_{T_D}(t)}{\partial \alpha_1}, \frac{\partial F_{T_D}(t)}{\partial \beta} \right)' \Bigg|_{t=\xi_p}. \tag{6.11}$$

It is straightforward to write down the detailed expression for each element in $\boldsymbol{\omega}(\boldsymbol{\theta})$. But they turn out to be quite complicated. When the approximate distribution (6.4) is used, the p-quantile can be specified as

$$\xi_p = \left[\frac{\exp(\alpha_0)}{4\lambda} \left(z_p + \sqrt{z_p^2 + 4D\lambda \exp(-2\alpha_0)} \right)^2 \right]^{1/\beta}. \tag{6.12}$$

Then the asymptotic variance is $(\nabla \xi_p)' \mathbf{I}^{-1}(\boldsymbol{\theta}) \nabla \xi_p$, where $\nabla \xi_p$ is the first derivative of ξ_p with respect to $\boldsymbol{\theta}$. The derivative $\nabla \xi_p$ is much easier to evaluate than $\omega(\boldsymbol{\theta})$. Therefore, we would recommend using (6.12) for the planning.

With the asymptotic variance of $\hat{\xi}_p$ on hand, the optimization problem can be formulated as follows.

$$\begin{aligned}
\textbf{Minimize} \quad & \operatorname{Avar}(\hat{\xi}_p) \\
\textbf{subject to} \quad & 0 \le x_j \le 1, j = 1, 2, \cdots, J, \\
& x_M = x_H \text{ and } x_0 \le x_j \le x_H, j = 1, 2, \cdots, J-1, \\
& \textstyle\sum_{j=1}^{J} N_j = N, \\
& 0 < N_j \le N, j = 1, 2, \cdots, J.
\end{aligned} \tag{6.13}$$

To solve the mixed nonlinear integer programming problem, some software packages such as Matlab can be used.

6.3 ADT Planning for the Random-Effects Model

6.3.1 The Random Volatility Model

Random-effects models are useful when there are substantial unit-to-unit differences due to variations in the raw materials and other factors. Wang and Xu Wang and Xu (2010) proposed a method to incorporate random effects in the inverse Gaussian process $\{Y(t); t > 0\}$ by letting $\lambda \sim$ Gamma(δ, γ^{-1}), $\delta, \gamma > 0$, with PDF

$$g(\lambda; \delta, \gamma) = \frac{\gamma^{\delta}\lambda^{\delta-1}}{\Gamma(\delta)}\exp(-\gamma\lambda), \lambda > 0, \tag{6.14}$$

where $\Gamma(\delta)$ is the Gamma function. This model is called random volatility model by (Ye et al., 2012). The unconditional distribution of $\mathbf{Y}(t)$ is given by

$$f_{Y(t)}(y) = \frac{\Gamma(\delta + 1/2)}{\Gamma(\delta)}\gamma^{\delta}\sqrt{\frac{\Lambda^2(t)}{2\pi y^3}} \cdot \left[\gamma + \frac{(y - \mu\Lambda(t))^2}{2\mu^2 y}\right]^{-\delta-1/2}, \tag{6.15}$$

It is readily shown that the mean of $Y(t)$ is still $\mu\Lambda(t)$ while the variance is given by $\gamma\mu^3\Lambda(t)/(\delta-1)$ when $\delta > 1$. To compute the distribution of the first-passage-time to the failure threshold D, we use the the normal distribution to approximate $[Y(t)|\lambda]$ and then marginalize over λ. Therefore, the CDF of T_D can be approximated as

$$F_{T_D}(t) \doteq F_{t_{2\delta}}\left(\frac{\delta^{1/2}(\mu\Lambda(t)-D)}{\mu\sqrt{\mu\Lambda(t)\gamma}}\right), \tag{6.16}$$

where $t_{2\delta}$ is the student t-distribution with 2δ degrees of freedom. Then the p-quantile of T_D is given by

$$\xi_p = \Lambda^{-1}\left(\frac{\mu\gamma}{4\delta}\left(t_{2\delta,p}+\sqrt{t_{2\delta,p}^2+4\delta D/(\gamma\mu^2)}\right)^2\right), \tag{6.17}$$

where $\Lambda^{-1}(\cdot)$ is again the inverse function of $\Lambda(\cdot)$, and $t_{2\delta,p}$ is the p-quantile of the student t-distribution with 2δ degrees of freedom.

6.3.2 Assumptions

Suppose a total number of N testing units are available. The normalized stress x would affect the degradation process. Similar to Section 6.2, we assume $\mu = h(x) = \exp(\alpha_0 + \alpha_1 x)$ while other parameters are independent of x. This is a legitimate stress-degradation accelerated relation, as both the mean degradation path $\mu\Lambda(t)$ and the variation of the degradation path

$\mu^3\gamma\Lambda(t)/(\delta-1)$ are increasing in x. Other assumptions about the ADT settings in Section 6.2.2 carry over to this section. That is, the respective observation time interval and number of measurements are τ_j and K_j for units under the j-th stress level, $j = 1, 2, \cdots, J$.

6.3.3 Statistical Inference

When the degradation data of N units under J normalized stress levels are available, we can use maximum likelihood estimation to estimate the parameters. The log-likelihood function, up to a constant, can be derived as

$$l(\boldsymbol{\theta}) = \sum_{j=1}^{J}\sum_{i=1}^{N_j} \left\{ \ln\Gamma(\delta + K_j/2) - \ln\Gamma(\delta) + \delta\ln\gamma + \sum_{k=1}^{K_j}\left(\ln\Lambda_{ijk} - \frac{3}{2}\ln y_{ijk}\right) - (\delta + \frac{K_j}{2})\ln\left[\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk} - e^{\alpha_0+\alpha_1 x_j}\Lambda_{ijk})^2}{2e^{2\alpha_0+2\alpha_1 x_j} y_{ijk}}\right]\right\}, \tag{6.18}$$

where $\boldsymbol{\theta}$ is the parameter to estimate, including $\alpha_0, \alpha_1, \delta$ and γ.

While the estimation of parameters can be performed through direct maximization of the likelihood function. An alternative optimization method which turns out to be efficient is through expectation maximization (EM) algorithm. Ye and Chen (Ye et al., 2012) have developed the EM algorithm for this random-volatility model when the testing conditions are identical. In the presence of more than one stress level, the EM algorithm can be developed in a similar vein. Denote λ_{ij} the realization of the unobserved random effect for the i-th unit under the j-th stress level, $i = 1, 2, \cdots, N_j$ and $j = 1, 2, \cdots, J$. Given the observed degradation data $\mathbf{Y}$ as well as the random effects λ_{ij}, the complete data log-likelihood can be expressed as

$$l_C(\boldsymbol{\theta}) = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\left[\frac{1}{2}\ln\lambda_{ij} + \ln\Lambda_{ijk} - \frac{\lambda_{ij}(y_{ijk} - e^{\alpha_0+\alpha_1 x_j}\Lambda_{ijk})^2}{2\mu^2 y_{ijk}}\right] + \sum_{j=1}^{J}\sum_{i=1}^{N_i}\left[\delta\ln\gamma + \delta\ln\lambda_{ij} - \ln\Gamma(\delta) - \gamma\lambda_{ij}\right]. \tag{6.19}$$

To invoke the EM algorithm, we need to compute $E[\lambda_{ij}|\mathbf{Y}]$ and $E[\ln\lambda_{ij}|\mathbf{Y}]$. Note that conditional on $\mathbf{Y}$, λ_{ij} still follows a gamma distribution with parameters

$$\tilde{\delta}_{ij} = K_j/2 + \delta \text{ and } \tilde{\gamma}_{ij} = \gamma + \sum_{k=1}^{K_j}\frac{[y_{ijk} - \Lambda_{ijk}\exp(\alpha_0 + \alpha_1 x_j)]^2}{2y_{ijk}\exp(2\alpha_0 + 2\alpha_1 x_j)}$$

Therefore, we have

$$E(\lambda_{ij}|\mathbf{Y}) = \tilde{\delta}_{ij}/\tilde{\gamma}_{ij} \text{ and } E(\ln\lambda_{ij}|\mathbf{Y}) = \psi(\tilde{\delta}_{ij}) - \ln(\tilde{\gamma}_{ij}), \tag{6.20}$$

where $\psi(\cdot)$ is the digamma function. The two expectations in (6.20) can be used to compute the Q-function at the E-step of an EM iteration, which is the expectation of $l_C(\boldsymbol{\theta})$ conditional of $\mathbf{Y}$ as well as the parameter estimates from the last EM iteration. Similar to (Ye et al., 2012), it is readily shown that the Q-function can be broken down into two parts. The first part involves a, b and parameters in $\Lambda(\cdot)$ and the second part is a simple Gamma log-likelihood function involving δ and γ. These two parts can be independently optimized to obtain the parameter estimates for the next iteration.

The Fisher information matrix for the parameters $\boldsymbol{\theta}$ can be derived by taking the expectation of the negative of the second derivative of (6.18). Detailed expression of $\mathbf{I}(\boldsymbol{\theta})$ depends on the parameters in $\Lambda(\cdot)$. When $\Lambda(t) = t^{\beta}$, $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \delta, \gamma, \beta)'$ and the information matrix is

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{pmatrix} E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \alpha_1}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \delta}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \gamma}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \beta}\right] \\ E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_0}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \delta}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \gamma}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta}\right] \\ E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta \partial \alpha_0}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta \partial \alpha_1}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta \partial \gamma}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta \partial \beta}\right] \\ E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \alpha_0}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \alpha_1}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \delta}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma^2}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \beta}\right] \\ E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha_0}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha_1}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \delta}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \gamma}\right] & E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta^2}\right] \end{pmatrix}. \tag{6.21}$$

The expressions for the elements in $\mathbf{I}(\boldsymbol{\theta})$ can be found in the appendix.

6.3.4 Optimal ADT planning

When $\Lambda(t) = t^{\beta}$, the p-quantile ξ_p under the use conditions can be specified based on (6.17) as

$$\xi_p = \left[\frac{\gamma \exp(\alpha_0)}{4\delta} \left(t_{2\delta,p} + \sqrt{t_{2\delta,p}^2 + 4\delta D \exp(-2\alpha_0)/\gamma} \right)^2 \right]^{1/\beta}. \tag{6.22}$$

The estimate of ξ_p can be obtained by substituting the MLE of $\boldsymbol{\theta}$ into (6.22). The asymptotic variance of $\hat{\xi}_p$ can be obtained based on the delta method as

$$\text{Avar}(\hat{\xi}_p) = (\nabla \xi_p)' \mathbf{I}^{-1}(\boldsymbol{\theta}) \nabla \xi_p \tag{6.23}$$

where $\nabla \xi_p$ is the first derivative of ξ_p with respect to $\boldsymbol{\theta}$, which is given by

$$\nabla \xi_p = \left(\frac{\partial \xi_p}{\partial \alpha_0}, \frac{\partial \xi_p}{\partial \alpha_1}, \frac{\partial \xi_p}{\partial \delta}, \frac{\partial \xi_p}{\partial \gamma}, \frac{\partial \xi_p}{\partial \beta} \right)'. \tag{6.24}$$

The detailed expression for each of the five elements can be found in the appendix.

With the asymptotic variance in hand, the optimization model is the same as (7.15). We can solve the model by using some commercial software to obtain the optimal stress levels as well as the optimal proportion of units allocated to each level.

6.4 Numerical example

The stress relaxation data in Yang (2007), (Example 8.7, pp.351) are used here to illustrate the proposed procedure. The stress relaxation is the loss of stress in a component subject to a constant strain over time. For example, the contacts of electrical connectors often fail due to excessive stress relaxation. The electrical connector is said to have failed if the stress relaxation exceeds 30%, i.e., $D = 30$. Data are collected under three temperature levels, i.e., 65°C, 85°C and 100°C. The time intervals between measurements are tabulated in Table 6.1 in the appendix. Note that the 7-th point of the second unit under 65°C (labelled □ in Yang (2007)) is removed to preserve the monotonicity of the stress relaxation. Table 6.2 in the appendix lists the measurement epochs under each temperature. The stress relaxation data and the measurement times are tabulated as follows in Table 6.1 and 6.2.

Yang (2007) used regression to fit each degradation path and extrapolated to the failure threshold to obtain the pseudo failure time for each unit. He then used log-normal distribution to fit the pseudo failure times under each temperature level and used the Arrhenius relationship to link the failure time distributions. Here we consider the stochastic process approach. In keeping with Yang (2007), we assume the normal use stress

TABLE 6.1: Stress relaxation data under three temperature levels.

Temperature	ID	Stress loss
65°C	1	2.12, 2.7, 3.52, 4.25, 5.55, 6.12, 6.75, 7.22, 7.68, 8.46, 9.46
	2	2.29 3.24, 4.16, 4.86, 5.74, 6.85, *, 7.40, 8.14, 9.25, 10.55
	3	2.4 3.61, 4.35, 5.09, 5.5, 7.03, 8.24, 8.81, 9.629, 10.27, 11.11
	4	2.31 3.48, 5.51, 6.2, 7.31, 7.96, 8.57, 9.07, 10.46, 11.48, 12.31
	5	3.14 4.33, 5.92, 7.22, 8.14, 9.07, 9.44, 10.09, 11.2, 12.77, 13.51
	6	3.59 5.55, 5.92, 7.68, 8.61, 10.37, 11.11, 12.22, 13.51, 14.16, 15
85°C	7	2.77, 4.62, 5.83, 6.66, 8.05, 10.61, 11.2, 11.98, 13.33, 15.64
	8	3.88 4.37, 6.29, 7.77, 9.16, 9.9, 10.37, 12.77, 14.72, 16.8
	9	3.18 4.53, 6.94, 8.14, 8.79, 10.09, 11.11, 14.72, 16.47, 18.66
	10	3.61 4.37, 6.29, 7.87, 9.35, 11.48, 12.4, 13.7, 15.37, 18.51
	11	3.42 4.25, 7.31, 8.61, 10.18, 12.03, 13.7, 15.27, 17.22, 19.25
	12	5.27 5.92, 8.05, 9.81, 12.4, 13.24, 15.83, 17.59, 20.09, 23.51
100°C	13	4.25, 5.18, 8.33, 9.53, 11.48, 13.14, 15.55, 16.94, 18.05, 19.44
	14	4.81 6.16, 7.68, 9.25, 10.37, 12.4, 15, 16.2, 18.24, 20.09
	15	5.09 7.03, 8.33, 10.37, 12.22, 14.35, 16.11, 18.7, 19.72, 21.66
	16	4.81 7.5, 9.16, 10.55, 13.51, 15.55, 16.57, 19.07, 20.27, 22.4
	17	5.64 6.57, 8.61, 12.5, 14.44, 16.57, 18.7, 21.2, 22.59, 24.07
	18	4.72 8.14, 10.18, 12.4, 15.09, 17.22, 19.16, 21.57, 24.35, 26.2

TABLE 6.2: Measurement times under three temperatures.

Temperature	Measurement time epochs (in hours)
65°C	108, 241, 534, 839, 1074, 1350, 1637, 1890, 2178, 2513, 2810
85°C	46, 108, 212, 408, 632, 764, 1011, 1333, 1517, 2586
100°C	46, 108, 212, 344, 446, 626, 729, 279, 1005, 1218

is $s_0 = 40°\text{C}$, the highest allowable stress is $s_H = 100°\text{C}$ and the transformed stress is $x = (1/s_0 - 1/s)(1/s_0 - 1/s_H)$. In addition, $\Lambda(t) = t^\beta$ and $\mu(x) = \exp(\alpha_0 + \alpha_1 x)$. Given this data, the approximation in (6.12) performs well. We first use the simple IG process to fit the degradation data. The estimated parameters are $\hat{\alpha}_0 = -1.88, \hat{\alpha}_1 = 1.73, \hat{\lambda} = 0.653$ and $\hat{\beta} = 0.449$, respectively. The maximum log-likelihood value is -222. To test the goodness-of-fit, we compare the estimated mean degradation paths estimated by the simple IG process model and by directly averaging over the samples under each stress level, as shown in Figure 6.1. The estimated mean paths using the IG process tally with the empirical estimate by direct average, which indicates a goodness-of-fit. In addition, (Wang and Xu, 2010) proposed using the χ^2 Q-Q plot to test the goodness-of-fit. The Q-Q plot is given in Figure 6.2(a). As can be seen from the Q-Q plot, there seems to be one outlier in the data. Except for this outlier, the Q-Q plot tends to be satisfactory. If we fit the data by the random-volatility model, the parameter estimates are $\hat{\alpha}_0 = -2.29, \hat{\alpha}_1 = 1.87, \hat{\gamma} = 3.76, \hat{\delta} = 2.02, \hat{\beta} = 0.492$, with a maximum log-likelihood value of -200. The χ_1^2 Q-Q plot, as suggested in Wang and Xu (2010) is given in 6.2(b). From the maximum log-likelihood values, the random-effects model tends to provide a better fit. But from Figures 6.1 and 6.2 (b), we can see that the simple model is also acceptable. In the following, we will determine the optimal ADT plans based on both models.

