UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ASSESSMENT OF SEMI-PARAMETRIC PROPORTIONAL INTENSITY MODELS APPLIED TO RECURRENT FAILURE DATA WITH MULTIPLE FAILURE TYPES FOR REPAIRABLE-SYSTEM RELIABILITY

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By
SHWU-TZY JIANG
Norman, Oklahoma
March, 2004

ASSESSMENT OF SEMI-PARAMETRIC PROPORTIONAL INTENSITY MODELS APPLIED TO RECURRENT FAILURE DATA WITH MULTIPLE FAILURE TYPES FOR REPAIRABLE-SYSTEM RELIABILITY

A Dissertation APPROVED FOR THE
SCHOOL OF INDUSTRIAL ENGINEERING

BY

Thomas L. Landers ______

Kuang-Hua Chang ______

Babur M. Pulat ______

Pakize S. Pulat ______

Teri R. Rhoads ______

Acknowledgements

Dr. Thomas L. Landers has provided guidance, encouragement, and financial assistantship needed to accomplish my dissertation research. He has also spent uncountable hours reviewing this dissertation with profession and enthusiasm. Through his teaching as a great faculty mentor, I have been equipped in a multidisciplinary environment that allows me to develop a broader vision to define research questions typically encountered in engineering and medical applications and devise a plan/methodology to fulfill the intended research goals. I am truly blessed to have him as my major advisor during the period of my doctoral studies.

Thanks to Dr. Rhoads for her hard work on the edits that has made this dissertation more complete. Thanks also to Drs. S. Pulat, Chang, and M. Pulat for their willingness to serve in my committee and their valuable comments towards this dissertation. Several staff members in the College of Engineering Dean's Office and the School of Industrial Engineering have provided a professional and convenient environment that has facilitated this research. I would like to express my deep appreciation to my husband and family for the great love and patience they have offered throughout my dissertation research/project. Thanks to my host families, Drs. Thomas and Patti Landers and Harold and Clara Jeffery, for sharing their love and friendship during my doctoral studies. I would like to give thanks to God for fulfilling all the needs in time. Without Him, this dissertation would not have been completed.

TABLE OF CONTENTS

		Page
Acknowledgements		iv
Table of Contents		v
List of Tables		viii
List of Figures		x
List of Symbols		xi
Acronyms		xiii
Abstract		xiv
1	Introduction	1
2	Literature review	5
2.1	Single-event models for non-repairable systems	5
2.1.1	Parametric survival models	6
2.1.2	Maximum likelihood estimators	14
2.1.3	Semi-parametric Cox models	16
2.1.4	Competing risk models	21
2.2	Recurring-event models for repairable systems	22
2.2.1	Parametric stochastic processes	23
2.2.2	Proportional intensity models	27
2.2.3	Robustness studies	48
2.2.4	Censoring of recurrent events	68
2.2.5	Multiple event types	78
2.3	References	83
3	Research methodology	86
3.1	Robustness of semi-parametric methods under right censoring	87
3.1.1	NHPP	87
3.1.2	HPP	92
3.2	Modeling of multiple failure types in recurrent events	93
3.2.1	Multi-dimensional covariate modeling	94
3.2.2	Discontinuous risk-free-intervals modeling	96
4	Semi-parametric proportional intensity models robustness for right-censored recurrent failure data	102
4.0	Abstract and Nomenclature (Acronyms and Notation)	102
4.1	Introduction	105
4.2	Semi-parametric Proportional Intensity models	107
4.3	Models and methods	112
4.3.1	Cox regression models	112
4.3.2	Semi-parametric PWP model	113
4.3.3	Semi-parametric AG model	114
4.3.4	Semi-parametric WLW model	115
4.3.5	Power-law intensity function	117
4.3.6	Method	117
4.4	Results	121
4.4.1	PWP-GT model results	121
4.4.2	Effect of heavy censoring	126
4.4.3	Complete data	127

TABLE OF CONTENTS (cont'd)

		Page
4.4.4	95% C.I. on $\hat{\beta}_i$	130
4.4.5	AG, PWP-TT, and WLW models	132
4.5	Conclusions	134
4.6	References	137
5	Robustness of Semi-parametric proportional intensity models for right-censored recurrent failure data from a stationary counting process	139
5.0	Abstract and Nomenclature (Acronyms and Notation)	139
5.1	Introduction	142
5.2	Semi-parametric Proportional Intensity models	144
5.3	Models and methods	150
5.3.1	Cox regression models	150
5.3.2	Semi-parametric PWP model	151
5.3.3	Semi-parametric AG model	152
5.3.4	Semi-parametric WLW model	152
5.3.5	Method	154
5.4	Results	158
5.4.1	Event-specific baseline models (PWP-GT, PWP-TT, and WLW)	158
5.4.2	Common baseline model (AG)	164
5.5	Conclusions	165
5.6	References	168
6	Semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals	170
6.0	Abstract and Nomenclature (Acronyms and Notation)	170
6.1	Introduction	173
6.2	Semi-parametric Proportional Intensity models	175
6.3	Models and methods	182
6.3.1	Cox regression models	182
6.3.2	Semi-parametric PWP model	183
6.3.3	Semi-parametric AG model	184
6.3.4	Semi-parametric WLW model	185
6.3.5	Power-law intensity function	187
6.3.6	Method	188
6.4	Results	192
6.4.1	PWP-GT model results	192
6.4.2	AG and WLW models results	196
6.5	Conclusions	200
6.6	References	202
7	Covariate proportional intensity modeling for recurrent data of two failure types (major and minor)	204
7.0	Abstract and Nomenclature (Acronyms and Notation)	204
7.1	Introduction	206
7.2	Semi-parametric Proportional Intensity models	208
7.3	Models and methods	210
7.3.1	Cox regression models	211

TABLE OF CONTENTS (cont'd)

		Page
7.3.2	Semi-parametric PWP model	211
7.3.3	Power-law intensity function	212
7.3.4	Method	213
7.4	Results	216
7.5	Conclusions	219
7.6	References	221
8	Conclusions and recommendations	223
8.1	Conclusions	223
8.1.1	Right-censoring recurring events on four PI models	223
8.1.2	Right-censoring recurring events on event-specific PI models	224
8.1.3	Discontinuous risk-free-intervals	224
8.1.4	Covariate PI modeling	224
8.2	Recommendations	225
8.2.1	Right-censoring recurring events on four PI models	225
8.2.2	Right-censoring recurring events on event-specific PI models	226
8.2.3	Discontinuous risk-free-intervals	226
8.2.4	Covariate PI modeling	227
8.3	Contributions	228
8.4	Future research	229
8.4.1	Right-censored recurrent events	229
8.4.2	Multiple failure types	229
8.4.3	Left-censored recurrent events	230
Appendix I	Semi-parametric proportional intensity models robustness for right-censored recurrent failure data	231
I.1	Experimental units (singular failure type)	231
I.2	Shape parameter (singular failure type)	235
I.3	95% C.I. Charts (major and minor types)	238
I.4	Experimental units (major and minor types)	244
I.5	Shape parameter (major and minor types)	247
Appendix II	Semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals	249
II.1	Experimental units	249
II.2	Shape parameter	256
Appendix III	Program for semi-parametric proportional intensity models robustness for right-censored recurrent failure data	257
Appendix IV	Program for robustness of semi-parametric proportional intensity models for right-censored recurrent failure data from a stationary counting process	260
Appendix V	Program for semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals	263
Appendix VI	Program for covariate proportional intensity modeling for recurrent data of two failure types (major and minor)	265
Appendix VII	Glossary	267