Suppose 10 units are available for the ADT test, In the ADT, we set $\tau_j = 24$ and $K_j = 14$ for all $j = 1, \cdots, J$. This means that we measure the degradation level once every day and the test lasts two weeks. Our planning involves selecting the stress level $(x_1, x_2, ..., x_{J-1})$ and the proportion of samples allocated to each testing level, $(N_1, N_2, ..., N_{J-1})$. Consider a two-level ADT plan, i.e., $J = 2$. Suppose we are interested in minimizing the

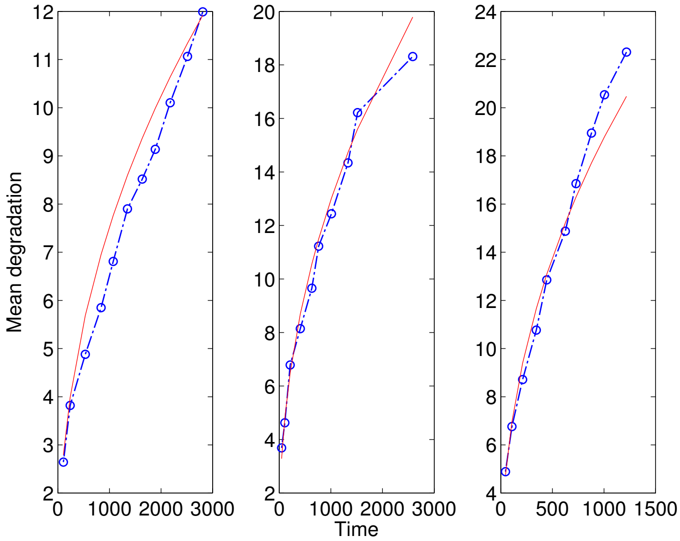


FIGURE 6.1: Estimated mean path under each stress level: 65° (left), 85° (middle), 100° (right). The dashed dotted line is based on direct average of the observed samples, and the solid line is the estimate based on the IG process.

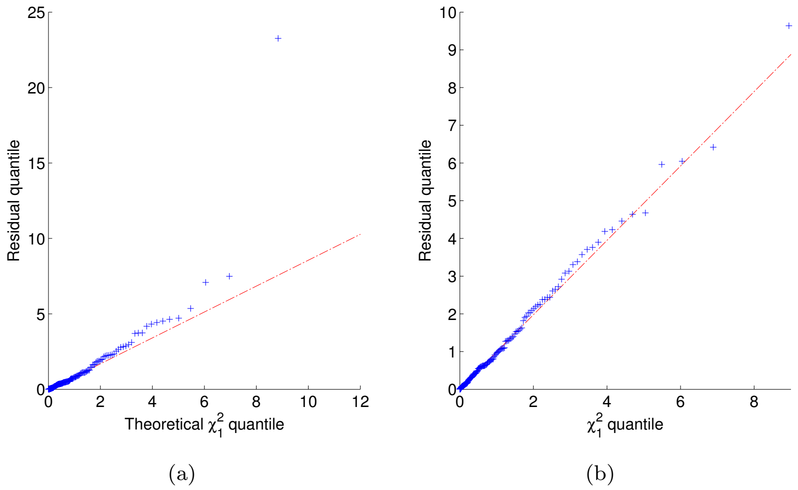


FIGURE 6.2: χ_1^2 Q-Q plot for the residuals fitted by the simple IG process.

asymptotic variance of B10, the 0.1-quantile of the failure time distribution at use conditions. Solving the optimization problem in (7.15) when $J = 2$ yields the optimal ADT design as shown in Table 6.3. It is interesting to observe that the optimal lower stress is 0. We suspect that this is because the degradation under the normal use conditions is fast enough so that the error caused by extrapolation to the failure threshold is small.

TABLE 6.3: Optimal two level ADT plan using IG process models.

	x_1	x_2	N_1	N_2	$\text{Std}(\hat{\xi}_p) \times 10^4$
Simple IG process	0	1	8	2	1.6604
IG process with random effects	0	1	2	8	0.374

We then investigate the deviation of the optimal ADT plan under the variation of parameters. We only consider the simple model, and the results of the random effects model are similar First, by varying α_1 from 1 to 4, the impacts of α_1 on the optimal ADT plan are presented in Tables 6.4. The optimal sample allocation differs when α_1 varies. The increase of the asymptotic variance of the life quantile is also observed when α_1 increases. On the other hand, the optimal stress level and sample allocation are insensitive to α_0 and λ, while the asymptotic variance of life quantile is greatly influenced by them. Figure 6.3 displays the natural logarithm of the minimum asymptotic standard deviation versus α_0 and λ. This figure shows that the optimal asymptotic variance is a decreasing function of λ.

TABLE 6.4: Optimal ADT plan with varying α_1 for the simple IG process.

α_1	x_1	x_2	N_1	N_2	$\text{Std}(\hat{\xi}_q(\boldsymbol{\theta})) \times 10^4$
1	0	1	9	1	1.5315
2	0	1	8	2	1.7223
3	0	1	7	3	2.0145
4	0	1	6	4	2.4506

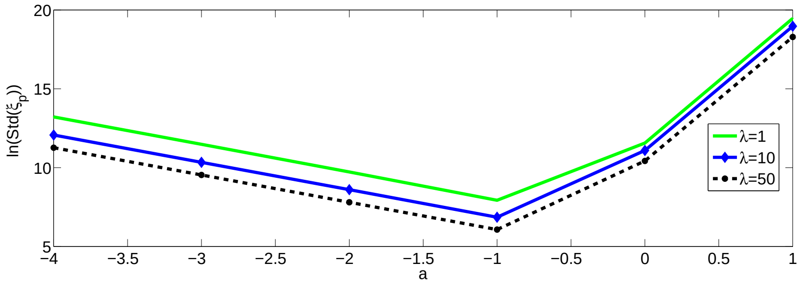


FIGURE 6.3: Minimized asymptotic standard deviation versus varying α_0 and λ.

The optimal plan depends on the true parameter values, i.e., $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \lambda, \beta)'$ for the simple model and $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \gamma, \delta, \beta)$ in the random effects model. The parameter values are unknown or subject to estimation errors at the test planning stage. So the sensitivity of the optimal ADT plan to the estimated parameters is studied. We only conduct the sensitivity analysis for the simple model, and the analysis for the random effects can be done in a similar vein. In particular, suppose the estimation bias for λ, α_0 and α_1 in the simple IG process are ϵ_1, ϵ_2 and ϵ_3 respectively. The optimal ADT plan is obtained under the various combinations of $(1+\epsilon_1)\lambda$, $(1+\epsilon_2)\alpha_0$ and $(1+\epsilon_3)\alpha_1$. It is found that overall speaking, the optimal ADT plan tends to be robust to estimation biases, given that the biases are not too large.

TABLE 6.5: Optimal ADT plan considering the estimation bias for λ, α_0 and α_1.

ϵ_1	ϵ_2	ϵ_3	x_1	x_2	N_1	N_2	$\text{Std}(\hat{\xi}_q(\boldsymbol{\theta})) \times 10^4$
+10%	+10%	+10%	0	1	8	2	2.2691
+10%	0	0	0	1	8	2	1.5903
+10%	-10%	-10%	0	1	8	2	1.1141
0	+10%	0	0	1	8	2	2.3235
0	0	-10%	0	1	8	2	1.6303
0	-10%	+10%	0	1	8	2	1.2146
-10%	+10%	-10%	0	1	8	2	2.3942
-10%	0	+10%	0	1	8	2	1.7865
-10%	-10%	0	0	1	8	2	1.2481
0	0	0	0	1	8	2	1.6635

6.5 Conclusion

This chapter has investigated the optimal constant-stress ADT plan based on the IG process. The objective is to minimize the asymptotic variance of the p-quantile under use conditions by properly specifying the stress levels and the allocation of test units to each level. Both the IG process with and without random effects were considered in this study. For both models, we assume that the degradation rate parameter μ is an increasing function of the stress. This is a legitimate assumption because degradation is most often hastened under severe working conditions. We applied the IG process model to fit the stress-relaxation data of a component and use the methods developed here to help with the ADT planning. The sensitivity analysis reveals that the optimal plan is quite robust to moderate departure of presumed model parameters.

This study has considered constant-stress ADT planning. An advantage of constant-stress ADT is that we can check the assumed stress-degradation relationship by separately estimating the parameters under each stress level. When the number of samples available for testing is extremely small, SSADT can be a good choice as long as it is believed that the underlying degradation is correct. MLEs of the model parameters and the asymptotic variance of the p-quantile ξ_p can be derived similar to the procedures presented in this paper. In addition to the random volatility model considered in this paper, Ye et al. (2012) further proposed two additional random effects models. The optimal ADT planning for these two models can be developed in a similar vein.

Chapter 7

ACCELERATED DEGRADATION TEST PLANNING CONSIDERING PRODUCT FIELD HETEROGENEITY

7.1 Introduction

As reviewed earlier, accelerated life test (ALT) is one of the most common approaches among the various accelerated test methods. Many ALT studies are devoted to planning optimum ALT and making statistical inferences. Their objectives are focused on the estimation precision of some critical reliability index (e.g., life quantile) when products are used under nominal and homogeneous conditions, see, e.g., Nelson (2005), Tang and Liu (2010). However, the field environment in which products actually operate is mostly complex and heterogeneous. For example, different customers may have different usage behaviours on the product. Therefore, when targeting product field performance, such as warranty costs, this discrepancy may bias the estimation and result in unacceptable outcome for the manufacturer. To overcome this deficiency, Liao and Elsayed (2006) generalized stress as a stochastic process and provided relevant statistical inference procedures. Recently Yang (2010) treated the stress (usage rate) as a normal random variable and studied a compromise ALT to predict warranty costs. However, these models result in intractable field lifetime distributions, which makes verification of the model assumptions very difficult.

In addition, the inference accuracy of ALT largely relies on unit failures. When the units are highly reliable, few failures may be observed within allowable testing time period. Nowadays an increasing number of highly reliable products are manufactured across a diverse set of industries. For example, a complex electronic system may consist of many components and maintaining high reliability for the entire system generally requires that the individual system components have extremely high reliability (Lu and Meeker, 1993). Within this context, ALT may be inaccurate or inefficient. On the other hand, although failures seldom occur, tested units may experience gradual degradation in terms of certain critical quality characteristic (Meeker and Escobar, 1998), which can be utilized to link with product failures and make useful predictions (Singpurwalla, 1995; Wang, 2010; Ye et al., 2011). Consequently, accelerated degradation test (ADT) has appeared recently and some studies can be found in Tang et al. (2004) and Tseng and Wen (2000), to name a few.

It is noted, however, that most existing ADT studies do not explicitly account for product reliability in the heterogeneous field conditions. Indeed, the bulk of research on ADT mainly focus on the investigation of candidate degradation models. Both general degradation path models (Meeker et al., 1998; Park and Yum, 1997) and stochastic process models (Tseng et al., 2009; Lim and Yum, 2011) are well implemented in stable and homogeneous conditions. When products in the field return to the manufacturer for repairs, the warranty claim data collected based on these returns suggests a higher variability in the product's failure times than the analysis based on accelerated tests. This gap is mainly attributed to the uncontrollable heterogeneous operating and use conditions in the field.

Although several studies have observed this deficiency (Pan, 2009; Meeker et al., 2009; Wang et al., 2012), few investigations are carried out on to cope with the ADT planning. This study investigates the ADT planning which utilizes both lab and filed information. In particular, we impose a tractable random-effect term on degradation modelling to explain the discrepancy caused by the heterogeneity in the field condition. Moreover, we focus on two objectives in determining an ADT plan, i.e., the asymptotic variance of the fraction failing and the life quantile. These two objectives are of particular interest and usefulness to manufacturers to determine product warranty.

The rest of this chapter is organized as follows. Section 7.2 presents the model formulation and statistical inference methods. Section 7.4 describes the optimization procedure. An application example is given in section 7.5 to illustrate the idea.

7.2 The model

In this study, we consider the situations in which the degradation/quality characteristic of the product is measurable in both lab test and the field. We assume that the underlying degradation of the product follows a gamma process, which is a commonly used model when the degradation path is monotone. Nevertheless, the practical operating environment is heterogeneous and affects the product degradation in the field. Therefore, a frailty term is imposed on unit's degradation to capture this uncontrollable effect, which naturally links the unit degradation measurement in lab and its actual degradation in the field. The detailed model development is investigated in this section.

7.2.1 Degradation in lab test

If the unit degradation is available and monotone, a gamma process is widely implemented to describe the time-dependent degradation increments, especially when the degradation is caused by intermittent external shocks. The degradation of a unit in lab test is assumed to follow a gamma process defined as follows.

Denote $Y(s)$ the measured degradation for an individual unit at time s, the gamma process is a stochastic process satisfying:

1. the increments $\Delta Y = Y(s + \Delta s) - Y(s)$ are independent;
2. ΔY has a gamma distribution $Ga\{\eta\Delta\Lambda, \xi^{-1}\}$, where $\Delta\Lambda = \Lambda(s + \Delta s) - \Lambda(s)$, and $\Lambda(s)$ is a given, monotone increasing function.

With the convention that $Y(0) = 0$ and $\Lambda(0) = 0$, the probability density of $Y(s)$ is $g(y; \eta, \xi^{-1}) = \Gamma(\eta\Lambda_s)^{-1}(\xi)^{\eta\Lambda_s} y^{\eta-1}\exp(-\xi y)$, where $\Lambda_s = \Lambda(s)$. Denote S the first passage time for a critical threshold D, then the CDF and PDF of S is give by

$$F_S(s) = P(S < s) = P(Y(s) > D) = 1 - \int_0^D g(y; \xi, \eta\Lambda_s)\, dy = \frac{\Gamma(\eta\Lambda_s, D\xi)}{\Gamma(\eta\Lambda_s)}, \tag{7.1}$$

where $\Gamma(p,q) = \int_q^\infty t^{p-1} e^{-t}\, dt$ is the incomplete gamma function.

$$f_S(s) = \frac{dF_S(s)}{ds} = \frac{\Gamma'(\eta\Lambda_s, D\xi)}{\Gamma(\eta\Lambda_s)} - \frac{\Gamma(\eta\Lambda_s, D\xi)}{\Gamma(\eta\Lambda_s)}\psi(\eta\Lambda_s). \tag{7.2}$$

where $\psi(\cdot)$ is the digamma function,

$\Gamma'(\eta\Lambda_s, D\xi) = \frac{d\Gamma(\eta\Lambda_s, D\xi)}{ds} = \eta\Lambda_s'(\ln(D\xi)\Gamma(\eta\Lambda_s, D\xi) + D\xi T(3, \eta\Lambda_s, D\xi))$, where $T(3, s, x)$ is the Meijer G-function, i.e., $T(m, s, x) = G_{m-1,m}^{m,0}\left({}_{s-1,-1,\cdots,-1}^{0,\cdots,0}|x\right)$.

The failure intensity function $h_S(s)$ is thus

$$h_S(s) = \frac{f_S(s)}{1 - F_S(s)} = \frac{\Gamma'(\eta\Lambda_s, D\xi)}{\gamma(\eta\Lambda_s, D\xi)} - \frac{\Gamma(\eta\Lambda_s, D\xi)}{\gamma(\eta\Lambda_s, D\xi)}\psi(\eta\Lambda_s). \tag{7.3}$$

7.2.2 Field Degradation with Random Effect

While lab test is conducted in a relatively stable environment, products in the field are usually exposed to a more complex environment where some uncontrollable factors take effects in unobservable ways, for example, the various usage rates across the customer population, the diverse geographical locations of products, etc. As a result, the perceived reliability of products in the field is often more varied than in the lab test. In literature, the random effect (or frailty) is proposed to account for the heterogeneities caused by the unobservable covariates whereby a random variable is imposed on the product hazard function (Meeker et al., 2009). For degradation modelling, however, this approach may result in intractability since the degradation model is much more complicated. Alternatively, the random-effect term can be imposed directly on the degradation path. In particular, ξ is replaced with ξz, where z is the random effect. This is legitimate since z affects both the mean and variance of gamma degradation process, which is capable to account for the heterogeneity of environment to a large extent. Note that this random-effect model is commonly used in degradation modelling and testing (Lawless and Crowder, 2004; Tsai et al., 2012). Here we implement this model to accommodate the discrepancy caused by the uncontrollable and heterogeneous field conditions.

With a random z, the tractability of the resulting lifetime distribution is of vital importance to validate the model with field data, especially when the data only consists of failure times. We find that when z belongs to the families of uniform, exponential or gamma, the field failure time distribution has a closed form. In this paper, z is assumed to follow a gamma distribution with parameter (k, θ). Marginalizing over z, $Y(t)$ has the CDF

(Lawless and Crowder, 2004)

$$P(Y(t) < d) = F_{2\eta\Lambda_t,2k}\left(\frac{kd}{\theta\xi\eta\Lambda_t}\right), \tag{7.4}$$

where $F_{x,y}$ is the distribution function of F-distribution with parameter (x, y).

The CDF of the first passage time T is (Tsai et al., 2012)

$$F_T(t) = P(T < t) = P(Y(t) > D) = 1 - F_{2\eta\Lambda_t,2k}\left(\frac{kD}{\theta\xi\eta\Lambda_t}\right) = \frac{B\left(\frac{D}{D+\theta\xi}; \eta\Lambda_t, k\right)}{B(\eta\Lambda_t, k)}, \tag{7.5}$$

where $B(x; \alpha_0, \alpha_1) = \int_x^1 s^{\alpha_0-1}(1-s)^{\alpha_1-1}\,ds$ is the upper incomplete beta function, and $B(\alpha_0, \alpha_1) = B(0; \alpha_0, \alpha_1)$ is the complete beta function.

Denote

$$\begin{aligned} M(x; \alpha_0, \alpha_1) = \quad & \tfrac{x^{\alpha_0}}{\alpha_0 B(\alpha_0,\alpha_1)}\Big[\Big(\psi(\alpha_0) - \psi(\alpha_0+\alpha_1) - \ln(x) + \tfrac{1}{\alpha_0}\Big) \\ & \cdot {}_2F_1(\{\alpha_0, 1-\alpha_1\}, \{1+\alpha_0\}; x) - \tfrac{x(1-\alpha_1)}{(1+\alpha_0)^2}\, {}_3F_2 \\ & \cdot(\{2-\alpha_1, 1+\alpha_0, 1+\alpha_0\}, \{2+\alpha_0, 2+\alpha_0\}; x)\Big] \end{aligned},$$

where ${}_mF_n$ is the confluent hypergeometric function defined by

${}_mF_n(\{a_1, \cdots, a_m\}, \{b_1, \cdots, b_n\}) = \sum_{i=0}^{\infty} \frac{(a_1)_i \cdots (a_m)_i}{(b_1)_i \cdots (b_n)_i} \frac{x^i}{i!}$, with Pochhammer symbol $(a)_i = \Gamma(a+i)/\Gamma(a)$, and $(a)_0 = 0$.

The PDF of T is

$$f_T(t) = \eta \Lambda_t' M \left(\frac{D}{D+\theta\xi}; \eta\Lambda_t, k \right). \tag{7.6}$$

Figure 7.1 shows that $f_T(t)$ exhibits different shapes when the parameters vary, indicating a flexibility of the random-effect model.

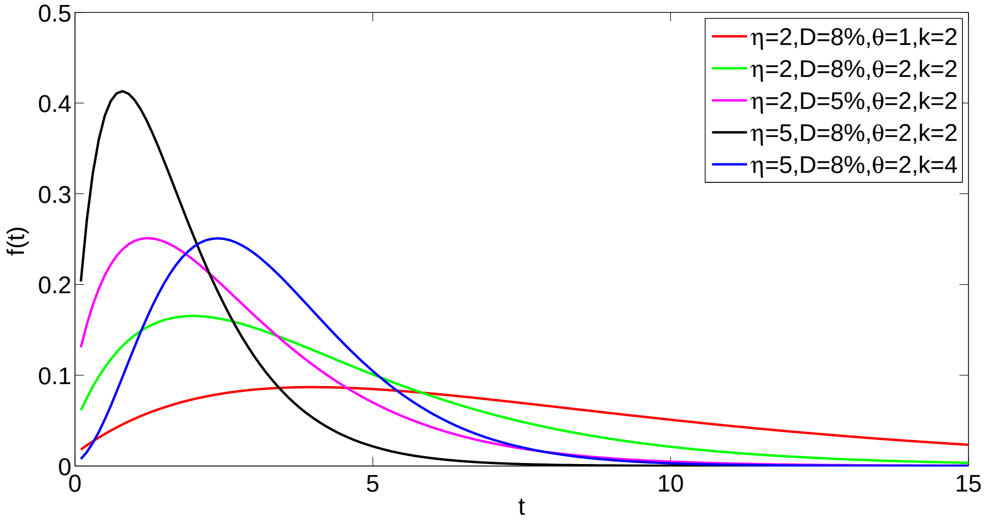


FIGURE 7.1: $f_T(t)$ under different parameter configurations.

7.3 Statistical inference

When the product is deployed to field operation, the data can be collected through a couple of ways, such as warranty returns, sensor devices embedded in products(Hong and Meeker, 2010), etc. Consequently, this field data can be analysed jointly with that obtained from lab test, as long as products in both lab and the field share the same inherent underlying degradation.

In this section, estimation procedures are developed when the field data contains degradation or failure times.

Suppose that P_1 units are put in the lab test, and the i-th unit is measured m_1 times at epochs $0 = t_{i1} < t_{i2} < \cdots < t_{im_1}$, $i = 1, \cdots, P_1$, with degradation measurements $X_i^1(t_{i1}), X_i^1(t_{i2}), \cdots, X_i^1(t_{im_1})$ respectively. Then the log-likelihood function of lab data, up to a constant, is

$$l(\eta_L, \xi|\text{Lab Data}) = \sum_{i=1}^{P_1}\sum_{j=1}^{m_1} (\eta\Lambda_{ij} - 1)\ \text{ln}x_{ij}^1 - \text{ln}\Gamma(\eta\Lambda_{ij}) + \eta\Lambda_{ij}\text{ln}\xi - x_{ij}^1\xi, \tag{7.7}$$

where η_L is η in lab test, $x_{ij}^1 = X_i^1(t_{ij}) - X_i^1(t_{i,j-1})$,$\Lambda_{ij} = \Lambda(t_{ij}) - \Lambda(t_{i,j-1})$.

7.3.1 Field degradation data

When the field information is incorporated, the obtained warranty data may be degradation data or only the failure time data. If it is possible to collect degradation data, and further assume that the field degradation data of P_2 units is recorded, and the i-th unit is measured at epochs $0 = t_{i1} < t_{j2} < \cdots < t_{im_2}$, with degradation $X_i^2(t_{i1}), X_i^2(t_{i2}), \cdots, X_i^2(t_{im_2})$. For simplicity and without loss of generality, we assume $t_{im_2} = t_{m_2}, i = 1, \cdots, P_2$. Using the frailty model in Section 2.2, the log-likelihood function of field degradation data is given by

$$\begin{aligned} l(\eta_F, k, \theta|\text{Field Data}) &= P_2 k \ln\theta + P_2 \ln(\Gamma(k+\eta\Lambda_{m_2})) + \textstyle\sum_{i=1}^{P_2}\sum_{j=1}^{m_2}(\Lambda_{ij}-1)(\ln x_{ij}^2 + k) \\ &- \textstyle\sum_{i=1}^{P_2}\sum_{j=1}^{m_2}\ln(\Gamma(\Lambda_{ij})) + P_2\ln(\Gamma(k)) - (\Lambda_{m_2}+\xi)\sum_{i=1}^{P_2}\ln(x_{im_2}^2+\theta). \end{aligned} \tag{7.8}$$

where η_F is η in the field, $x_{ij}^2 = X_i^2(t_{ij}) - X_i^2(t_{i,j-1})$, $\Lambda_{m_2} = \Lambda(t_{m_2})$.

7.3.2 Field life data

On the other hand, in some situations the measurement of product degradation cannot be implemented in the field, for example, when the data is collected through the warranty returns. Instead, failure time data is obtained. Denote the failure time $\mathbf{X}^3 = (x_1^3, x_2^3, \cdots, x_{m_3}^3)$, then $x_i^3, i = 1, \cdots, m_2$ has the PDF as in (7.6), and the log-likelihood function is

$$l(\eta_F, k, \theta|\text{Field Data}) = \sum_{i=1}^{m_3} \ln(\eta\Lambda'(x_i^3)) + \ln\left(M\left(\frac{D}{D+\theta\xi}; \eta\Lambda(x_i^3), k\right)\right). \tag{7.9}$$

The use of both lab and field data should be made of in statistical inference. Note that if the frailty model holds, the shape parameter η in lab and in the field are expected to be the same. This equality can be examined through likelihood ratio test or score test with null hypothesis $H_0 : \eta_L = \eta_F$ versus alternative hypothesis $H_a : \eta_L \neq \eta_F$. If H_0 is accepted, both the lab and field data can be utilized through the following total log-likelihood function

$$l(\eta, \xi, k, \theta|\text{All Data}) = l(\eta, \xi|\text{Lab Data}) + l(\eta, \xi, k, \theta|\text{Field Data}). \quad (7.10)$$

The MLE of parameters can be computed by either some direct search methods or the EM algorithm. The detailed procedure of EM algorithm is given in the appendix.

7.4 The ADT planning

Following the estimation of product reliability using both lab and field data, as well as other evaluations, some refinement procedures may be performed by the manufacturer and lead to a new design of products. Consequently, an ADT needs to be implemented to quickly obtain the reliability information of the new generation. It is most likely that new products do not inherit the same underlying degradation process. However, it is believed that the operating condition in the field is much less deviated for similar products. In other words, the new product has the same frailty Z as the old generation. With this fact, an optimal ADT plan can be developed.

Suppose the degradation of new products in lab test follows the gamma process specified in Section 2.1. Denote C_0 the nominal stress of new products, C_M the maximum stress level allowed, and C a variable level between C_0

and C_M. In line with the convention of ADT design using gamma process (e.g.Tseng et al. (2009)), we assume that the shape parameter η depends on the variable stress level,

$$\eta = \exp(r_0 + r_1\beta), \tag{7.11}$$

where β is the standardized stress given by

$$\begin{aligned} \beta &= \frac{\ln C - \ln C_0}{\ln C_M - \ln C_0} \quad \text{for the power law relation,} \\ &= \frac{1/C_0 - 1/C}{1/C_0 - 1/C_M} \quad \text{for the Arrhenius relation,} \\ &= \frac{C - C_0}{C_M - C_0} \quad \text{for the exponential relation.} \end{aligned}$$

Note that the standardized nominal stress and maximum stress is $\beta_0 = 0$ and $\beta_M = 1$ respectively.

7.4.1 The Fraction Failing

A two-level plan is considered in this ADT, and the variables to be optimally determined are the stress levels (β_1, β_2) and sample allocation (π_1, π_2). Usually the higher stress level is specified as the highest allowable level, i.e. $\beta_2 = \beta_M$. Besides, $\pi_1 + \pi_2 = 1$. Therefore, the plan involves determining β_1 and π_1. It is noted that this type of plan is uncovered in previous ADT studies using gamma process. In addition, when the heterogeneity of field operation is incorporated, the plan needs to be further revised.

One of the main objectives of ADT is to predict product reliability in the field. Since most products are under warranty, if the warranty period w is given, the manufacturer is usually concerned with the fraction failings during the warranty period. Denote p_w the failure probability within w, then from (7.5), it is obtained that $p_w = F_T(w)$.

The asymptotic variance of p_w can be obtained by the delta method, i.e. $AV(\hat{p}_w) = (\nabla p_w)'\mathcal{I}^{-1}\nabla p_w$, where ∇p_w is the first derivative of p_w with respect to (r_0, r_1, ξ) under the nominal stress, $\mathcal{I}^{-1}$ is the inverse of Fisher information matrix. The details of ∇p_w is given as follows.

$$
\begin{aligned}
\frac{\partial p_w}{\partial r_0} &= \Lambda_w M\left(\tfrac{D}{D+\theta\xi}; \exp(r_0)\Lambda_w, k\right), \\
\frac{\partial p_w}{\partial r_1} &= 0, \\
\frac{\partial p_w}{\partial \xi} &= \tfrac{D}{\theta^2\xi^2 B(k,\eta\Lambda_w)}\left(\tfrac{\theta\xi}{D+\theta\xi}\right)^{k+2} \\
&\quad \cdot \Big[\left(\tfrac{D+\theta\xi}{\theta\xi}\, {}_2F_1\left(\{k, 1-\eta\Lambda_w\}, \{1+k\}; \tfrac{\theta\xi}{D+\theta\xi}\right)\right) \\
&\quad + \tfrac{1-\eta\Lambda_w}{1+k}\, {}_2F_1\left(\{1+k, 2-\eta\Lambda_w\}, \{2+k\}; \tfrac{\theta\xi}{D+\theta\xi}\right)\Big]
\end{aligned}
\tag{7.12}
$$

The Fisher information is given by

$$\mathcal{I} = \begin{pmatrix} E\left[-\frac{\partial^2 l}{\partial r_0^2}\right] & E\left[-\frac{\partial^2 l}{\partial r_0 \partial r_1}\right] & E\left[-\frac{\partial^2 l}{\partial r_0 \partial \xi}\right] \\ & E\left[-\frac{\partial^2 l}{\partial r_1^2}\right] & E\left[-\frac{\partial^2 l}{\partial r_1 \partial \xi}\right] \\ & & E\left[-\frac{\partial^2 l}{\partial \xi^2}\right] \end{pmatrix}. \tag{7.13}$$

where l is the log-likelihood function conducted in lab ADT under various stress levels. In particular, suppose that the i-th unit under the j-th stress level is measured at $t_{ijk} = k\tau_j$ with observations $Y_{ij}(t_{ijk})$, $k = 0, 1, \cdots, K_j$. Let $y_{ijk} = Y_{ij}(t_{ijk}) - Y_{ij}(t_{ij,k-1})$ be the observed increments and $\Lambda_{ijk} = \Lambda(t_{ijk}) - \Lambda(t_{ij,k-1})$. Then the log-likelihood function, up to a constant, can be expressed as

$$l = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}(\eta_j\Lambda_{ijk} - 1)\ln y_{ijk} - \ln(\Gamma(\eta_j\Lambda_{ijk})) + \eta_j\Lambda_{ijk}\ln\xi - y_{ijk}\xi \tag{7.14}$$

where $\eta_j = \exp(r_0 + r_1\beta_j)$.

The detailed derivation of elements in $\mathcal{I}$ can be found in the appendix.