LIST OF TABLES

Table		Page
2.1	Taxonomy of risk interval and risk set for each model	48
2.2	The results of ($\hat{\delta}$, $\hat{\beta}_0$, $\hat{\beta}_1$) using the Lawless method	52
2.3	The parametric Lawless method	53
2.4	Average $\hat{MTTF}$ obtained from the PWP method	54
2.5	The semi-parametric PWP method	55
2.6	The summary of the robustness test using the lawless method	57
2.7	Qureshi results	57
2.8	The summary of the robustness test using the PWP method	57
2.9	Qureshi results	57
2.10	GLM summary (Qureshi (1991))	58
2.11	($\hat{\theta}$, $\hat{\beta}_0$, $\hat{\beta}_1$) obtained from the Lawless method	62
2.12	The parametric Lawless method	63
2.13	Average $\hat{MTTF}$ obtained from the PWP method	64
2.14	Average $\hat{MTTF}$ obtained from the PWP method (corrected)	65
2.15	The semi-parametric PWP method	66
2.16	The summary of the robustness test using the Lawless method	67
2.17	Vithala results	67
2.18	The summary of the robustness test using the PWP method	67
2.19	Vithala results	68
2.20	GLM summary (Vithala (1994))	68
3.1	Three-factor experimental design: (U,δ,P_c)	91
3.2	Two-factor experimental design: (U,P_c)	93
3.3	Two-factor experimental design: (U,δ)	96
3.4	Simulation dataset example	101
3.5	Three-factor experimental design: (U,δ,R)	101
4.1	Summary of PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit)	124
4.2	Performance metrics (PWP-GT) in three power-law intensity functions	128
4.3	Performance metrics (PWP-GT) in three sample sizes, $\delta=1.5$	128
4.4	95% C.I. on $\hat{\beta}_i$, $(\delta,P_c)=(1.0,1.0)$, for three sample sizes, 60, 120, and 180	131
4.5	Summary of semi-parametric AG model results for $\hat{\beta}$ ($\delta=1$, HPP)	133
5.1	Performance metrics of $\hat{\beta}_i$ in a HPP case (PWP-GT, PWP-TT, and WLW)	159
5.2	95% C.I. of AG estimates for $\hat{\beta}$ ($\delta=1$, HPP)	164
6.1	Summary of PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit)	194
6.2	Performance metrics of the AG and WLW models in HPP, U=20	198
6.3	Performance metrics of the AG and WLW models in HPP, U=60	198
6.4	Performance metrics of the AG and WLW models and the 95%	199

C.I. in HPP, U=120

7.1 Summary of PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit) 217

LIST OF FIGURES

Fig.		Page
2.1	Exponential hazard function with $\lambda = 2$	7
2.2	Weibull hazard function with combination of δ and υ, where υ=5	8
2.3	Type I smallest value of extreme value hazard functions	10
2.4	Gamma hazard function with different shape parameters	11
2.5(a)-(b)	Lognormal distribution with parameters (μ, σ)	12
2.6	Log-Logistic hazard function with parameters (p, λ)	13
2.7(a)-(c)	Risk interval formulations (Kelly and Lim(2000))	47
2.8	A hypothetical example from Kelly and Lim (2000)	76
4.1(a)-(c)	Risk interval formulations (Kelly and Lim [23])	111
4.2(a)-(c)	PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures /unit), $\delta = 0.5, 1.0$, and 1.5	125
4.3	Performance comparisons of 3 and 6 replicates (PWP-GT) $\delta = 1.5$, $P_c = 1.0$	126
4.4	Error plots (PWP-GT) in three shape parameters ($\delta = 1.5, 1.0$, and 0.5) and three sample sizes (U=60,120, and 180) at $\delta = 1.5$	129
4.5(a)-(c)	95% C.I. on PWP-GT estimates at each failure count for three sample sizes 60,120, and 180, $(\delta, P_c) = (1.0, 1.0)$	131
4.6(a)-(c)	AG, PWP-TT, and WLW model results, U=60, 120, and 180 ($\delta = 1.0$)	133
5.1(a)-(c)	Risk interval formulations (Kelly and Lim [23])	149
5.2	Comparison of three event-specific estimators of the covariate effect, $(U, P_c) = (120, 0.4)$	158
5.3(a)-(c)	Comparison of PWP-GT, PWP-TT, and WLW model results for estimating $\hat{\beta}_i$ (4 failures/unit)	161
5.4(a)-(c)	Comparison of three event-specific models at three sample sizes (refer also to Table 5.1)	163
5.5	95% C.I. of AG estimates for $\hat{\beta}$	165
6.1(a)-(c)	Risk interval formulations (Kelly and Lim [26])	181
6.2	3-D plots, PWP-GT model results for sample size of 120	195
6.3	HPP, AG and WLW estimates vs. Gap time ratio, U=120	199
7.1	3-D plots, PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit)	218

LIST OF SYMBOLS

C_{ki}	Censoring time for the i^{th} subject of the k^{th} type of failures
D	Major overhaul duration
$E[N(t)]$	Expected number of failures in $(0,t]$
$E(t_n)$	Expected (cumulative) time of the n^{th} failure
e, e_n	Relative errors of proportional intensity model estimates
$f(t)$	Probability density function of failure time distribution
F	Major failure event number
$\mathbf{G}_{..}$	Information matrix (2 by 2) in the Newton-Raphson algorithm
h	Interval between t and $t+h$, limit to time zero
$h(t)$	Hazard function; proportional hazard function
$h_0(t)$	Baseline hazard function
I_0	Number of sample units in class ϕ
I_1	Number of sample units in class 1
i.i.d.	Independent and identically distributed
$L(.)$	Likelihood function
N	Successive failure count
$N(t)$	Random variable for the number of failures in $(0,t]$; a counting process
n	An integer counting successive failure times; a stratification indicator subscript
p, λ	Parameters of a log-logistic distribution
P_c	Censoring probability
$Q(a,b)$	Crude probability of failure in $[a,b)$ in the competing risk model
$q(a,b)$	Net probability of failure in $[a,b)$ in the competing risk model
R	Gap time ratio that controls the magnitude of D
R_i	Treatment indicator in the multi-dimensional covariate modeling
$R(t)$	Reliability function
$S(y \mid z)$	Survival function of y give z in the PH model, y: the logarithm of lifetime, z: covariate
s.d.	Standard deviation
T_i	Cumulative time of failure event i
T_e	End of an event
T_g	Global time of an event (clock hour)
T_s	Beginning of an event
t_{ij}	Observed failure time corresponding to sample unit i and failure count j
t_F	Cumulative failure time of the first major failure
t_{F+1}	Cumulative time after performing a major overhaul D
ΔT	Elapsed time of an event