Based on the above results, the optimization problem of ADT planning can be formulated as,

$$\begin{aligned} &\textbf{Minimize} \quad \mathrm{AV}(\hat{p}_w(\beta_1, \pi_1)) \\ &\textbf{subject to} \quad 0 \leq \beta_1 < 1, \\ &\qquad\qquad\qquad 0 \leq \pi_1 < 1. \end{aligned} \tag{7.15}$$

7.4.2 The p-th life quantile

The life quantile of product reliability is another important quantity. It is used to help the manufacturer to determine the warranty length. Thus the precision of life quantile estimate is of interest in many ADT studies, where the objective is focused on minimizing the asymptotic variance (standard deviation) of p-th life quantile, i.e. t_p.

$$\mathrm{Avar}(\hat{t}_p) = \frac{\mathbf{g}_{\boldsymbol{\theta}}' \mathbf{I}^{-1}(\boldsymbol{\theta}) \mathbf{g}_{\boldsymbol{\theta}}}{N(f_{T_D}(t_p(\boldsymbol{\theta})))^2} \tag{7.16}$$

where N is the sample size of test, $\mathbf{g}_{\boldsymbol{\theta}}' = \left(\frac{\partial F_T(t)}{\partial \alpha_0}, \frac{\partial F_T(t)}{\partial \alpha_1}, \frac{\partial F_T(t)}{\partial \eta} \right)$. The detailed derivations are given in the appendix. The planning procedure is the same as (7.15).

7.5 Numerical example

7.5.1 Model goodness-of-fit and parameter estimation

The carbon film resistor in is usually subject to accelerated degradation test in lab during its design and production phase. Rather than failures, only the degradation of quality characteristic is observed during the allowable testing time. The degradation behaves in terms of the increase of its resistance, and the resistor fails when the increase exceeds a critical threshold. Consequently, inference procedures can be derived to estimate the product's reliability, which is addressed in previous studies, such as Park and Padgett (2005), Padgett and Tomlinson (2004).

We first check the adequacy of the gamma process model for the lab test data. The lab data is generated by simulation under a similar parameter configuration as in Tseng et al. (2009), which recently studied a cost-effective ADT plan for the carbon film resistor in the lab environment. In particular, the parameter values are

$$(r_0, r_1, \xi) = (-8.3907, 3.6622, 0.0625). \tag{7.17}$$

Note that this value is obtained after standardization based on the result in Tseng et al. (2009). The standardization is carried out as in (7.4) with $C_0 = 50°C$ and $C_M = 183°C$. Besides, $\Lambda(t) = t$. We then generate 10 samples of degradation using (7.17) as shown in Figure 7.2.

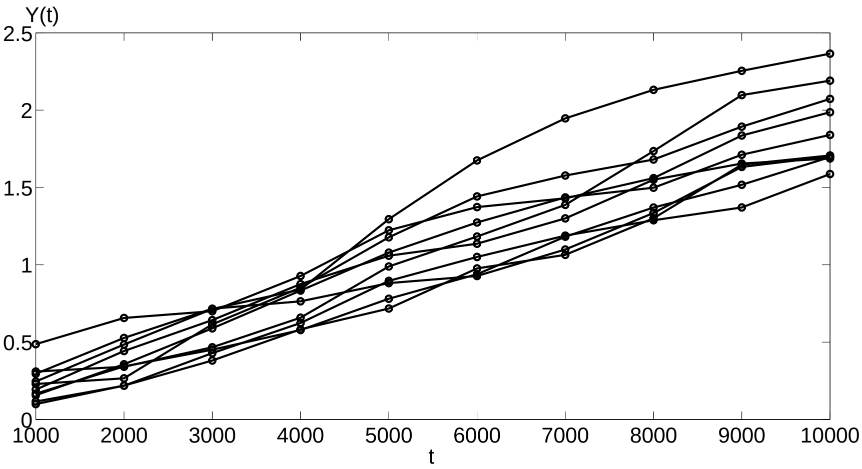


FIGURE 7.2: Simulated degradation paths of carbon film resistors.

The adequacy of gamma process model is examined using graphical quantile-quantile (Q-Q) plot. As mentioned in Wang and Xu (2010), for gamma process, $\sqrt[3]{\Delta x/(\xi\eta\Delta\Lambda)}$ is approximately normally distributed with mean $1-1/(9\eta\Delta\Lambda)$ and variance $1/(9\eta\Delta\Lambda)$. From the Q-Q plot depicted in Figure 7.3, we find that the gamma process fit the lab test data quite well.

On the other hand, when the carbon film resistor is put in field operation, its degradation is affected by the heterogeneous environment. The gamma process above may no longer provide a good description. As such, the frailty model proposed in this study may be a better choice. To illustrate this, we take the degradation data of carbon film resistor in Meeker and

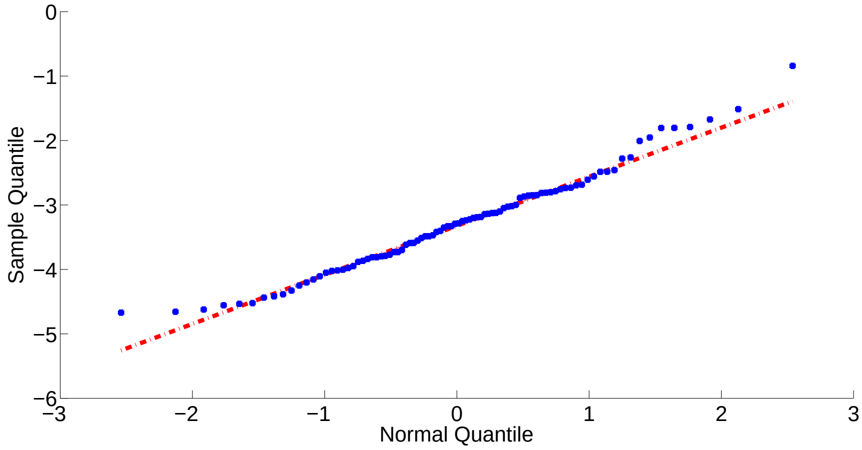


Figure 7.3: Q-Q plot for the simulated data versus the normal quantile.

Escobar (1998) as the field data and compare the fitness of gamma process and the frailty model. Specifically, the distribution functions of threshold-failure time based on both models are examined and fit to its nonparametric estimate which utilizes the pseudo failure times by extrapolating from the degradation data. The failure threshold D is assumed 8%. The comparison under 83°C and 133°C is shown in Figure 7.4. As can be seen, the frailty model appears a superior choice to the gamma process.

The estimated parameters of the frailty model under 83°C and 133°C are $\hat{r}_0 = -8.05, \hat{r}_1 = 3.94, \hat{k} = 12.9, \hat{\theta} = 0.449$. The maximum log-likelihood value is 49.68. When the gamma process model is fit to the data, the maximum log-likelihood value is 43.57. Alternatively, if we use the Akaike information criterion (AIC) to compare the two models, the respective AIC values for the gamma process model and the frailty model are -81.14 and -91.36. With a smaller AIC value, the frailty model has a better fit.

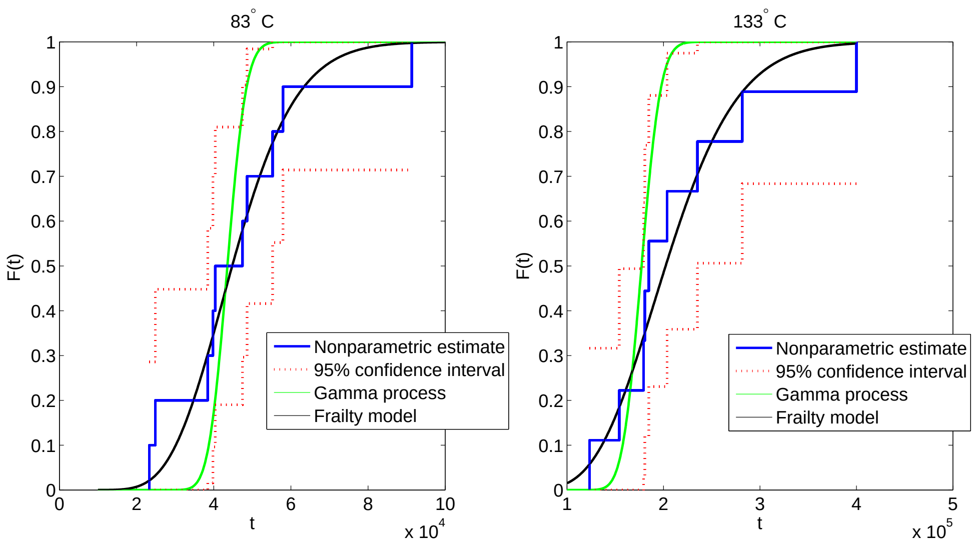
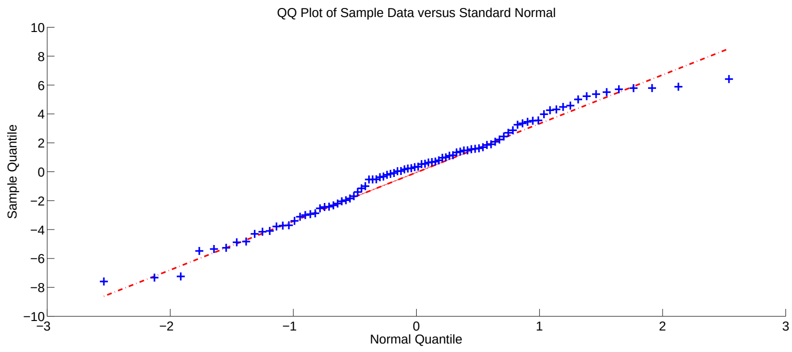
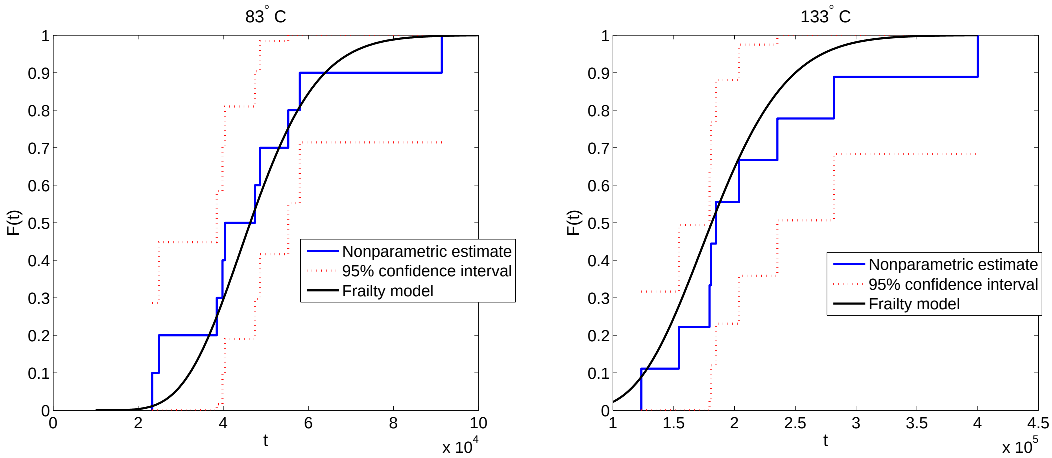


FIGURE 7.4: Comparison of distribution functions of threshold failure time under different models.

As seen from (7.17) and the estimated parameters for the frailty model, r_0 and r_1 are close respectively in both models, indicating the equality of η_L and η_F. The likelihood ratio method is used to test the null hypothesis $H_0 : \eta_L = \eta_F$ against the alternative hypothesis $H_a : \eta_L \neq \eta_F$. The likelihood ratio statistic value is 2.7286 with a p-value 0.2556. Therefore, there is no strong reason to reject the null hypothesis, and the frailty model is appropriate to account for the heterogeneity caused in operating environment. Consequently, both the lab and field data can be utilized as in (7.10). The estimation result is $\hat{r}_0 = -8.2388, \hat{r}_1 = 3.5599, \hat{k} = 15.713, \hat{\theta} = 0.8308, \hat{\xi} = 0.1554$. With these updated parameters, the fitness of model to the data in lab and the field can be checked and the result is depicted in Figure 7.5.



(a) Lab data



(b) Field data

FIGURE 7.5: Q-Q plot fit to the lab data and CDF fit to the field data using the updated parameters.

7.5.2 Optimal ADT planing

In this ADT planning, we focus on the new generation of products that comes out with a number of changes to the old generation by the manufacturer. However, when sold to the customer, the new generation should operate under the same field environment subject to similar heterogeneous effects. Consequently, we assume that the heterogeneity can still be modelled by the frailty Z which is gamma distributed with parameter $(k, \theta) =$

(15.713, 0.83). Suppose that the management is interested in knowing the fraction failings and one-dimensional warranty cost of products within a two year warranty. The available testing sample size is 20, and the measurement is performed in every 240 hours with totally 20 measurements for each unit. Since the units are tested in the lab, their degradation follows the gamma process under the nominal stress condition, and the planning values are assumed $r_0 = -8.24, r_1 = 3.56, \xi = 0.155$.

Under the above settings, the optimal ADT plan can be obtained which determines the sample allocation and the lower stress level. Specifically, if the objective is the fraction failings, the optimal test plan is $(\pi_1^*, \beta_1^*) = (0.5062, 0.1459)$. On the other hand, if the life quantile is of interest with $p = 10$, the optimal test plan is $(\pi_1^*, \beta_1^*) = (0.6328, 0.1197)$. Figure 7.6 shows the contour plot of the fraction failing with respect to π_1 and β_1.

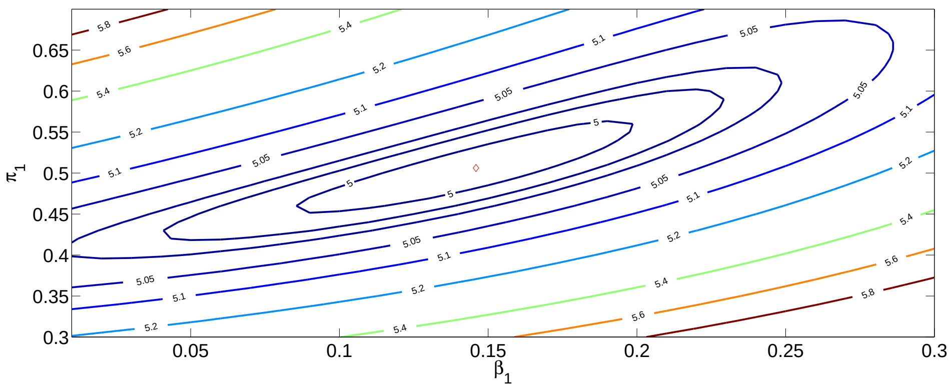


FIGURE 7.6: Contour plot of the asymptotic variance of fraction failings.

Chapter 8

CONCLUSION AND FUTURE WORK

Motivated by several practical problems, this thesis studies planning and optimization of maintenance and accelerated degradation test under complex failure processes. The proposed models facilitate some in-depth discussions for implementing the two important reliability programs in product's life cycle. In the following, we first summarize the main findings of this thesis. Then we conclude the thesis by elaborating some future research topics.

8.1 Main findings

In this thesis, we first investigate the failure behaviour of repairable systems and the subsequent maintenance planning in Chapter 3. In view of the non monotonic behaviour of system failure intensity in practice, we discuss the most commonly observed bathtub shape. In addition, we consider the reliability growth which may trigger a substantial change of failure intensity at system maintenance or other time epochs. To incorporate the heterogeneity within a fleet of systems, we introduce a random effect term imposed on failure intensity. Statistical inference procedures are addressed, and optimal maintenance planning is investigated. We then illustrate the proposed model using real dataset and reveal the model superiority in fitting some practical dataset by comparison with existing model in the literature.

In Chapter 4 we study a repairable manufacturing system and develop a bivariate maintenance policy. The manufacturing system adopts the economic production lot sizing principle and introduce new cost factors such as production set-up and inventory. The developed cost-effective maintenance policy aims at balancing a variety of costs incurred for production set-up, inventory and maintenance in the long run. Our model indicates a necessity of well planned maintenance in the real applications.

For some systems whose operation can be monitored, condition-based maintenance emerges as a more dynamic and adaptive technique compared with data driven or experience based approaches. Equipped with proper degradation model, Chapter 5 investigate the maintenance inspection policy for systems with multiple dependent failure modes. An alternative perspective is provided in modelling the dependency between failure modes. Then the system reliability is analysed followed by the cost-effective maintenance policy in which the inspection period is determined. The determination of inspection period is especially useful when the inspection cost is high. The regression nature of the proposed model enables its flexibility in fitting a number of dataset.

Chapter 6 studies the planning of accelerated degradation test using the inverse gaussian process model. With clear physical interpretation, the inverse gaussian process serves as an important class of models within the Lévy process family. When planning the test, we consider the heterogeneity that may be observed among test units by incorporating a random effect term. The optimal test stress levels are determined to minimize the asymptotic life quantile. The adequacy of model fitness to real dataset is illustrated with statistical test methods.

We then explore the accelerated degradation test considering field heterogeneity in Chapter 7. This study address the important application of accelerated degradation test to predict the field performance of products. We propose statistical inference methods that jointly utilizes both lab test and field use information. Then an optimization problem is formulated to minimize both the fractional failings and the p-th life quantile of the product. The proposed model is helpful for practical management decisions.

8.2 Future research topics

The current work can be extended in several directions:

1. For the non monotonic behaviour of failure intensity of repairable systems, Chapter 3 covers the bathtub type. Other types of non monotonicity can be investigated, such as upside down bathtub, roller-coaster type, etc. Dataset in real applications should be identified accordingly to illustrate the models.

2. As studied in Chapter 4, the maintenance is age based driven by statistical analysis of failure data. While the manufacturing system is usually safety critical, its real time health monitoring should be enabled. Then a cost-effective condition based maintenance can be implemented.

3. Chapter 5 determines the inspection frequency in maintenance planning given the degradation threshold and dependency of failure mode. An important extension would be to implement preventive maintenance wherein the system failure can be avoided to a large extent.

4. A constant stress ADT is planned in Chapter 6. When sample size of test units is relatively small, step stress ADT will be a good alternative provided the correct depiction of underlying degradation path.

5. As indicated in Chapter 7, the gap between in-lab reliability test and product dynamic operating environment needs to be investigated. So far we only consider a generic frailty model. Other issues such as new failure modes can also be incurred.

Overall, the future research direction on maintenance and degradation test is increasingly focused on data-driven modelling, with the enrichment of large-scale data acquisition and storage technologies (Meeker and Hong, 2014). Morever, the proposed models can be fully evaluated by the test data available.

Appendix A

Proofs of Lemma 4.1, Propoition 4.1, 4.2.

A.1 Proof of Lemma 4.1:

From Eq.(4.1),we obtain

$$
\begin{aligned}
HC_i &= \tfrac{hp(p-d)}{2d}\left(\int_0^{T_i} t^2 f_{Z_i}(t)\,dt + \int_{T_i}^{\infty} T_i^2 f_{Z_i}(t)\,dt\right)\\
&= \tfrac{hp(p-d)}{2d}\left(t^2 F_{Z_i}(t)|_0^{T_i} - \int_0^{T_i} 2tF_{Z_i}(t)\,dt + T_i^2 F_{Z_i}(t)|_{T_i}^{\infty}\right)\\
&= \tfrac{hp(p-d)}{2d}\left(T_i^2 - \int_0^{T_i} 2tF_{Z_i}(t)\,dt\right)\\
&= \tfrac{hp(p-d)}{d}\int_0^{T_i} t\left(1 - F_{Z_i}(t)\right)\,dt\\
&= \tfrac{hp(p-d)}{d}\int_0^{T_i} t\overline{F}_{Z_i}(t)\,dt\\
&= \tfrac{hp(p-d)}{d}\int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt
\end{aligned}
$$

Similarly, we have $DE = pc_d E[B] \sum_{i=1}^{N} \beta_i \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt$,

$MC = \sum_{i=1}^{N} \left[(c_r - c_p)\, F_{Z_1}\left(\alpha^{i-1}T_i\right) + c_p\right] + r$,

$CL = \frac{p}{d} \sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt$.

So when $N = 1$, AVC can be simplified as

$$AVC(T_1, 1) = \frac{\frac{hp(p-d)}{d} \int_0^{T_1} t\overline{F}_{Z_1}(t)\, dt + (c_r - c_p)\, F_{Z_1}(T) + k + c_p + r}{\frac{p}{d} \int_0^{T_1} \overline{F}_{Z_1}(t)\, dt} + pc_d d \beta_1 E[B] \tag{A.1}$$

and is finite.

When $N \to \infty$,

$$AVC(T, \infty) = \lim_{N\to\infty} \frac{1}{\frac{p}{d}\sum_{i=1}^{N}\int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt} \Bigg[\frac{hp(p-d)}{d} \sum_{i=1}^{N} \int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\, dt$$
$$+ pc_d E[B] \sum_{i=1}^{N} \beta_i \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt + \sum_{i=1}^{N} (c_r - c_p)\, F_{Z_1}\left(\alpha^{i-1}T_i\right) + N(k + c_p) + r \Bigg]$$

For the dominator, it is seen that $\int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt = 0$, for $\alpha > 1$, when i is sufficiently large. So $\frac{p}{d} \sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\, dt$ approaches some constant. But the numerator is greater than $N(k + c_p)$ which approaches infinity when $N \to \infty$ and $AVC(T, \infty) \to \infty$. Therefore, the optimal number of production runs N^* is finite, i.e. $N^* < \infty$.

For $\alpha < 1$, if $\beta_i, i = 1, 2, ..., N$ is nondecreasing,

$$AVC(T, \infty) > \lim_{N\to\infty} \frac{pc_d E[B] \sum_{i=1}^{N} \beta_i \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt}{\frac{p}{d}\sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt}$$
$$> \lim_{N\to\infty} \frac{pc_d E[B] \sum_{i=1}^{N} \beta_1 \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt}{\sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt} = pc_d E[B] N \beta_1 \to \infty$$

Therefore, N^* is also finite.

To prove the existence of optimal T_is, taking derivative of $AVC(\{T_i\}, N)$ with respect to T_i, one gets

$$\frac{\partial AVC(\{T_i\},N)}{\partial T_i} = \frac{1}{\frac{p}{d}CL^2(\{T_i\},N)} \Bigg\{ \Bigg[\frac{hp(p-d)}{d} T_i \overline{F}_{Z_1}(\alpha^{i-1}T_i) + (c_r - c_p)\,\alpha^{i-1} f_{Z_1}(\alpha^{i-1}T_i)$$
$$+pc_d\beta_i E[B]\overline{F}_{Z_1}(\alpha^{i-1}T_i)\Bigg] \sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt - \Bigg[\frac{hp(p-d)}{d} \sum_{i=1}^{N} \int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$$
$$+pc_d E[B] \sum_{i=1}^{N} \beta_i \int_0^{T_i} \overline{F}_{Z_i}(\alpha^{i-1}t)\,dt + \sum_{i=1}^{N} (c_r - c_p)\, F_{Z_1}(\alpha^{i-1}T_i) + N(k + c_p) + r\Bigg]$$
$$\bar{F}_{Z_1}(\alpha^{i-1}T_i)\Bigg\}$$

When $T_i \to 0$, the numerator approaches $-(Nc_r + Nk + r)$,because $\int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$ and $\int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$ both approach 0, for $1 \le i \le N$. Hence, we have $\frac{\partial AVC(\{T_i\},N)}{\partial T_i} \to \frac{-(Nc_r+Nk+r)}{\frac{p}{d}CL^2(\{T_i\},N)} < 0$.

When $T_i \to \infty$, both $f_{Z_1}(\alpha^{i-1}T_i)$ and $\overline{F}_{Z_1}(\alpha^{i-1}T_i)$ approach 0, for $1 \le i \le N$, so the numerator approaches

$$\frac{hp(p-d)}{d} T_i \overline{F}_{Z_1}(\alpha^{i-1}T_i) \sum_{i=1}^{N} \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$$
$$+ (c_r - c_p)\, f_{Z_1}(\alpha^{i-1}T_i) \int_0^{T_i} \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$$
$$- \frac{hp(p-d)}{d} \overline{F}_{Z_1}(\alpha^{i-1}T_i) \sum_{i=1}^{N} \int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt \overline{F}_{Z_1}(\alpha^{i-1}T_i)$$
$$= \frac{hp(p-d)}{d} \overline{F}_{Z_1}(\alpha^{i-1}T_i) \sum_{i=1}^{N} \int_0^{T_i} (T_i - t) \overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$$
$$+ (c_r - c_p)\, f_{Z_1}(\alpha^{i-1}T_i) \int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt$$

So $\frac{\partial AVC(\{T_i\},N)}{\partial T_i} > 0$,when $T_i \to \infty$.

Since the functions are continuous and differentiable, there exists $T_i = T_i^*, i = 1, 2, ..., N$, such that $\frac{\partial AVC(\left\{T_i^*\right\},N)}{\partial T_i^*} = 0$. Thus the proof is completed.$\square$

A.2 Proof of Proposition 4.1:

When $T_i = T_i^*, i = 1, 2, ..., N, \frac{\partial AVC(\{T_i\},N)}{\partial T_i}|_{T_i = T_i^*} = 0$.

$$\frac{\partial AVC(\{T_i\},N)}{\partial T_i} = \frac{\frac{\partial TC(\{T_i\},N)}{\partial T_i} CL(\{T_i\},N) - \frac{\partial CL(\{T_i\},N)}{\partial T_i} TC(\{T_i\}),N}{CL^2(\{T_i\},N)}$$
$$= \frac{\frac{\partial TC(\{T_i\},N)}{\partial T_i} - \frac{\partial CL(\{T_i\},N)}{\partial T_i} AVC(\{T_i\}),N)}{CL(\{T_i\},N)}$$

Since $CL(\{T_i\}, N) > 0$, we have

$$\frac{\partial TC(\{T_i\}, N)}{\partial T_i} - \frac{\partial CL(\{T_i\}, N)}{\partial T_i} AVC(\{T_i\}), N)|_{T_i = T_i^*} = 0, \text{for } i = 1, 2, ..., N \tag{A.7}$$

Substituting (4.2),(4.6),(4.7),(4.8),(4.9) into (A.7) and after simplifications, the following equation set is obtained,

$$\tfrac{hp(p-d)}{d} T_i + pc_d \beta_i E[B] + (c_r - c_p)\lambda_{Z_i}(T_i) = \tfrac{p}{d} AVC(\{T_i\}, N), \text{for } i = 1, 2, ..., N. \tag{A.8}$$

Note that the *RHS* of (A.8) is constant for $1 \leq i \leq N$.

When $\alpha > 1$, if $\beta_i, i = 1, 2, ..., N$ is nondecreasing, the *LHS* does not decrease when i increases because the first term is constant with i, the second term is nondecreasing function of i because $\beta_i, i = 1, 2, ..., N$ is nondecreasing and third term is nondecreasing functions of i, because from (4.3) and (4.4), $\lambda_{Z_i}(T_i) = \alpha^{i-1}\lambda_{Z_i}(\alpha^{i-1}T_i)$, when $\alpha > 1$, it is noncreasing function of i. On the other hand, the *LHS* is an increasing function of T_i. Therefore, to keep the value consistent to the *RHS*, when i increases, T_i must decrease.

When $\alpha < 1$, if $\beta_i, i = 1, 2, ..., N$ is nonincreasing, the *LHS* does not increase when i increases because the first term is constant with i, the second term is nonincreasing function of i because $\beta_i, i = 1, 2, ..., N$ is nonincreasing and third term is nonincreasing functions of i, because when $\alpha < 1$, $\lambda_{Z_i}(T_i) =$

$\alpha^{i-1}\lambda_{Z_i}(\alpha^{i-1}T_i)$ is nonincreasing function of i. On the other hand, the *LHS* is an increasing function of T_i. Therefore, to keep the value consistent to the *RHS*, when i increases, T_i must increase. Thus proof is thus completed.□

A.3 Proof of Proposition 4.2:

From the above derivations, it is obtained that

$$AVC(\{T_i\},N) = \frac{1}{\frac{p}{d}\sum_{i=1}^{N}\int_0^{T_i}\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt}\left[\frac{hp(p-d)}{d}\sum_{i=1}^{N}\int_0^{T_i} t\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt \right.$$
$$\left. +pc_dE[B]\sum_{i=1}^{N}\beta_i\int_0^{T_i}\overline{F}_{Z_i}(t)\,dt + \sum_{i=1}^{N}(c_r-c_p)\,F_{Z_1}(\alpha^{i-1}T_i) + N(k+c_p) + r\right] \tag{A.9}$$

Since $\overline{F}_{Z_1}(\alpha^{i-1}t) \leq 1$, for $i \leq i \leq N$, it is found that

$$AVC(\{T_i\},N) \leq \frac{\frac{hp(p-d)}{d}\sum_{i=1}^{N}\int_0^{T_i} t\,dt + pc_dE[B]\sum_{i=1}^{N}\beta_iT_i + \sum_{i=1}^{N}(c_r-c_p) + N(k+c_p) + r}{\frac{p}{d}\sum_{i=1}^{N}\int_0^{T_i}\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt}$$
$$= \frac{\frac{hp(p-d)}{2d}NT_i^2 + pc_dE[B]\sum_{i=1}^{N}\beta_iT_i + N(c_r+k) + r}{\frac{p}{d}\sum_{i=1}^{N}\int_0^{T_i}\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt} \tag{A.10}$$

The *RHS* of (A.11) is an increasing function of α.

On the other hand, since $\overline{F}_{Z_1}(\alpha^{i-1}t) \geq 0$,

$$AVC(\{T_i\},N) \geq \frac{N(c_p+k)+r}{\frac{p}{d}\sum_{i=1}^{N}\int_0^{T_i}\overline{F}_{Z_1}(\alpha^{i-1}t)\,dt} \tag{A.11}$$

The *RHS* of (A.11) is also an increasing function of α. Therefore, bounded by two increasing functions of α, the claim of Proposition 4.2 is obtained.□

Appendix B

Derivations of (6.9),(6.21) and (6.24)

The elements of $\mathbf{I}(\boldsymbol{\theta})$ in (6.9) are derived as follows,

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2}\right] = \frac{\sum_{j=1}^{J} N_j K_j}{2\lambda^2}, \tag{B.1}$$

Since $E[y_{ijk}] = \mu\Lambda_{ijk} = e^{\alpha_0+\alpha_1 x_j}\Lambda_{ijk}$,

$$\begin{aligned} E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0^2}\right] &= \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \lambda e^{-\alpha_0-\alpha_1 x_j}(2E[y_{ijk}]e^{-\alpha_0-\alpha_1 x_j} - \Lambda_{ijk}) \\ &= \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \lambda e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk}, \end{aligned} \tag{B.2}$$

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_1^2}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\lambda x_j^2 e^{-\alpha_0-\alpha_1 x_j}(2E[y_{ijk}]e^{-\alpha_0-\alpha_1 x_j} - \Lambda_{ijk}) = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\lambda x_j^2 e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk}, \tag{B.3}$$

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\alpha_0}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} e^{-\alpha_0-\alpha_1 x_j}(\Lambda_{ijk} - E[y_{ijk}]e^{-\alpha_0-\alpha_1 x_j}) = 0, \tag{B.4}$$

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\alpha_1}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} x_j e^{-\alpha_0-\alpha_1 x_j}(\Lambda_{ijk} - E[y_{ijk}]e^{-\alpha_0-\alpha_1 x_j}) = 0, \tag{B.5}$$

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0\partial\alpha_1}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\lambda x_j e^{-\alpha_0-\alpha_1 x_j}(2E[y_{ijk}]e^{-\alpha_0-\alpha_1 x_j} - \Lambda_{ijk}) = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\lambda x_j e^{-\alpha_0-\alpha_1 x_j}\Lambda_{ijk}, \tag{B.6}$$

When $\Lambda(t) = t^\beta$, $\Lambda_{ijk} = t_{ijk}^\beta - t_{ij,k-1}^\beta = (k\tau)^\beta - ((k-1)\tau)^\beta$. Then

$$\frac{\partial\Lambda_{ijk}}{\partial\beta} = \beta\left[(k\tau)^{\beta-1} - ((k-1)\tau)^{\beta-1}\right], \tag{B.7}$$

and

$$\frac{\partial^2\Lambda_{ijk}}{\partial\beta^2} = \left[(k\tau)^{\beta-1} - ((k-1)\tau)^{\beta-1}\right] + \beta(\beta-1)\left[\left((k\tau)^{\beta-2} - ((k-1)\tau)^{\beta-2}\right)\right]. \tag{B.8}$$

Therefore,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\partial \Lambda_{ijk}}{\partial \beta}\left(\frac{1}{\Lambda_{ijk}} + \frac{\lambda}{y_{ijk}}(y_{ijk}e^{-\alpha_0-\alpha_0 x_i} - \Lambda_{ijk})\right), \quad \text{(B.9)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \lambda} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\partial \Lambda_{ijk}}{\partial \beta}\frac{y_{ijk}e^{-\alpha_0-\alpha_1 x_j} - \Lambda_{ijk}}{y_{ijk}}. \quad \text{(B.10)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha_0} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\partial \Lambda_{ijk}}{\partial \beta}(-\lambda e^{-\alpha_0-\alpha_1 x_j}). \quad \text{(B.11)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha_1} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\partial \Lambda_{ijk}}{\partial \beta}(-\lambda x_j e^{-\alpha_0-\alpha_1 x_j}). \quad \text{(B.12)}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta^2} \quad & = \textstyle\sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} -\frac{1}{\Lambda_{ijk}^2}\left(\frac{\partial \Lambda_{ijk}}{\partial \beta}\right)^2 + \frac{1}{\Lambda_{ijk}}\frac{\partial^2 \Lambda_{ijk}}{\partial \beta^2} + \lambda \\ e^{-\alpha_0-\alpha_1 x_j}\frac{\partial^2 \Lambda_{ijk}}{\partial \beta^2} \quad & -\frac{\lambda}{y_{ijk}}\left[\left(\frac{\partial \Lambda_{ijk}}{\partial \beta}\right)^2 + \Lambda_{ijk}\frac{\partial^2 \Lambda_{ijk}}{\partial \beta^2}\right]\end{aligned} \quad \text{(B.13)}$$

Taking expectation yields,

$$\begin{aligned}E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \beta}\right] & = E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \lambda}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} -\frac{\partial \Lambda_{ijk}}{\partial \beta}\left(e^{-\alpha_0-\alpha_1 x_j} - \Lambda_{ijk}E\left(\frac{1}{y_{ijk}}\right)\right) \\ & = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\partial \Lambda_{ijk}}{\partial \beta}\frac{1}{\lambda \Lambda_{ijk}}.\end{aligned} \quad \text{(B.14)}$$

where we substitute into the fact that $E\left(\frac{1}{y_{ijk}}\right) = \frac{1}{\lambda\Lambda_{ijk}^2} + \frac{1}{\mu\Lambda_{ijk}}$.