U	Sample size (number of units)
$\mathbf{U}.$	Score vector (1 by 2) in the Newton-Raphson algorithm
$\widetilde{X}$	Observation time
$Y_i^{(n)}$	An at-risk indicator in the AG model
Y_{F-1}	Gap time associated with a minor failure prior to a major failure
$\mathbf{Z}$	Vector of covariates
$\mathbf{Z}_1(t)$	A two-dimensional covariate for major type failures
$\mathbf{Z}_2(t)$	A two-dimensional covariate for minor type failures
$\boldsymbol{\beta}_n$	$(k \times 1)$ vector of stratum-specific regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)$
$\boldsymbol{\beta}_{1n}$	Regression coefficient for major failure events, n: event count
$\boldsymbol{\beta}_{2n}$	Regression coefficient for minor failure events, n: event count
δ	Shape parameter of a power-law NHPP
δ_1	Shape parameter of the major type events
δ_2	Shape parameter of the minor type events
δ, υ	Parameters of a Weibull distribution
Φ	Standardized normal distribution
Γ	Gamma function
γ, θ	Parameters of an extreme value distribution
η, γ	Parameters of a gamma distribution
$\lambda(t; \mathbf{z})$	Proportional intensity function
$\lambda_0(t)$	Proportional baseline intensity function
$\lambda_{0n}(t)$	Event-specific baseline intensity function
λ	Parameter of an exponential distribution
μ	Mean parameter; scale parameter of a log-linear NHPP
μ, σ	Parameters of a lognormal distribution
θ	Shape parameter of a log-linear NHPP; correlation among recurring events
σ	Standard deviation parameter
υ_0	Baseline value of scale parameter of a power-law NHPP
υ_1	Alternate value of a power-law NHPP
^	Denotes an estimator
'	Denotes the transpose of a vector
Δ	Indicator of a failure or censored time; limit to time zero

ACRONYMS

AG	Andersen and Gill model
C.I.	Confidence interval
CR	Competing risk
CM	Corrective maintenance
CMTBF	Cumulative mean time between failures
DROCOF	Decreasing rate of occurrence of failures
HPP	Homogeneous Poisson process
IMTBF	Instantaneous mean time between failures
IROCOF	Increasing rate of occurrence of failures
i.i.d	Independent and identically distributed
LWA	Lee, Wei, and Amato model
MTTF	Mean time to failure
MAD	Mean absolute deviation
MSE	Mean squared error
m.l.e.	Maximum likelihood estimator
NHPP	Non-homogeneous Poisson Process
PH	proportional hazards
PI	Proportional intensity
PM	Preventive maintenance
PWP	Prentice, Williams, and Peterson model
PWP-GT	Prentice, Williams, and Peterson-gap time model
PWP-TT	Prentice, Williams, and Peterson-total time model
ROCOF	Rate of occurrence of failures
WLW	Wei, Lin, and Weissfeld model

ASSESSMENT OF SEMI-PARAMETRIC PROPORTIONAL INTENSITY MODELS APPLIED TO RECURRENT FAILURE DATA WITH MULTIPLE FAILURE TYPES FOR REPAIRABLE-SYSTEM RELIABILITY

Abstract

The class of semi-parametric proportional intensity (PI) models applies to recurrent failure event modeling for a repairable system with covariates. Abundant federal funding received in biostatistics/medical research has advanced the PI models to become well developed and widely referenced. PI models for medical applications could also apply to recurring failure/repair data in engineering problems. Wider engineering use of these models requires better understanding of applications, performance, and methods to accommodate important situations such as censoring, maintenance intervals, and multiple failure types.

Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001) have examined robustness of the Prentice-Williams-Peterson-gap time (PWP-GT) model for the case of an underlying Non-homogeneous Poisson Process (NHPP) with power-law and log-linear intensity functions and complete (uncensored) data. However, the phenomenon of censoring is generally present in field data. This research has extended their work to the important case of right-censorship and has examined other semi-parametric PI models (Prentice-Williams-Peterson-total time (PWP-TT), Andersen-Gill (AG), and Wei-Lin-Weissfeld (WLW)). The experimental design in this research has incorporated three levels of censorship severity (light, moderate, and severe) to evaluate these four proposed PI models.

Certain systems experience a substantial period of downtime due to performing maintenance (i.e. major overhaul) following a major failure. This discontinuity in observation time has been a concern in the accuracy of estimating the covariate effect. Therneau and Hamilton (1997) proposed a discontinuous risk-free-intervals method for biomedical applications that could also apply to this engineering problem. This study has recommended selecting appropriate PI models and the more favorable engineering applications range for the overhaul duration based on the sample size and shape parameter. This research has examined two levels of the overhaul duration (short and long) to evaluate the PI models.

Major and minor failure events are commonly seen in industry, where minor failure rate is typically higher than major failure rate. Most researchers have formulated the problem as univariate and pooled the major and minor failures as though they are identical. Lin (1993, 1994) proposed a covariate PI modeling approach to handle the recurrent data with multiple failure types. Although covariates are typically used to incorporate treatment effects, a covariate may be defined to conceptually model multiple failure types in the special case where the proportional intensities rule holds. This study has examined covariate PI modeling as an approach for explicit treatment of multiple (two) recurrent failure types (major and minor) with complete data.

The PWP-GT and AG models prove to outperform the PWP-TT and WLW models in the robustness studies on right-censoring severity and multiple failure types. The AG model performs well in the HPP case. The results of examining

the PI models in the discontinuous risk-free-intervals modeling indicate that the PWP-GT model performs better in the short overhaul duration than the long overhaul duration. The AG model performs consistently well in the small sample size (20) regardless of the overhaul duration in an HPP case. The WLW model performance improves as the overhaul duration increases.
Keywords: repairable system reliability, proportional intensity model, recurrent events, covariates, right-censored recurrent events, major repairs, overhauls, multiple failure types, covariate proportional intensity modeling, power-law NHPP, risk interval, Prentice-Williams-Peterson, Andersen-Gill, Wei-Lin-Weissfeld

1 Introduction

Aircraft, automobiles, and process machine tools are examples of systems designed to be repairable. These systems undergo during their lifetimes multiple recurrent unscheduled failure and repair cycles and/or scheduled preventive maintenance or overhaul cycles. This research addresses statistical modeling of recurrent failure events in repairable systems reliability, by building on previous work of Qureshi (1991, 1994) and Vithala (1994). They examined the robustness of a semi-parametric Prentice-Williams-Peterson-gap time (PWP-GT) model for estimating the covariate effect where the underlying stochastic process is a Non-homogeneous Poisson Process (NHPP) with power-law or log-linear intensity function, respectively. Both Qureshi and Vithala restricted their studies to the case of complete (uncensored) data.

This research provides a thorough review of the relevant literature (Chapter 2) on the parametric survival models, semi-parametric Cox proportional hazards (PH) method for single failure event (non-repairable systems); and both the parametric Lawless (proportional intensity) and semi-parametric proportional intensity (PI) models for recurrent events, including the PWP-GT (1981) model examined by Qureshi and Vithala as well as the PWP-total time (PWP-TT-1981), Andersen-Gill (AG-1982), and Wei-Lin-Weissfeld (WLW-1989) models. The literature review also reports the published work on right-censoring and multiple event types in PI models for recurrent events. A limited verification is reported for Qureshi (power-law form) and Vithala (log-linear form) results, applying the parametric Lawless and semi-parametric PWP-GT methods to recurrent data.

Two modeling extensions are examined for the case of multiple event types: multi-dimensional covariate (Lin (1993)) and discontinuous risk-free-intervals (Therneau and Hamilton (1997)).

The proposed research methodology (Chapter 3) addresses four research objectives to answer the following two research questions regarding the PI models robustness for the case of an underlying recurrent failure event process that is NHPP with power-law intensity:

(1) How do the PWP-GT, PWP-TT, AG, and WLW methods compare in performance under right-censoring?

(2) How do the multi-dimensional covariate and discontinuous risk-free-intervals methods perform in estimating the regression coefficients for two failure types (major and minor)?

Four research objectives are raised:

(1) Examine the semi-parametric PI models robustness as a function of right-censoring severity measured by BIAS, MAD, and MSE. The special case of common baseline intensity function (PWP-TT and WLW models) is investigated to compare with the AG model.

(2) Examine the robustness of the four reliability estimates (PWP-GT, PWP-TT, AG, and WLW) as a function of right-censoring severity, for the special case of a stationary counting process. BIAS, MAD, and MSE are employed to measure the robustness of the three event-specific PI models (PWP-GT, PWP-TT, and WLW), while the common baseline model (AG) estimates the general covariate effect.

(3) Examine multi-dimensional covariate modeling to deal with two types of complete (uncensored) recurrent events.

(4) Examine risk-free-intervals within an NHPP process where there are two event types (major and minor) and the overhaul duration following a major failure is substantial.

The methodology includes plans for generation of simulated data sets, design of experiments, and robustness measurements. After the investigation and comparison of the four semi-parametric PI models, PWP-GT is proven the best event-specific model in handling recurrent data with power-law intensity function under right-censoring, and thus is chosen to further investigate the second research question regarding multiple failure types.

This structure of this dissertation is organized as follows. Chapters 1-3 present introduction, literature review, and the proposed research methodology, respectively. Chapter 3 proposes the four research objectives motivated by the two primary research questions and provides a plan/method for each objective/question. Chapters 4-5 investigate the four semi-parametric PI models under right-censoring for an NHPP and HPP, respectively. Chapter 6 addresses the discontinuous risk-free-intervals problem. Chapter 7 studies the covariate PI modeling to handle recurring failure events with two failure types. Chapter 8 summarizes the conclusions from Chapters 4 to 7. Appendix I provides for other relevant charts and tables from Chapter 4 (right-censoring) for singular failure type and from Chapter 7 for two failure types (major and minor). Appendix II provides for the other relevant charts and tables from Chapter 6 for the overhaul duration/maintenance interval problem. Appendixes III to VI present the programming codes to perform the four methodologies/plans regarding the four research problems raised from Chapters 4 to 7. Appendix VII (Glossary) is also provided for the definition of each terminology used in this study.

2 Literature Review

A system may be classified as either non-repairable or repairable. Consumer electronics provides good examples of non-repairable systems and replacement. The aircraft industry provides good examples of repairable system maintenance. A repairable system can be restored from failures to perform a desired function by repair actions other than replacement of the entire system (Ascher and Feingold (1984)). This chapter reviews the literature on reliability assessment for both non-repairable and repairable systems. Non-repairable systems produce single-event failure data, and repairable systems produce recurring-event failure data. In a non-repairable system, a unit is replaced when a failure occurs that renews the hazard rate function each time. However, the unit in a repairable system is repaired rather than replacement of the unit when a failure occurs, and thus can fail two or more times. If it is a successful repair, the intensity function is improved to the degree between as-bad-as-old (minimal repair) and as-good-as-new (replacement).

2.1 Single-event models for non-repairable systems

This section reviews the literature on non-repairable systems with four divisions: 1) Parametric survival models 2) Maximum likelihood estimators, 3) Semi-parametric Cox models, and 4) competing risk models. The subsection of parametric survival models introduces several commonly encountered distributions in reliability. Lawless (1982) has reviewed the parametric method to obtain maximum likelihood estimators using the Newton-Raphson iterations. A semi-parametric Cox proportional hazards (PH) model is used to obtain the parameter estimator through the partial maximum likelihood function. The combination of a decreasing failure rate and an increasing failure rate produces a bathtub function achieved by using the competing risk models.

2.1.1 Parametric survival models

The reliability information provided in this chapter for each distribution contains: density function $f(t)$, reliability function $R(t)$, and the hazard function $h(t)$. This section reviews several commonly encountered distributions in reliability: Exponential, Weibull, Extreme value, Gamma, Lognormal, and Log-logistic distributions, followed by a numerical illustration of plots for each distribution.

Many applications, including customer arrivals, bank service time, and machine breakdowns, have been modeled using the exponential distribution. Exponential distribution displays the memoryless property of constant hazard rate and underlies the Homogeneous Poisson Process (HPP). Figure 2.1 illustrates a constant hazard function based on the exponential distribution with parameter 0.5. The distribution of the interarrival times for an HPP counting process follows an exponential distribution. The density function $f(t)$, reliability function $R(t)$, and the hazard function $h(t)$ are as follows:

$$f(t) = \lambda \exp(-\lambda t)$$
$$R(t) = \exp(-\lambda t)$$
$$h(t) = \lambda,$$

where $\lambda =$ hazard rate.

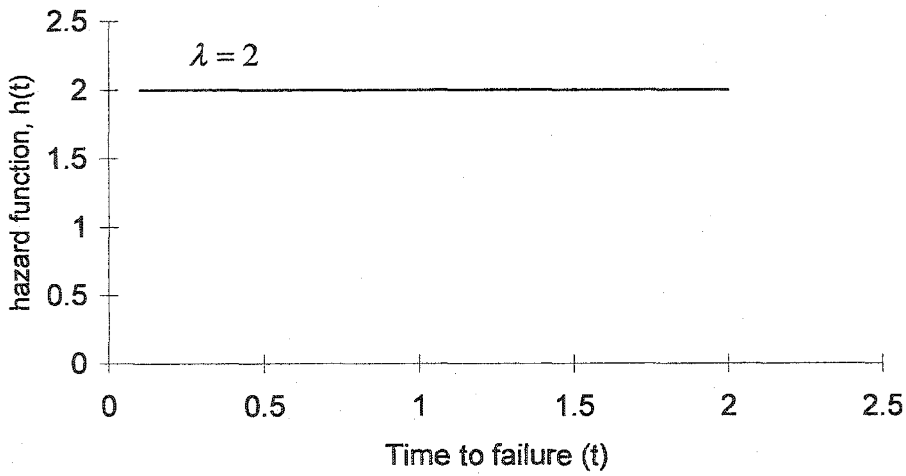


Figure 2.1. Exponential hazard function with $\lambda = 2$

The Weibull distribution is a generalization of the exponential distribution, and capable of modeling a constant, strictly increasing, and strictly decreasing hazard functions. The Weibull functions are:

$$f(t) = \frac{\delta}{\upsilon}(\frac{t}{\upsilon})^{\delta-1}\exp(-\frac{t}{\upsilon})^{\delta}, t > 0$$

$$R(t) = \exp[-(\frac{t}{\upsilon})^{\delta}], t > 0$$

$$h(t) = \frac{\delta}{\upsilon}(\frac{t}{\upsilon})^{\delta-1}, t > 0,$$

where

$\delta =$ shape parameter,

$\upsilon =$ scale parameter.

When $\delta = 1$, the Weibull distribution becomes the Exponential distribution. Figure 2.2 presents three cases of a Weibull hazard function: (δ, υ) = (0.5,5), (1,5), and (3,5).