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0\partial\beta}\right] = E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta\partial\alpha_0}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\frac{\partial\Lambda_{ijk}}{\partial\beta}\lambda e^{-\alpha_0-\alpha_1 x_j}. \quad \text{(B.15)}$$

$$E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha_0\partial\beta}\right] = E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta\partial\alpha_1}\right] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\frac{\partial\Lambda_{ijk}}{\partial\beta}\lambda x_j e^{-\alpha_0-\alpha_1 x_j}. \quad \text{(B.16)}$$

$$\begin{aligned} E\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta^2}\right] &= \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\frac{1}{\Lambda_{ijk}^2}\left(\frac{\partial\Lambda_{ijk}}{\partial\beta}\right)^2 - \frac{1}{\Lambda_{ijk}}\frac{\partial^2\Lambda_{ijk}}{\partial\beta^2} - \lambda e^{-\alpha_0-\alpha_1 x_j}\frac{\partial^2\Lambda_{ijk}}{\partial\beta^2} \\ &+\lambda\left(\frac{1}{\lambda\Lambda_{ijk}^2} + \frac{\exp(-\alpha_0-\alpha_1 x_j)}{\Lambda_{ijk}}\right)\left[\left(\frac{\partial\Lambda_{ijk}}{\partial\beta}\right)^2 + \Lambda_{ijk}\frac{\partial^2\Lambda_{ijk}}{\partial\beta^2}\right] \\ &= \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\left(\frac{2}{\Lambda_{ijk}^2} + \frac{\lambda\exp(-\alpha_0-\alpha_1 x_j)}{\Lambda_{ijk}}\right)\left(\frac{\partial\Lambda_{ijk}}{\partial\beta}\right)^2 \end{aligned} \quad \text{(B.17)}$$

For the random volatility model, the derivations of elements $\mathbf{I}(\boldsymbol{\theta})$ in (6.21) is as below.

The log-likelihood function in (6.18) is written below,

$$l(\boldsymbol{\theta}) = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\left\{\ln\Gamma(\delta+\frac{K_j}{2}) - \ln\Gamma(\delta) + \delta\ln\gamma + \sum_{k=1}^{K_j}\left(\ln\Lambda_{ijk} - \frac{3}{2}\ln y_{ijk}\right)\right.$$
$$\left. -(\delta+\frac{K_j}{2})\ln\left[\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right]\right\} \tag{B.18}$$

where $\mu_j = \exp(\alpha_0 + \alpha_1 x_j)$.

The first derivatives of $l(\boldsymbol{\theta})$ with respect to the parameters are obtained as follows,

First, note that the following equations hold for α_0 and α_1,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial\alpha_0} = \sum_{j=1}^{J}\exp(\alpha_0+\alpha_1 x_j)\frac{\partial l(\boldsymbol{\theta})}{\partial\mu_j}, \tag{B.19}$$

and

$$\frac{\partial l(\boldsymbol{\theta})}{\partial\alpha_1} = \sum_{j=1}^{J}x_j\exp(\alpha_0+\alpha_1 x_j)\frac{\partial l(\boldsymbol{\theta})}{\partial\mu_j} \tag{B.20}$$

where

$$\frac{\partial l(\boldsymbol{\theta})}{\partial\mu_j} = \sum_{i=1}^{N_j}(\delta+\frac{K_j}{2})\frac{\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{2\mu_j^2}}{\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}}. \tag{B.21}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial\delta} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\psi(\delta+\frac{K_j}{2}) - \psi(\delta) + \ln\gamma - \ln\left[\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right]. \tag{B.22}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\frac{\delta}{\gamma} - (\delta + \frac{K_j}{2})\frac{1}{\left[\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right]}. \tag{B.23}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\left(\sum_{k=1}^{K_j}\frac{1}{\Lambda_{ijk}}\frac{\partial \Lambda_{ijk}}{\partial \beta} + (\delta + \frac{K_j}{2})\frac{\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{y_{ijk}}\frac{\partial \Lambda_{ijk}}{\partial \beta}}{\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}}\right). \tag{B.24}$$

where $\frac{\partial \Lambda_{ijk}}{\partial \beta}$ is as derived in (**??**).

The second derivatives of $l(\boldsymbol{\theta})$ are now readily derived as follows,

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j \partial \beta} = \sum_{i=1}^{N_j}(\delta + \frac{K_j}{2})\frac{\sum_{k=1}^{K_j} -\frac{1}{2\mu_j^2}\frac{\partial \Lambda_{ijk}}{\partial \beta}\left(\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right)}{\left(\gamma + \sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right)^2}$$

$$\frac{-\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{2\mu_j^2}\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{y_{ijk}}\frac{\partial \Lambda_{ijk}}{\partial \beta}}{} \tag{B.25}$$

Therefore,

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \beta} = \sum_{j=1}^{J}\exp(\alpha_0 + \alpha_1 x_j)\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j \partial \beta}, \tag{B.26}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta} = \sum_{j=1}^{J}x_j\exp(\alpha_0 + \alpha_1 x_j)\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j \partial \beta}, \tag{B.27}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j^2} =$$

$$\sum_{i=1}^{N_j}(\delta+\frac{K_j}{2})\frac{\sum_{k=1}^{K_j}\frac{y_{ijk}-2\mu_j\Lambda_{ijk}}{2\mu_j^4}\left(\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right)-\left(\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{2\mu_j^2}\right)^2}{\left(\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right)^2}. \tag{B.28}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0^2} = \sum_{j=1}^{J}\exp(\alpha_0+\alpha_1 x_j)(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu_j}+\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j^2}) \tag{B.29}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_0 \partial \alpha_1} = \sum_{j=1}^{J}x_j\exp(\alpha_0+\alpha_1 x_j)(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu_j}+\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j^2}) \tag{B.30}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1^2} = \sum_{j=1}^{J}x_j^2\exp(\alpha_0+\alpha_1 x_j)(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu_j}+\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_j^2}) \tag{B.31}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta^2} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\psi_1(\delta+\frac{K_j}{2})-\psi_1(\delta). \tag{B.32}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma^2} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}-\frac{\delta}{\gamma^2}+(\delta+\frac{K_j}{2})\frac{1}{\left[\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}\right]^2}. \tag{B.33}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta \partial \mu_j} = \sum_{i=1}^{N_j}-\frac{\sum_{k=1}^{K_j}\frac{y_{ijk}\mu_j^{-1}-\Lambda_{ijk}}{2\mu_j^2}}{\gamma+\sum_{k=1}^{K_j}\frac{(y_{ijk}\mu_j^{-1}-\Lambda_{ijk})^2}{2y_{ijk}}}. \tag{B.34}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\delta\partial\alpha_0} = \sum_{j=1}^{J} \exp(\alpha_0 + \alpha_1 x_j) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\delta\partial\mu_j} \tag{B.35}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\delta\partial\alpha_1} = \sum_{j=1}^{J} x_j \exp(\alpha_0 + \alpha_1 x_j) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\delta\partial\mu_j} \tag{B.36}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\delta\partial\gamma} = \sum_{j=1}^{J} \sum_{i=1}^{N_j} \frac{1}{\gamma} - \frac{1}{\left[\gamma + \sum_{k=1}^{K_j} \frac{(y_{ijk}\mu_j^{-1} - \Lambda_{ijk})^2}{2y_{ijk}}\right]}. \tag{B.37}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma\partial\mu_j} = \sum_{i=1}^{N_j} (\delta + \frac{K_j}{2}) \frac{\sum_{k=1}^{K_j} \frac{y_{ijk}\mu_j^{-1} - \Lambda_{ijk}}{2\mu_j^2}}{\left(\gamma + \sum_{k=1}^{K_j} \frac{(y_{ijk}\mu_j^{-1} - \Lambda_{ijk})^2}{2y_{ijk}}\right)^2} \tag{B.38}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma\partial\alpha_0} = \sum_{j=1}^{J} \exp(\alpha_0 + \alpha_1 x_j) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma\partial\mu_j} \tag{B.39}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma\partial\alpha_1} = \sum_{j=1}^{J} x_j \exp(\alpha_0 + \alpha_1 x_j) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma\partial\mu_j} \tag{B.40}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta^2} &= \sum_{j=1}^{J} \sum_{i=1}^{N_j} \left(\sum_{k=1}^{K_j} -\frac{1}{\Lambda_{ijk}^2} \left(\frac{\partial\Lambda_{ijk}}{\partial\beta}\right)^2 + \frac{1}{\Lambda_{ijk}} \frac{\partial^2\Lambda_{ijk}}{\partial\beta^2} \right. \\
&+ (\delta + \frac{K_j}{2}) \frac{\sum_{k=1}^{K_j} \left(\frac{-1}{y_{ijk}} (\frac{\partial\Lambda_{ijk}}{\partial\beta})^2 + \frac{y_{ijk}\mu_j^{-1} - \Lambda_{ijk}}{y_{ijk}} \frac{\partial^2\Lambda_{ijk}}{\partial\beta^2} \right)}{\gamma + \sum_{k=1}^{K_j} \frac{(y_{ijk}\mu_j^{-1} - \Lambda_{ijk})^2}{2y_{ijk}}} \\
&\left. - \frac{\left(\sum_{k=1}^{K_j} \frac{y_{ijk}\mu_j^{-1} - \Lambda_{ijk}}{y_{ijk}} \frac{\partial\Lambda_{ijk}}{\partial\beta} \right)^2}{\left(\gamma + \sum_{k=1}^{K_j} \frac{(y_{ijk}\mu_j^{-1} - \Lambda_{ijk})^2}{2y_{ijk}} \right)^2} \right).
\end{aligned} \tag{B.41}$$

Due to the complex form of the above second derivatives, their expectations can be evaluated numerically by using the RS sums method proposed in (?).

The detailed expressions of $\nabla\xi_p$ in (6.24) are derived as follows.

$$\frac{\partial \xi_p}{\partial \alpha_0} = \frac{1}{2\beta}\left(\frac{1}{4}\left(\sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}}t_{2\delta,p} + \sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}}\right)\right)^{2/\beta-1} \cdot \left[\frac{1}{2}\sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}}t_{2\delta,p} + \frac{1}{2}\left(\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}\right)^{-1/2} \left(\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 - \frac{4D}{\exp(\alpha_0)}\right) \tag{B.42}$$

$$\frac{\partial \xi_p}{\partial \alpha_1} = 0 \tag{B.43}$$

$$\frac{\partial \xi_p}{\partial \delta} = \frac{1}{2\beta}\left(\frac{1}{4}\left(\sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}}t_{2\delta,p} + \sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}}\right)\right)^{2/\beta-1} \cdot \left[-\frac{1}{2}\sqrt{\frac{\exp(\alpha_0)\gamma}{\delta^{3/2}}}t_{2\delta,p} + \frac{1}{2}\left(\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}\right)^{-1/2} \left(-\frac{\exp(\alpha_0)\gamma}{\delta^2}t_{2\delta,p}^2\right)$$

(B.45)

(B.46)

$$\begin{aligned}\frac{\partial \xi_p}{\partial \gamma} &= \frac{1}{2\beta}\left(\frac{1}{4}\left(\sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}}t_{2\delta,p} + \sqrt{\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}}\right)\right)^{2/\beta-1} \\ &\cdot \left[\frac{1}{2}\sqrt{\frac{\exp(\alpha_0)}{\delta\gamma}}Ft_{2\delta,p} + \frac{1}{2}\left(\frac{\exp(\alpha_0)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(\alpha_0)}\right)^{-1/2}\frac{\exp(\alpha_0)}{\delta}t_{2\delta,p}^2\right]\end{aligned} \tag{B.47}$$

$$\begin{aligned}\frac{\partial \xi_p}{\partial \beta} &= -\frac{2}{\beta^2}\ln\left(\frac{1}{4}\left(\sqrt{\frac{\exp(a)\gamma}{\delta}}t_{2\delta,p} + \sqrt{\frac{\exp(a)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(a)}}\right)\right) \\ &\cdot \left(\frac{1}{4}\left(\sqrt{\frac{\exp(a)\gamma}{\delta}}t_{2\delta,p} + \sqrt{\frac{\exp(a)\gamma}{\delta}t_{2\delta,p}^2 + \frac{4D}{\exp(a)}}\right)\right)^{2/\beta}\end{aligned} \tag{B.48}$$

Appendix C

Derivations of elements in (7.13) and statistical inference using EM algorithm

The derivation of elements in (7.13) is as follows,

$$\frac{\partial l}{\partial r_0} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \eta_j\Lambda_{ijk}\text{ln}y_{ijk} - \eta_j\Lambda_{ijk}\psi_0(\eta_j\Lambda_{ijk}) + \eta_j\Lambda_{ijk}\text{ln}\xi,$$

$$\frac{\partial l}{\partial r_1} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \beta_j\eta_j\Lambda_{ijk}\text{ln}y_{ijk} - \beta_j\eta_j\Lambda_{ijk}\psi_0(\eta_j\Lambda_{ijk}) + \beta_j\eta_j\Lambda_{ijk}\text{ln}\xi,$$

$$\frac{\partial l}{\partial \xi} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\eta_j\Lambda_{ijk}}{\xi} - y_{ijk},$$

$$\frac{\partial^2 l}{\partial r_0^2} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \eta_j\Lambda_{ijk}\ln y_{ijk} - (\eta_j\Lambda_{ijk})^2\psi_1(\eta_j\Lambda_{ijk}) - \eta_j\Lambda_{ijk}\psi_0(\eta_j\Lambda_{ijk}) + \eta_j\Lambda_{ijk}\ln\xi,$$

$$\frac{\partial^2 l}{\partial r_0 \partial r_1} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \beta_j\eta_j\Lambda_{ijk}\ln y_{ijk} - \beta_j(\eta_j\Lambda_{ijk})^2\psi_1(\eta_j\Lambda_{ijk}) - \beta_j\eta_j\Lambda_{ijk}\psi_0(\eta_j\Lambda_{ijk}) + \beta_j\eta_j\Lambda_{ijk}\ln\xi,$$

$$\frac{\partial^2 l}{\partial r_0 \partial \xi} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\eta_j\Lambda_{ijk}}{\xi},$$

$$\frac{\partial^2 l}{\partial r_1^2} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \beta_j^2\eta_j\Lambda_{ijk}\ln y_{ijk} - \beta_j^2(\eta_j\Lambda_{ijk})^2\psi_1(\eta_j\Lambda_{ijk}) - \beta_j^2\eta_j\Lambda_{ijk}\psi_0(\eta_j\Lambda_{ijk}) + \beta_j^2\eta_j\Lambda_{ijk}\ln\xi,$$

$$\frac{\partial^2 l}{\partial r_1 \partial \xi} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\beta_j\eta_j\Lambda_{ijk}}{\xi},$$

$$\frac{\partial^2 l}{\partial \xi^2} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} -\frac{\eta_j\Lambda_{ijk}}{\xi^2}.$$

where $\psi_0(\cdot)$ is the digamma function, $\psi_1(\cdot)$ is the trigamma function.