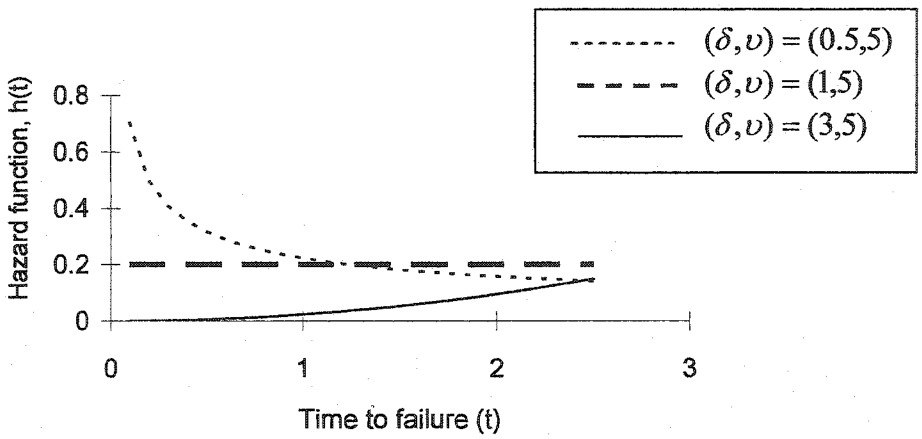


Figure 2.2. Weibull hazard function with combination of δ and υ, where υ=5

The extreme value probability model is an asymptotic distribution, which originates from $Y_{\max} = \max(Y_1, Y_2, ..., Y_n)$ or $Y_{\min} = \min(Y_1, Y_2, ..., Y_n)$, where $Y_1, Y_2, ..., Y_n$ denote random variates. Applications of extreme value model in industry include the distribution of the smallest extreme value (e.g., breaking strength) or the distribution of the largest extreme value (e.g., maximum load). Examples of the extreme value model application in the reliability field include corrosion level and breaking strength. Figure 2.3 is a hazard function plot that demonstrates the distribution for the smallest extreme value.

$Type\mathrm{I}$

$$f(t) = \frac{1}{\theta}\exp[-\exp[\frac{t-\gamma}{\theta}]\exp[\frac{t-\gamma}{\theta}]]$$

$$F(t) = 1 - \exp[-\exp[\frac{t-\gamma}{\theta}]], -\infty < t < \infty, \theta > 0$$

$$h(t) = \frac{1}{\theta}\exp[\frac{t-\gamma}{\theta}]$$

*Type*II

$$f(t)=\frac{\delta}{\theta}\exp[-[-\frac{t-\gamma}{\theta}]^{-\delta}][-[\frac{t-\gamma}{\theta}]^{-\delta-1}]$$

$$F(t)=1-\exp[-[-\frac{t-\gamma}{\theta}]^{-\delta}],-\infty<t<\gamma,\theta>0,\delta>0$$

$$h(t)=-\frac{\delta}{\theta}[-[\frac{t-\gamma}{\theta}]^{-\delta-1}$$

*Type*III

$$f(t)=\frac{\delta}{\theta}\exp[-[\frac{t-\gamma}{\theta}]^{\delta}][-[\frac{t-\gamma}{\theta}]^{\delta-1}]$$

$$F(t)=1-\exp[-[\frac{t-\gamma}{\theta}]^{\delta}],\gamma<t<\infty,\theta>0,\delta>0$$

$$h(t)=-\frac{\delta}{\theta}[\frac{t-\gamma}{\theta}]^{-\delta-1},$$

where
$\gamma=$ location parameter (minimum life),
$\theta=$ scale parameter (characteristic life),
$\delta=$ shape parameter.

The type III Extreme value distribution is the three-parameter Weibull distribution, and the setting of $\theta=0$ in the type III Extreme value distribution yields the two-parameter Weibull model. Figure 2.3 presents three cases of a Type I smallest value of extreme value hazard function: (γ,θ) = (1,5), (2,5), and (3,5).

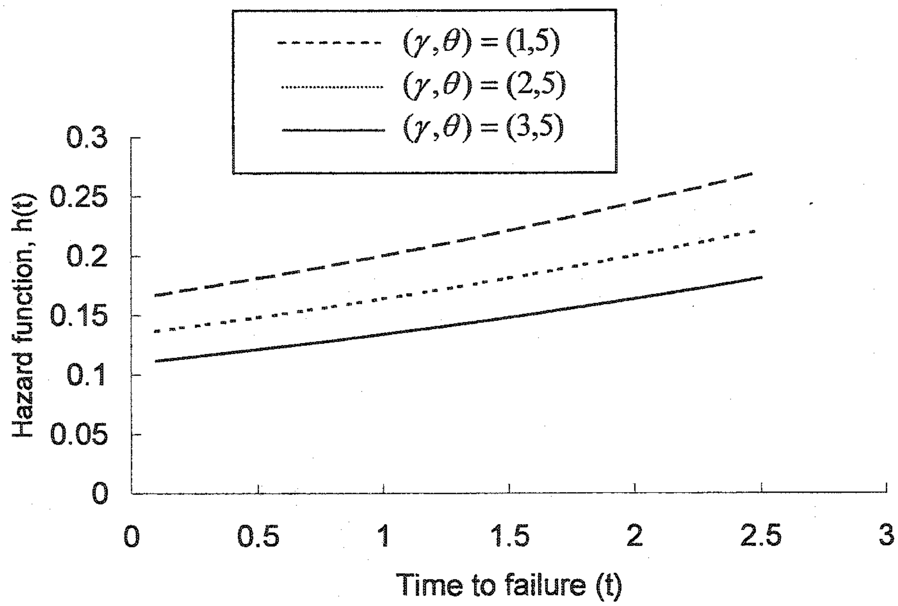


Figure 2.3. Type I smallest value of extreme value hazard functions

The Gamma distribution can model constant, increasing and decreasing hazard rates by controlling the shape parameter setting. The gamma function is defined as $\Gamma(k)=(k-1)!$, where k is an integer. The exponential case arises when parameter η equals to 1. The hazard function (Figure 2.4) is monotonically increasing when η is greater than 1, and monotonically decreasing when η is less than 1. Both Weibull and gamma have a constant hazard function when the shape parameter is set to 1. However, as the shape parameter is greater than one (a strictly increasing hazard function), the Weibull distribution has a faster deterioration rate of hazard function than the Gamma distribution. Figure 2.4 presents three cases of a gamma hazard function: (η,γ) = (0.5,2), (1,2), and (2,2).

$$f(t) = \frac{\lambda(\lambda t)^{\eta-1}\exp[-\lambda t]}{\Gamma(\eta)}, \eta, \lambda, t > 0$$

$$R(t) = 1 - I(t) = 1 - \frac{1}{\Gamma(\eta)} \int_0^x u^{\eta-1}\exp[-u]du$$

$$h(t) = \frac{\lambda(\lambda t)^{\eta-1}\exp(-\lambda t)\Gamma(\eta)^{-1}}{1 - I(\eta, \lambda t)}$$

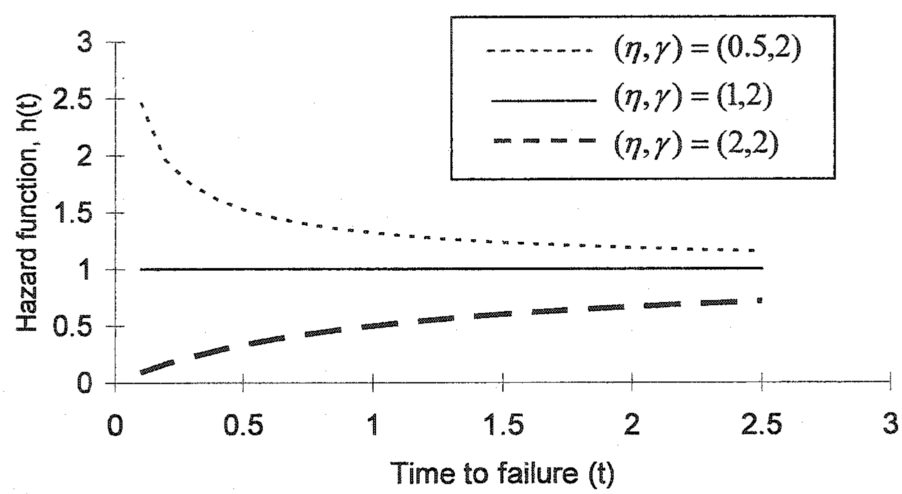


Figure 2.4. Gamma hazard function with different shape parameters

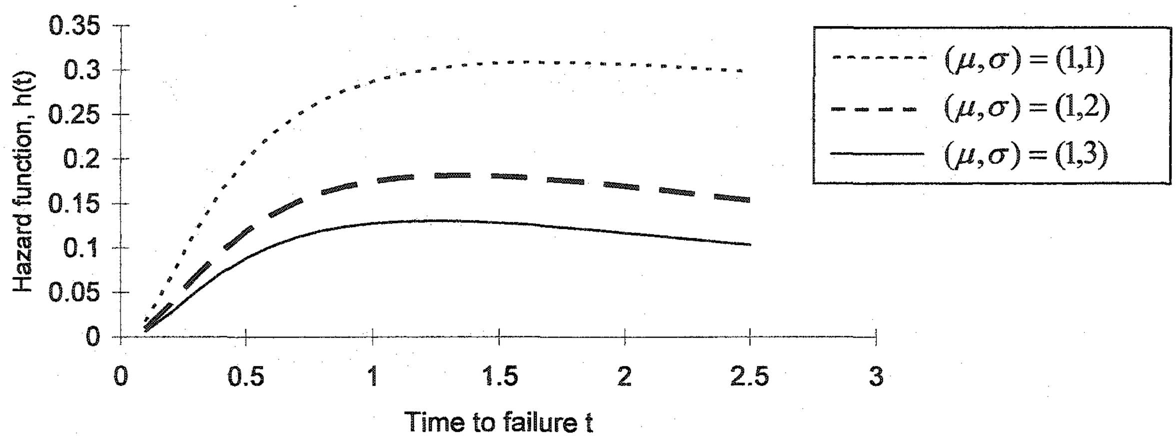
The negative range of time values for the normal Gaussian distribution makes it less suitable for application in reliability. The Lognormal distribution also has merit in reliability modeling of failure mechanisms that are synergistic (multiplicative). The Lognormal model (Figure 2.5(a): (μ,σ) = (1,1), (1,2), and (1,3) and Figure 2.5(b): (μ,σ) = (1,20), (10,20), and (100,20)) relates to the normal distribution in that $z = LN(t)$. A variable t is lognormally distributed if $z = LN(t)$ is normally distributed, where LN denotes the natural logarithm. The Lognormal functions are:

$$f(t)=\frac{1}{\sigma t\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2\right], t>0$$

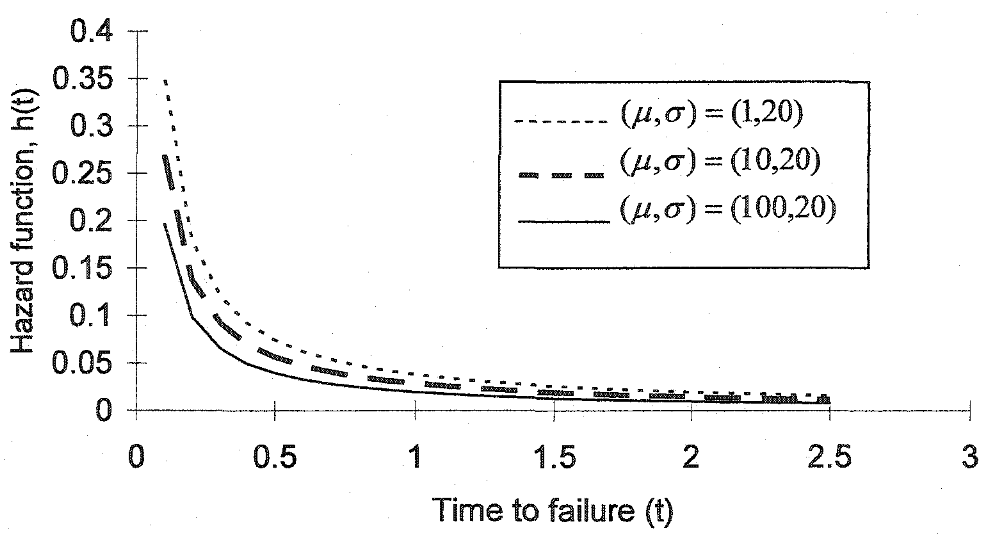
$$R(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_t^{\infty}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2\right]dz, t>0\, where \quad z=\ln t$$

$$h(t)=\frac{\frac{1}{t\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]}{-\Phi(z)}, t>0\, where \quad z=\frac{\ln t-\mu}{\sigma},$$

where
μ = mean parameter,
σ = standard deviation parameter.



(a) Three levels of variance σ^2 with μ=1



(b) Three levels of mean μ with σ=20

Figure 2.5(a)-(b) Lognormal distribution with parameters (μ,σ)

Although the Log-logistic model has simple functional expressions, the hazard function is able to handle multiple stages of life-cycle failure patterns (see Figure 2.6). When p is equal to or greater than 1, the hazard function of the log-logistic model is a monotonically decreasing function. However, there is an unique property in the log-logistic model; when p is less than 1, the hazard function will increase from 0 to a peak point, and then decrease monotonically. The Log-logistic functions are:

$$f(t) = \frac{\lambda p(\lambda t)^{p-1}}{[1+(\lambda t)^p]^2}, p, \lambda, t > 0$$

$$R(t) = 1 - S(t) = 1 - \frac{1}{1+(\lambda t)^p}$$

$$h(t) = \frac{\lambda p(\lambda t)^{p-1}}{1+(\lambda t)^p}.$$

Figure 2.6 presents three cases of a log-logistic hazard function: (p, λ) = (2.5,20), (1,20), and (0.5,20).

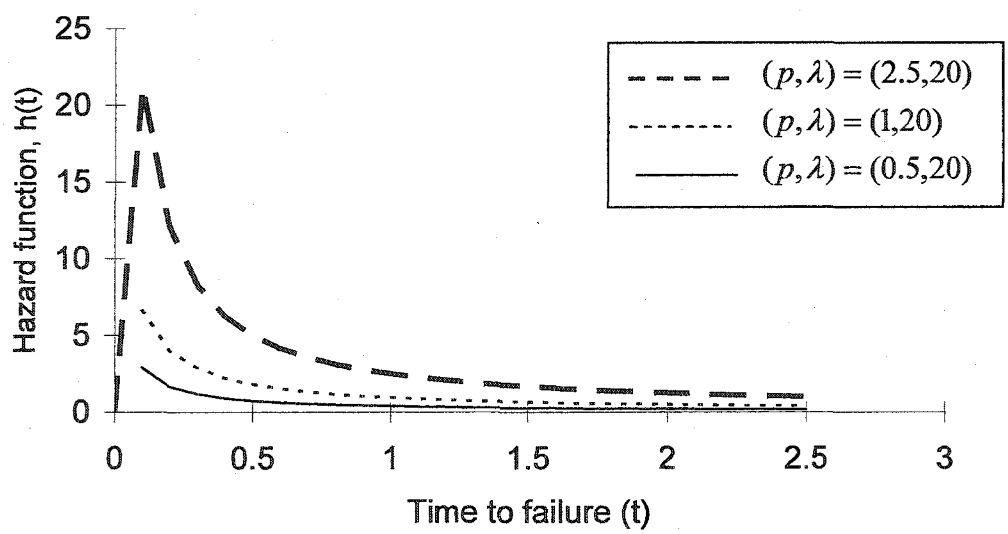


Figure 2.6. Log-Logistic hazard function with parameters (p, λ)

2.1.2 Maximum likelihood estimators

This section reviews the maximum likelihood estimator (m.l.e.) for both Exponential and Weibull underlying distributions proposed by Lawless. The density function of the exponential distribution (with parameter $\theta = 1/\lambda$) is:

$f(t) = \theta^{-1} e^{-t/\theta}, t > 0$.