Since $E(y_{ijk}) = \eta_j\Lambda_{ijk}/\xi$, $E(\ln(y_{ijk})) = \psi_0(\eta_j\Lambda_{ijk}) - \ln\xi$, it is obtained that

$$E[-\frac{\partial^2 l}{\partial r_0^2}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} (\eta_j\Lambda_{ijk})^2\psi_1(\eta_j\Lambda_{ijk}),$$

$$E[-\frac{\partial^2 l}{\partial r_0 \partial r_1}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \beta_j(\eta_j\Lambda_{ijk})^2\psi_1(\eta_j\Lambda_{ijk}),$$

$$E[-\frac{\partial^2 l}{\partial r_0 \partial \xi}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} -\frac{\eta_j \Lambda_{ijk}}{\xi},$$

$$E[-\frac{\partial^2 l}{\partial r_1^2}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \beta_j^2 (\eta_j \Lambda_{ijk})^2 \psi_1(\eta_j \Lambda_{ijk}),$$

$$E[-\frac{\partial^2 l}{\partial r_1 \partial \xi}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} -\frac{\beta_j \eta_j \Lambda_{ijk}}{\xi},$$

$$E[-\frac{\partial^2 l}{\partial \xi^2}] = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j} \frac{\eta_j \Lambda_{ijk}}{\xi^2}.$$

The EM algorithm procedure is given below,

When the degradation data of J stress levels with N_j units at each level are available, the log-likelihood function, up to a constant, can be derived as,

$$l_r = \prod_{j=1}^{J}\prod_{i=1}^{N_j}\left[\prod_{k=1}^{K_j} \frac{y_{ijk}^{\eta\Lambda_{ijk}-1}\xi^{\eta\Lambda_{ijk}}}{\Gamma(\eta\Lambda_{ijk})} \cdot \frac{\Gamma(\eta\sum_{k=1}^{K_j}\Lambda_{ijk}+k)\theta^k}{(\xi\sum_{k=1}^{K_j} y_{ijk}+\theta)^{(\eta\sum_{k=1}^{K_j}\Lambda_{ijk}+k)}\Gamma(k)}\right].$$

Tsai et al. (2012) has developed the EM algorithm when the stress level is identical. In the presence of more than one stress level, the EM algorithm can be developed as follows.

Denote ξ_{ij} the realization of the frailty for the i-th unit under the j-th stress level, $i = 1, 2, \cdots, N_j$, and $j = 1, 2, \cdots, J$. Given the observed degradation data $\mathbf{Y}$ and the frailty ξ_{ij}, the complete data log-likelihood can be expressed as

$$
\begin{aligned}
l_c = \; & \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}[(\eta_j\Lambda_{ijk}-1)\ln y_{ijk} - \ln(\Gamma(\eta_j\Lambda_{ijk})) + \eta_j\Lambda_{ijk}\ln\xi_{ij} - y_{ijk}\xi_{ij}] \\
& + \sum_{j=1}^{J}\sum_{i=1}^{N_j}[k\ln\theta + (k-1)\ln\xi_{ij} - \ln\Gamma(k) - \theta\xi_{ij}]
\end{aligned}
$$

E-step:

To compute $E[\xi_{ij}|\mathbf{Y}]$ and $E[\ln\xi_{ij}|\mathbf{Y}]$, note that conditional on $\mathbf{Y}$, ξ_{ij} still follows a gamma distribution with parameters

$$\tilde{k}_{ij} = \eta_j \sum_{k=1}^{K_j}\Lambda_{ijk} + k \text{ and } \tilde{\theta}_{ij} = \sum_{k=1}^{K_j} y_{ijk} + \theta.$$

Therefore, we have

$$E[\xi_{ij}|\mathbf{Y}] = \tilde{k}_{ij}/\tilde{\theta}_{ij} \text{ and } E[\ln\xi_{ij}|\mathbf{Y}] = \psi(\tilde{k}_{ij}) - \ln(\tilde{\theta}_{ij}).$$

The above two expectations can be used to compute the Q-function at the E-step of an EM iteration. Maximization step can be realized by solving two separate equations which involve (r_0, r_1) and (k, θ) respectively. In particular, the following M-step is performed.

M-Step: From the expression of l_c, the first term depends on r_0 and r_1 only and the second term depends on k and θ only. So in the maximization step we can optimize the updated parameters $r_0^{(u+1)}, r_1^{(u+1)}, k^{(u+1)}, \theta^{(u+1)}$

separately. Specifically, $r_0^{(u+1)}, r_1^{(u+1)}$ is obtained by solving

$$\begin{cases} \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\left[\eta_j\Lambda_{ijk}\ln y_{ijk} - \eta_j\Lambda_{ijk}\psi(\eta_j\Lambda_{ijk}) + \eta_j\Lambda_{ijk}\ln\xi_{ij}\right] = 0, \\ \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_j}\beta_j\left[\eta_j\Lambda_{ijk}\ln y_{ijk} - \eta_j\Lambda_{ijk}\psi(\eta_j\Lambda_{ijk}) + \eta_j\Lambda_{ijk}\ln\xi_{ij}\right] = 0. \end{cases} \tag{C.1}$$

and $k^{(u+1)}, \theta^{(u+1)}$ is obtained by solving:

$$\begin{cases} \sum_{j=1}^{J}\sum_{i=1}^{N_j}\left[\ln\theta + \ln\xi_{ij} - \psi(k)\right] = 0, \\ \sum_{j=1}^{J}\sum_{i=1}^{N_j}\left[k/\theta - \xi_{ij}\right] = 0. \end{cases} \tag{C.2}$$

Appendix D

Candidate's publication list arising from the PhD work

Chen, L., Huang, B., and Xie, M. "Maintenance in an unreliable production system with imperfect production", submitted, under review.

Chen, L. and Tang, L.C. "A piecewise constant failure intensity model and related optimal maintenance planning", *IEEE Transactions on Reliability*, revised and submitted.

Chen, L., Ye, Z.S., Tang, L.C. and Xie, M. "Accelerated degradation test planning using the inverse Gaussian process", *IEEE Transactions on Reliability*, accepted.

Chen, L.,Ye, Z.S., and Xie, M. "Joint maintenance and spare provisioning policy for k-out-of-n systems", *Asia-Pacific Journal of Operational Research*, accepted.

Tang, L.C. and Chen, L. "A Piecewise Constant Failure Intensity Model", *Reliability and Maintainability Symposium (RAMS)*, 2014, accepted.

Chen, L.,Ye, Z.S. and Huang,B. "Condition-based maintenance for systems under dependent competing failures", *Proceedings of IEEM 2011,* pp.1586-1590.

Chen, L. and Huang, B. "A study of maintenance in a manufacturing system with machine breakdown and imperfect production", INFORMS Annual Meeting, October,2012, Phenix, Arizona, USA.

References

Arab, A., Rigdon, S., and Basu, A. (2012), "Bayesian Inference for the Piecewise Exponential Model for the Reliability of Multiple Repairable Systems," *Journal of Quality Technology*, 44(1), 28–38.

Ascher, H. and Feingold, H. (1984), *Repairable Systems Reliability: Modelling, Inference, Misconceptions and Their Causes*, New York,NY,USA: Marcel Dekker.

Ascher, H. E. (1968), "Evaluation of Repairable System Reliability Using the"Bad-As-Old"Concept," *Reliability, IEEE Transactions on*, 17 (2), 103–110.

Attardi, L., Guida, M., and Pulcini, G. (2005), "A mixed-Weibull regression model for the analysis of automotive warranty data," *Reliability Engineering & System Safety*, 87(2), 265–273.

Bae, S. J., Kuo, W., and Kvam, P. H. (2007), "Degradation models and implied lifetime distributions," *Reliability Engineering & System Safety*, 92(5), 601–608.

Barlow, R. E. and Proschan, F. (1965), *Mathematical theory of reliability*, John Wiley, New York.

Barlow, R. E., Proschan, F., and Hunter, L. C. (1996), *Mathematical Theory of Reliability*, vol. 17, Siam Philadelphia.

Ben-Daya, M. (2002), "The economic production lot-sizing problem with imperfect production processes and imperfect maintenance," *International Journal of Production Economics*, 76(3), 257–264.

Berman, M. (1981), "Inhomogeneous and modulated gamma processes," *Biometrika*, 68(1), 143–152.

Blischke, W. R. and Murthy, D. P. (2003), *Case Studies in Reliability and Maintenance*, vol. 377, Wiley-Interscience.

Boone, T., Ganeshan, R., Guo, Y., and Ord, J. (2000), "The impact of imperfect processes on production run times," *Decision Sciences*, 31(4), 773–787.

Boulanger, M. and Escobar, L. A. (1994), "Experimental design for a class of accelerated degradation tests," *Technometrics*, 36(3), 260–272.

Braun, W., Li, W., and Zhao, Y. (2005), "Properties of the geometric and related processes," *Naval Research Logistics*, 52(7), 607–616.

Brown, M. and Proschan, F. (1983), "Imperfect repair," *Journal of Applied Probability*, 851–859.

Cassady, C. R., Bowden, R. O., Liew, L., and Pohl, E. A. (2000), "Combining preventive maintenance and statistical process control: a preliminary investigation," *Iie Transactions*, 32(6), 471–478.

Chakraborty, T., Giri, B., and Chaudhuri, K. (2009), "Production lot sizing with process deterioration and machine breakdown under inspection schedule," *Omega*, 37(2), 257–271.

Chan, J., Lam, Y., and Leung, D. (2004), "Statistical inference for geometric processes with gamma distributions," *Computational Statistics and Data Analysis*, 47(3), 565–581.

Chen, D. and Trivedi, K. S. (2005), "Optimization for condition-based maintenance with semi-Markov decision process," *Reliability Engineering & System Safety*, 90(1), 25–29.

Cheng, T. (1991), "An economic order quantity model with demand-dependent unit production cost and imperfect production processes," *IIE transactions*, 23(1), 23–28.

Chhikara, R. (1988), *The Inverse Gaussian Distribution: Theory: Methodology, and Applications*, vol. 95, CRC.

Chien, Y.-H., Sheu, S.-H., Zhang, Z. G., and Love, E. (2006), "An extended optimal replacement model of systems subject to shocks," *European Journal of Operational Research*, 175(1), 399–412.

Christer, A. and Wang, W. (1992), “A model of condition monitoring of a production plant,” *The International Journal Of Production Research*, 30(9), 2199–2211.

Coetzee, J. (1996), “Reliability degradation and the equipment replacement problem,” in *Proceeding of International Conference of Maintenance Societies (ICOMS-96)*.

Coetzee, J. L. (1997), “The role of NHPP models in the practical analysis of maintenance failure data,” *Reliability Engineering & System Safety*, 56(2), 161–168.

Cook, R. J. and Lawless, J. F. (2007), *The statistical analysis of recurrent events*, Springer.

Cox, D. R. (1955), “Some statistical methods connected with series of events,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 129–164.

Crowder, M. and Lawless, J. (2007), “On a scheme for predictive maintenance,” *European Journal of Operational Research*, 176(3), 1713–1722.

Deloux, E., Castanier, B., and Bérenguer, C. (2009), “Predictive maintenance policy for a gradually deteriorating system subject to stress,” *Reliability Engineering & System Safety*, 94(2), 418–431.

Dieulle, L., Berenguer, C., Grall, A., and Roussignol, M. (2003a), "Sequential condition-based maintenance scheduling for a deteriorating system," *European Journal of operational research*, 150(2), 451–461.

— (2003b), "Sequential condition-based maintenance scheduling for a deteriorating system," *European Journal of Operational Research*, 150(2), 451–461.

Doyen, L. and Gaudoin, O. (2004), "Classes of imperfect repair models based on reduction of failure intensity or virtual age," *Reliability Engineering & System Safety*, 84(1), 45–56.

Efron, B. and Tibshirani, R. (1994), *An Introduction to the Bootstrap*, vol. 57, Chapman & Hall/CRC.

El-Ferik, S. (2008), "Economic production lot-sizing for an unreliable machine under imperfect age-based maintenance policy," *European Journal of Operational Research*, 186(1), 150–163.

Eroglu, A. and Ozdemir, G. (2007), "An economic order quantity model with defective items and shortages," *International Journal of Production Economics*, 106(2), 544–549.

Finkelstein, M. (2007), "Shocks in homogeneous and heterogeneous populations," *Reliability Engineering & System Safety*, 92(5), 569–574.

Fries, A. and Sen, A. (1996), "A survey of discrete reliability-growth models," *Reliability, IEEE Transactions on*, 45(4), 582–604.

Fuqing, Y. and Kumar, U. (2012), "A general imperfect repair model considering time-dependent repair effectiveness," *Reliability, IEEE Transactions on*, 61(1), 95–100.

Gaudoin, O., Yang, B., and Xie, M. (2003), "A simple goodness-of-fit test for the power-law process, based on the Duane plot," *Reliability, IEEE Transactions on*, 52(1), 69–74.

Ge, Z., Li, X., Jiang, T., and Huang, T. (2011), "Optimal design for step-stress accelerated degradation testing based on D-optimality," in *Reliability and Maintainability Symposium (RAMS), 2011 Proceedings-Annual*, IEEE, pp. 1–6.

Ge, Z. Z., Li, X. Y., Zhang, J. R., and Jiang, T. M. (2010), "Planning of Step-Stress Accelerated Degradation Test with Stress Optimization," *Advanced Materials Research*, 118, 117–408.

Gilardoni, G. and Colosimo, E. (2007), "Optimal maintenance time for repairable systems," *Journal of Quality Technology*, 39(1), 48–53.

Grall, A., Dieulle, L., Bérenguer, C., and Roussignol, M. (2002), "Continuous-time predictive-maintenance scheduling for a deteriorating system," *Reliability, IEEE Transactions on*, 51(2), 141–150.

Groenevelt, H., Pintelon, L., and Seidmann, A. (1992), "Production lot sizing with machine breakdowns," *Management Science*, 38(1), 104–123.

Guida, M. and Pulcini, G. (2009), "Reliability analysis of mechanical systems with bounded and bathtub shaped intensity function," *Reliability, IEEE Transactions on*, 58(3), 432–443.

Hamada, M., Wilson, A., Reese, C., and Martz, H. (2008), *Bayesian reliability*, Springer.

Hong, Y. and Meeker, W. Q. (2010), "Field-failure and warranty prediction based on auxiliary use-rate information," *Technometrics*, 52(2), 148–159.

— (2013), "Field-Failure Predictions Based on Failure-time Data with Dynamic Covariate Information," *Technometrics*, (just-accepted).

Hopp, W. and Spearman, M. (2008), *Factory physics, 3rd*, Boston: McGraw-Hill/Irwin.

Huang, W. and Askin, R. G. (2003), "Reliability analysis of electronic devices with multiple competing failure modes involving performance aging degradation," *Quality and Reliability Engineering International*, 19(3), 241–254.

Huynh, K. T., Barros, A., Berenguer, C., and Castro, I. T. (2011), "A periodic inspection and replacement policy for systems subject to competing failure modes due to degradation and traumatic events," *Reliability Engineering & System Safety*, 96(4), 497–508.

Jaturonnatee, J., Murthy, D., and Boondiskulchok, R. (2006), "Optimal preventive maintenance of leased equipment with corrective minimal repairs," *European Journal of Operational Research*, 174(1), 201–215.

Kharoufeh, J. P., Finkelstein, D. E., and Mixon, D. G. (2006), "Availability of periodically inspected systems with Markovian wear and shocks," *Journal of Applied Probability*, 43(2), 303–317.

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

Klutke, G.-A. and Yang, Y. (2002), "The availability of inspected systems subject to shocks and graceful degradation," *Reliability, IEEE Transactions on*, 51(3), 371–374.

Kumar, U. and Klefsjö, B. (1992), "Reliability analysis of hydraulic systems of LHD machines using the power law process model," *Reliability Engineering and System Safety*, 35(3), 217–224.

Lam, Y. (2007), "A geometric process maintenance model with preventive repair," *European Journal of Operational Research*, 182(2), 806–819.

— (2009), "A Geometric Process delta-Shock Maintenance Model," *IEEE Transactions on Reliability*, 58(2), 389–396.

Lam, Y. and Zhang, Y. L. (2004), "A shock model for the maintenance problem of a repairable system," *Computers & Operations Research*, 31 (11), 1807–1820.

Lawless, J. and Crowder, M. (2004), "Covariates and random effects in a gamma process model with application to degradation and failure," *Lifetime Data Analysis*, 10(3), 213–227.

Lawless, J. F., Çiğşar, C., and Cook, R. J. (2012), "Testing for monotone trend in recurrent event processes," *Technometrics*, 54(2), 147–158.

Lee, H. and Rosenblatt, M. (1987), "Simultaneous determination of production cycle and inspection schedules in a production systems," *Management Science*, 1125–1136.

Lee, L. (1980), "Testing adequacy of the Weibull and log linear rate models for a Poisson process," *Technometrics*, 22(2), 195–199.

Li, W. and Pham, H. (2005), "An inspection-maintenance model for systems with multiple competing processes," *Reliability, IEEE Transactions on*, 54(2), 318–327.

Liao, C. and Tseng, S. (2006), "Optimal design for step-stress accelerated degradation tests," *Reliability, IEEE Transactions on*, 55(1), 59–66.

Liao, G., Chen, Y., and Sheu, S. (2009), "Optimal economic production quantity policy for imperfect process with imperfect repair and maintenance," *European Journal of Operational Research*, 195(2), 348–357.

Liao, H. (2009), "Optimal design of accelerated life testing plans for periodical replacement with penalty," *Naval Research Logistics (NRL)*, 56 (1), 19–32.

Liao, H. and Elsayed, E. A. (2006), "Reliability inference for field conditions from accelerated degradation testing," *Naval Research Logistics (NRL)*, 53(6), 576–587.

Liao, H., Elsayed, E. A., and Chan, L.-Y. (2006), "Maintenance of continuously monitored degrading systems," *European Journal of Operational Research*, 175(2), 821–835.

Lim, H. (2012), "Optimal Design of Accelerated Degradation Tests under the Constraint of Total Experimental Cost in the Case that the Degradation Characteristic Follows a Wiener Process," *Journal of the Korean Society for Quality Management*, 40(2), 117–125.

Lim, H. and Yum, B. (2011), "Optimal design of accelerated degradation tests based on Wiener process models," *Journal of Applied Statistics*, 38 (2), 309–325.

Lindqvist, B., Elvebakk, G., and Heggland, K. (2003), "The trend-renewal process for statistical analysis of repairable systems," *Technometrics*, 45 (1), 31–44.

Liu, X. (2012), "Planning of accelerated life tests with dependent failure modes based on a gamma frailty model," *Technometrics*, 54(4), 398–409.

Liu, X., Li, J., Al-Khalifa, K. N., Hamouda, A. S., Coit, D. W., and Elsayed, E. A. (2012), "Condition-based maintenance for continuously monitored degrading systems with multiple failure modes," *IIE Transactions*, (just-accepted).

Lu, C. J. and Meeker, W. O. (1993), "Using degradation measures to estimate a time-to-failure distribution," *Technometrics*, 35(2), 161–174.

Lu, S., Tu, Y.-C., and Lu, H. (2007), "Predictive condition-based maintenance for continuously deteriorating systems," *Quality and Reliability Engineering International*, 23(1), 71–81.

Maillart, L. M. (2006), "Maintenance policies for systems with condition monitoring and obvious failures," *IIE Transactions*, 38(6), 463–475.

Makis, V. and Jardine, A. K. (1992), "Optimal replacement in the proportional hazards model," *Infor*, 30(1), 172–183.

Meeker, W. and Escobar, L. (1998), "Statistical Methods for Reliability Data," *New York*.

Meeker, W., Escobar, L., and Lu, C. (1998), "Accelerated degradation tests: modeling and analysis," *Technometrics*, 40(2), 89–99.

Meeker, W. Q., Escobar, L. A., and Hong, Y. (2009), "Using accelerated life tests results to predict product field reliability," *Technometrics*, 51 (2), 146–161.

Meeker, W. Q. and Hong, Y. (2014), "Reliability meets big data: Opportunities and challenges," *Quality Engineering*, 26(1), 102–116.

Mun, B. M. and Bae, S. J. (2011), "Optimal maintenance policy of repairable system with bathtub shaped intensity," in *Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2011 International Conference on*, IEEE, pp. 431–435.

Muralidharan, K. (2002), "Reliability inferences of modulated power-law process," *Reliability, IEEE Transactions on*, 51(1), 23–26.

Murthy, D. and Kobbacy, K. (2008), *Complex System Maintenance Handbook*, Springer.

Nelder, J. A. and Mead, R. (1965), "A simplex method for function minimization," *The computer journal*, 7(4), 308–313.

Nelson, W. (1988), "Graphical analysis of system repair data." *Journal of Quality Technology*, 20(1), 24–35.

— (2005), "A bibliography of accelerated test plans," *Reliability, IEEE Transactions on*, 54(2), 194–197.

Nelson, W. B. (1990), *Accelerated testing: statistical models, test plans, and data analysis*, vol. 344, Wiley-Interscience.

Padgett, W. and Tomlinson, M. A. (2004), "Inference from accelerated degradation and failure data based on Gaussian process models," *Lifetime Data Analysis*, 10(2), 191–206.

Pan, R. (2009), "A Bayes approach to reliability prediction utilizing data from accelerated life tests and field failure observations," *Quality and Reliability Engineering International*, 25(2), 229–240.

Park, C. and Padgett, W. (2005), "Accelerated degradation models for failure based on geometric Brownian motion and gamma processes," *Lifetime Data Analysis*, 11(4), 511–527.

Park, J.-I. and Yum, B.-J. (1997), "Optimal design of accelerated degradation tests for estimating mean lifetime at the use condition," *Engineering Optimization+ A35*, 28(3), 199–230.

Peng, C. and Tseng, S. (2010), "Progressive-stress accelerated degradation test for highly-reliable products," *Reliability, IEEE Transactions on*, 59 (1), 30–37.

Peng, H., Feng, Q., and Coit, D. W. (2009), "Simultaneous quality and reliability optimization for microengines subject to degradation," *Reliability, IEEE Transactions on*, 58(1), 98–105.

— (2010), "Reliability and maintenance modeling for systems subject to multiple dependent competing failure processes," *IIE transactions*, 43 (1), 12–22.

Pulcini, G. (2001), "A bounded intensity process for the reliability of repairable equipment." *Journal of Quality Technology*, 33(4), 480–492.

— (2008), "Repairable system analysis for bounded intensity functions and various operating conditions," *Journal of Quality Technology*, 40(1), 78–96.

— (2014), "A Model-Driven Approach for the Failure Data Analysis of Multiple Repairable Systems Without Information on Individual Sequences," *Reliability, IEEE Transactions on.*

Rigdon, SE, M. X. B. K. (1998), "Statistical inference for repairable systems using the power law process," *Journal of Quality Technology*, 30(4), 395–400.

Rigdon, S. and Basu, A. (2000), *Statistical Methods for the Reliability of Repairable Systems*, Wiley New York, NY.

Ross, S. (1983), *Stochastic processes*, vol. 23, John Wiley and Sons, New York.

— (1996), *Stochastic Processes*, vol. 2, John Wiley & Sons Hoboken, NJ, USA.

Rouse, G. and Kelly, J. (2011), "Electricity Reliability: Problems, Progress, and Policy Solutions," *Galvin Electricity Initiative, Chicago, IL.*

Salameh, M. and Jaber, M. (2000), "Economic production quantity model for items with imperfect quality," *International journal of production economics*, 64(1), 59–64.

Samatl-Pa, G. and Taner, M. (2009), "The role of repair strategy in warranty cost minimization: An investigation via quasi-renewal processes," *European Journal of Operational Research*, 197(2), 632–641.

Sana, S. (2010), "An economic production lot size model in an imperfect production system," *European Journal of Operational Research*, 201(1), 158–170.

Sen, A. (1998), "Estimation of current reliability in a Duane-based reliability growth model," *Technometrics*, 40(4), 334–344.

Sen, A. and Bhattacharyya, G. (1993), "A piecewise exponential model for reliability growth and associated inferences," *Advances in Reliability*, 331–355.

Sheu, S. and Chen, J. (2004), "Optimal lot-sizing problem with imperfect maintenance and imperfect production," *International Journal of Systems Science*, 35(1), 69–77.

Shi, Y., Escobar, L. A., and Meeker, W. Q. (2009), "Accelerated destructive degradation test planning," *Technometrics*, 51(1), 1–13.

Shi, Y. and Meeker, W. Q. (2012), "Bayesian Methods for Accelerated Destructive Degradation Test Planning," *Reliability, IEEE Transactions on*, 61(1), 245–253.

Si, X.-S., Wang, W., Hu, C.-H., and Zhou, D.-H. (2011), "Remaining useful life estimation–A review on the statistical data driven approaches," *European Journal of Operational Research*, 213(1), 1–14.

— (2014), "Estimating Remaining Useful Life With Three-Source Variability in Degradation Modeling," .

Singpurwalla, N. (1995), "Survival in dynamic environments," *Statistical Science*, 10(1), 86–103.

Singpurwalla, N. D. (2006), "The hazard potential: introduction and overview," *Journal of the American Statistical Association*, 101(476), 1705–1717.

Tang, L. and Olorunniwo, F. (1989), "A maintenance model for repairable systems," *Reliability Engineering & System Safety*, 24(1), 21–32.

Tang, L., Yang, G., and Xie, M. (2004), "Planning of step-stress accelerated degradation test," in *Reliability and Maintainability, 2004 Annual Symposium-RAMS*, IEEE, pp. 287–292.

Tang, L. C. and Chang, D. S. (1995), "Reliability prediction using nondestructive accelerated-degradation data: case study on power supplies," *Reliability, IEEE Transactions on*, 44(4), 562–566.

Tang, L. C. and Liu, X. (2010), "Planning and inference for a sequential accelerated life test," *Journal of quality technology*, 42(1), 103–118.

Thompson, W. (1988), *Point Process Models with Applications to Safety and Reliability*, Chapman and Hall New York.

Tsai, C., Tseng, S., and Balakrishnan, N. (2012), "Optimal Design for Degradation Tests Based on Gamma Processes With Random Effects," *Reliability, IEEE Transactions on*, 61(2), 604–613.

Tsao, Y.-C., Chen, T.-H., and Zhang, Q.-H. (2013), "Effects of maintenance policy on an imperfect production system under trade credit," *International Journal of Production Research*, 51(5), 1549–1562.

Tseng, S. (1996), "Optimal preventive maintenance policy for deteriorating production systems," *IIE transactions*, 28(8), 687–694.

Tseng, S., Balakrishnan, N., and Tsai, C. (2009), “Optimal step-stress accelerated degradation test plan for gamma degradation processes,” *Reliability, IEEE Transactions on*, 58(4), 611–618.

Tseng, S. and Wen, Z. (2000), “Step-stress accelerated degradation analysis for highly reliable products,” *Journal of Quality Technology*, 32(3).

Tseng, S., Yeh, R., and Ho, W. (1998), “Imperfect maintenance policies for deteriorating production systems,” *International journal of production economics*, 55(2), 191–201.

Tseng, S.-T. and Yu, H.-F. (1997), “A termination rule for degradation experiments,” *Reliability, IEEE Transactions on*, 46(1), 130–133.

Van Noortwijk, J. (2009), “A survey of the application of gamma processes in maintenance,” *Reliability Engineering & System Safety*, 94(1), 2–21.

Wang, G. and Zhang, Y. (2005), “A shock model with two-type failures and optimal replacement policy,” *International Journal of Systems Science*, 36(4), 209–214.

Wang, H. (2002), “A survey of maintenance policies of deteriorating systems,” *European Journal of Operational Research*, 139(3), 469–489.

Wang, H. and Pham, H. (1996), “A quasi renewal process and its applications in imperfect maintenance,” *International journal of systems science*, 27(10), 1055–1062.

— (2006), "Availability and maintenance of series systems subject to imperfect repair and correlated failure and repair," *European Journal of Operational Research*, 174(3), 1706–1722.

Wang, L., Pan, R., Li, X., and Jiang, T. (2012), "A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information," *Reliability Engineering & System Safety*.

Wang, P., Lin, Y., Chen, Y., and Chen, J. (2009), "An optimal production lot-sizing problem for an imperfect process with imperfect maintenance and inspection time length," *International Journal of Systems Science*, 40(10), 1051–1061.

Wang, W. (2012), "A simulation-based multivariate Bayesian control chart for real time condition-based maintenance of complex systems," *European Journal of Operational Research*, 218(3), 726–734.

Wang, W.-b. (2000), "A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance," *International Journal of Production Research*, 38(6), 1425–1436.

Wang, X. (2010), "Wiener processes with random effects for degradation data," *Journal of Multivariate Analysis*, 101(2), 340–351.

Wang, X. and Xu, D. (2010), "An inverse Gaussian process model for degradation data," *Technometrics*, 52(2), 188–197.

Wasan, M. (1968), "On an inverse Gaussian process," *Scandinavian Actuarial Journal*, 1968(1-2), 69–96.

Wee, H.-M., Wang, W.-T., and Yang, P.-C. (2013), "A production quantity model for imperfect quality items with shortage and screening constraint," *International Journal of Production Research*, 51(6), 1869–1884.

Yang, G. (2007), *Life Cycle Reliability Engineering*, Wiley.

— (2010), "Accelerated life test plans for predicting warranty cost," *Reliability, IEEE Transactions on*, 59(4), 628–634.

Yang, Q., Hong, Y., Chen, Y., and Shi, J. (2012), "Failure profile analysis of complex repairable systems with multiple failure modes," *Reliability, IEEE Transactions on*, 61(1), 180–191.

Yang, Z. M., Djurdjanovic, D., and Ni, J. (2008), "Maintenance scheduling in manufacturing systems based on predicted machine degradation," *Journal of Intelligent Manufacturing*, 19(1), 87–98.

Ye, Z. S., Tang, L. C., and Xu, H. Y. (2011), "A distribution-based systems reliability model under extreme shocks and natural degradation," *Reliability, IEEE Transactions on*, 60(1), 246–256.

Ye, Z.-S., Wang, Y., Tsui, K.-L., and Pecht, M. (2012), "Degradation data analysis using Wiener processes with measurement errors," *Reliability,IEEE Transactions on.*, 62(4), 772–780.

Yeh, L. (1992), "Nonparametric inference for geometric processes," *Communications in Statistics-Theory and Methods*, 21(7), 2083–2105.

Yeh, L. and Chan, S. (1998), "Statistical inference for geometric processes with lognormal distribution," *Computational Statistics and Data Analysis*, 27(1), 99–112.

Yeung, T. G., Cassady, C. R., and Schneider, K. (2007), "Simultaneous optimization of [Xbar] control chart and age-based preventive maintenance policies under an economic objective," *IIE Transactions*, 40(2), 147–159.

Yu, H.-F. and Tseng, S.-T. (1998), "On-line procedure for terminating an accelerated degradation test," *Statistica Sinica*, 8, 207–220.

Zhou, J., Djurdjanovic, D., Ivy, J., and Ni, J. (2007), "Integrated reconfiguration and age-based preventive maintenance decision making," *IIE Transactions*, 39(12), 1085–1102.