The point estimator $\hat{\theta}$ can be produced by the maximum likelihood method. The following derivation of the point estimator is taken from Lawless (1982). Suppose $t_1 \le t_2 \le ... \le t_n$ are n samples drawn from an exponential distribution representing the first and only failure time for each unit in the sample. Assume all t_i are independent and identically distributed (i.i.d.). The likelihood is expressed as follows.

$$L(\theta) = \frac{1}{\theta^n} \exp\left(-\sum_{i=1}^{n} \frac{t_i}{\theta} \right).$$

The log likelihood is:

$$\log L = -n \log \theta - \frac{\sum_{i=1}^{n} t_i}{\theta}.$$

The maximum likelihood estimator is obtained by taking the derivative of the log likelihood function, setting equal to zero, and solving for $\hat{\theta}$:

$$\frac{d \log L}{d\theta} = -n \times \frac{1}{\theta} + \frac{\sum_{i=1}^{n} t_i}{\theta^2}.$$

To obtain $\hat{\theta}$, let $\frac{d \log L}{d\theta} = 0$.

$$\hat{\theta} = \frac{\sum_{i=1}^{n} t_i}{n}.$$

Likewise, if the density function for the two-parameter (δ, υ) Weibull is

$$f(t) = \frac{\delta}{\upsilon}(\frac{t}{\upsilon})^{\delta-1}\exp(-\frac{t}{\upsilon})^{\delta}, t > 0.$$

The Weibull parameters (δ, υ) are transformed into the extreme value parameters (b, u) by

$$b = 1/\delta$$
$$u = \log(\upsilon).$$

The maximum likelihood method produces two point estimators $(\hat{\delta}, \hat{\upsilon})$ for the shape and scale parameters (δ, υ) through maximizing the likelihood. The derivation of the point estimators is taken from Lawless (1982). Suppose $t_1 \le t_2 \le ... \le t_r$ are n samples drawn from a Weibull distribution representing the first and only failure time for each sample. Assume all t_i are independently and identically distributed. After the logarithm transformation $X = \log(t)$, the joint likelihood function $L(u, b)$ can be expressed by

$$L(u,b) = \frac{1}{b^r}\exp\left(\sum_{i=1}^{r}\frac{x_i - u}{b} - \sum_{i=1}^{r}\exp\frac{x_i - u}{b}\right).$$

The log likelihood is

$$\log L = -r\log b + \left(\sum_{i=1}^{r}\frac{x_i - u}{b} - \sum_{i=1}^{r}\exp\frac{x_i - u}{b}\right).$$

The maximum likelihood estimators are obtained by taking the partial derivatives of the log likelihood function, setting each equal to zero, and solving for $\hat{u}$ and $\hat{b}$.

$$\frac{\partial \log L}{\partial u} = -\frac{r}{b} + \frac{1}{b}\sum_{i=1}^{r}\exp(\frac{x_i - u}{b})$$

$$\frac{\partial \log L}{\partial b} = \frac{-r}{b} - \sum_{i=1}^{r}\frac{(x_i - u)}{b^2} + \sum_{i=1}^{r}\frac{x_i - u}{b^2}\times\exp(\frac{x_i - u}{b}).$$

Estimators $\hat{u}$ and $\hat{b}$ can be expressed as

$$e^{\hat{u}} = \left[\frac{1}{r}\sum_{i=1}^{r}\exp\left(\frac{x_i}{\hat{b}}\right)\right]^{\hat{b}}$$

$$\frac{\sum_{i=1}^{r} x_i \exp(\frac{x_i}{b})}{\sum_{i=1}^{r}\exp(\frac{x_i}{\hat{b}})} - \hat{b} - \frac{1}{r}\sum_{i=1}^{r} x_i = 0.$$

The Newton-Raphson iteration is employed to obtain the mle's.

2.1.3 Semi-parametric Cox models

The Proportional hazards (PH) lifetime model is used to account for covariate effects for lifetime data. Cox (1972) developed the proportional hazards method for the two-sample problem that introduced explanatory variables. The PH model is the product of a baseline hazard function and an exponential link function composed of explanatory variables, also called concomitant or covariate variables. The baseline hazard functions and covariates are discussed in this section, following the PH theories and applications. A Weibull PH model is then used to illustrate the proportional hazards model.

Under the proportionality assumption, the ratio of the hazard functions $h_A(t \mid z_A) / h_B(t \mid z_B)$ of two sample units, A and B, is constant over time, where A and B represent two levels of a covariate to form two strata of a population. Let sample A have a series of failure times ($T_{A(i)}$) and its estimated mean time between failures ($MTBF_{A(i)}$). Likewise, sample B has ($T_{B_{(i)}}$) and ($MTBF_{B(i)}$). The plots of $\log(MTBF)$ vs. $\log(T)$ for both samples A and B shall present two straight and parallel lines if the proportionality assumption holds.

The definition of a hazard rate based on Cox is the instantaneous failure rate (hazard rate) between t and $t+\Delta t$ under the condition that this individual has survived after time point t, which can be expressed as

$$h(t) = \lim_{\Delta t \to 0+} \frac{P(t \le T < t+\Delta t \mid T \ge t)}{\Delta t}.$$

The explanatory variables $\tilde{\mathbf{z}} = (z_1, z_2, \ldots, z_p)$ are included in the form of an exponential link function in the regression model:

$$h(t;\tilde{z}) = \exp(\tilde{z}\tilde{\beta})h_0(t),$$

where $h_o(t)$ is a baseline hazard function.

$\tilde{\boldsymbol{\beta}}$ is a vector of regression coefficients corresponding to the vector of explanatory variables $\tilde{\mathbf{z}}$. Derived from a conditional likelihood method, the $\tilde{\boldsymbol{\beta}}$ estimator can be obtained from the score vector $\mathbf{U}(\beta)$ and observed information matrix $\mathbf{I}(\beta)$ by solving the following two equations through the Newton-Raphson iteration (Cox (1972)).

$$\mathbf{U}_\xi(\beta) = \frac{\partial L(\beta)}{\partial \beta_\xi} = \sum_{i=1}^{k} (z_{\xi\, i} - \frac{\sum z_{\xi\, l} \exp(z_l \beta)}{\sum \exp(z_l \beta)})$$

$$\mathbf{I}_{\xi\eta}(\beta) = -\frac{\partial^2 L(\beta)}{\partial \beta_\xi \partial \beta_\eta} = \sum_{i=1}^{k} \left(\frac{\sum z_{\xi\, l} z_{\eta l} \exp(z_l \beta)}{\sum \exp(z_l \beta)} - \frac{\sum z_{\xi\, l} \exp(z_l \beta)}{\sum \exp(z_l \beta)} \times \frac{\sum z_{\eta l} \exp(z_l \beta)}{\sum \exp(z_l \beta)} \right).$$

The main purpose of the PH model is to investigate the relationship between the distribution of failure time t_i and covariate variables $\tilde{\mathbf{z}}$, where the regression coefficients $\widetilde{\boldsymbol{\beta}}$ measure the covariate effects. The most common case to assume baseline hazard function $h_o(t)$ in a parametric method is a Weibull form, which leads the PH model in the following equation (Leemis (1995) and Lawless (1982)).

$$h(t;\widetilde{\mathbf{z}}) = k \times \lambda^k \times t^{k-1}\Psi(\widetilde{z}) = \frac{\delta}{\upsilon_0}\left(\frac{t}{\upsilon_0}\right)^{\delta-1} \exp\left(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}}\right).$$

Covariates stratify the population, and the associated regression coefficients $\widetilde{\boldsymbol{\beta}}$ represent the covariate effect. The procedures to estimate $\boldsymbol{\beta}$ can be approached in two ways, parametric and semi-parametric. The analytical procedure is termed as parametric when the baseline function $h_o(t)$ is specified (e.g., Weibull); otherwise, if $h_o(t)$ is left as arbitrary, the methodology is termed semi-parametric. From the time of Cox's foundational work, PH model covariates have been used to evaluate the effects of innovative treatment protocals in clinical trials. Covariates can be classified into four categories: external, internal, constant, and time-varying. An external covariate is determined in advance and is not affected by the treatment, whereas an internal covariate is affected by the treatment. Qureshi (1991) proposed a hypothetical clinical case: a time-varying covariate $z(t)$ takes values of 0,1,2,3,4,5,6 to represent the health situation of a subject who is an AIDS infected patient. $z(t) = 0$ means no clinical evidence is provided, while $z(t) = 6$ represents the death of a subject. The values of time-varying covariates $z(t)$ change during the period of observation, while constant covariates remain at a fixed value that does not change throughout the experiment.

To illustrate this Weibull PH model, the process to derive the score vector and information matrix for a parametric Weibull PH model, provided by Lawless, is summarized below.

Let T be the observed lifetime, and let $y = \log T$.

The p.d.f. of y, given z is

$$f(y \mid z) = \frac{1}{\sigma}\exp\left[\frac{y-\mu(z)}{\sigma} - \exp(\frac{y-\mu(z)}{\sigma})\right], -\infty < y < \infty .$$

Let $\mu(z) = z\beta$,

and $y = z\beta + \sigma A$.

The survival function of y, given z is

$$S(y \mid z) = \exp\left[-\exp(\frac{y_i - z_i\beta}{\sigma})\right],$$

where A is a standard extreme value distribution with p.d.f.

$f(A) = \exp(A - \exp(A))$.

The likelihood function with observed and censored data is

= p.d.f. of observed data$\times$survival function of censored data, which is

$$L(\beta,\sigma) = \prod_{i\in D}\frac{1}{\sigma}\exp\left[\frac{y_i - z_i\beta}{\sigma} - \exp(\frac{y_i - z_i\beta}{\sigma})\right]\prod_{i\in C}\exp\left[-\exp(\frac{y_i - z_i\beta}{\sigma})\right],$$

where (D,C) is a (log lifetime, log censoring time).

It can be simplified as

$$L(\beta,\sigma) = \left(\frac{1}{\sigma}\right)^r \times \exp\left[\sum_{i\in D}\frac{y_i - z\ \beta}{\sigma} - \sum_{i\in D}\exp(\frac{y_i - z_i\beta}{\sigma}) - \sum_{i\in C}\exp(\frac{y_i - z_i\beta}{\sigma})\right],$$

where r is the number of observed data and ($n-r$) is the number of censored data.

After logarithm transformation, the likelihood function becomes

$$\log L = l = -r\log\sigma + \sum_{i\in D}\frac{y_i - z_i\beta}{\sigma} - \sum_{i=1}^{n}\exp(\frac{y_i - z_i\beta}{\sigma}).$$

Since $A_i = \dfrac{y_i - z_i\beta}{\sigma}$,

$$\log L = -r\log\sigma + \sum_{i\in D} A_i - \sum_{i=1}^{n}\exp A_i .$$

Assume that the covariates are z_i, for all $i = 1,2,\ldots,p$ corresponding to regression coefficients β_i, for all $i = 1,2,\ldots,p$.

The first and the second derivatives of l with respect to β,σ are

$$\frac{\partial\log L}{\partial\beta_l} = \frac{-1}{\sigma}\sum_{i\in D} z_{il} + \frac{1}{\sigma}\sum_{i=1}^{n} z_{il}\times\exp A_i$$

$$\frac{\partial\log L}{\partial\sigma} = \frac{-r}{\sigma} - \sum_{i\in D}\frac{y_i - z_i\beta}{\sigma^2} + \sum_{i=1}^{n}\frac{y_i - z_i\beta}{\sigma^2}\exp A_i$$

$$= \frac{-r}{\sigma} - \frac{1}{\sigma}\sum_{i\in D} A_i + \frac{1}{\sigma}\sum_{i=1}^{n} A_i\exp A_i$$

$$\frac{\partial^2\log L}{\partial\beta_l\beta_s} = \frac{-1}{\sigma^2}\sum_{i=1}^{n} z_{il} z_{is}\exp A_i$$

$$\frac{\partial^2\log L}{\partial\sigma^2} = \frac{r}{\sigma^2} + \frac{2}{\sigma^2}\sum_{i\in D} A_i - \frac{2}{\sigma^2}\sum_{i=1}^{n} A_i\exp A_i - \frac{1}{\sigma^2}\sum_{i=1}^{n} A_i^2\exp A_i$$

$$\frac{\partial^2\log L}{\partial\beta_l\partial\sigma} = \frac{1}{\sigma^2}\sum_{i\in D} z_{il} - \frac{1}{\sigma^2}\sum_{i=1}^{n} z_{il}\exp A_i - \frac{1}{\sigma^2}\sum_{i=1}^{n} z_{il} A_i\exp A_i .$$

The score vector at θ is defined $\mathbf{U}_i(\theta) = \dfrac{\partial\log L}{\theta}$, where θ is a $1\times p$ matrix. Set the score vector equal to zero; that is, $\dfrac{\partial\log L}{\beta} = 0, \dfrac{\partial\log L}{\sigma} = 0$.

The score vector $\mathbf{U}(\theta)$ at θ can be expanded as a Taylor series given the first guess at $\hat{\theta}$ is θ_0, the starting point of Newton-Raphson iteration. The first order of the Taylor series will become

$$\mathbf{U}(\theta) \cong U(\theta_0) + G(\theta_0)(\theta - \theta_0),$$

where $\mathbf{G}(\theta)$ is the ($p+1$) by ($p+1$) observed information matrix, with elements

$$\mathbf{G}_{ij}(\theta) = \frac{\partial^2 \log L}{\partial \theta_i \theta_j},$$

e.g., $$\mathbf{G}_{ij}(\theta) = \begin{bmatrix} \frac{\partial^2 \log L}{\partial \beta_l \partial \beta_s} & \frac{\partial^2 \log L}{\partial \beta_l \partial \beta \sigma} \\ \frac{\partial^2 \log L}{\partial \beta_s \partial \sigma} & \frac{\partial^2 \log L}{\partial \sigma^2} \end{bmatrix}.$$

To obtain the m.l.e's. (maximum likelihood estimator), we set the score vector equal to zero.

$$\mathbf{U}(\theta) = 0 = U(\theta_0) + G(\theta_0)(\theta - \theta_0)$$

$$\hat{\theta} \cong \theta_0 - G(\theta_0)^{-1} U(\theta_0).$$

Iteratively, the next guess will be generated through the same procedures until the termination rule is met. The termination rule may be defined as $\Delta\theta$ <10^{-5}. As for any m.l.e.'s, the estimates $\hat{\beta}$ are assumed to be distributed asymptotically normal.

2.1.4 Competing risk models

Leemis (1995) discusses the competing risks (CR) model for combining multiple distributions to achieve a bathtub-shaped hazard function. A typical example from Leemis is to use a DFR (decreasing failure rate) Weibull distribution to model manufacturing defect failures and an IFR (increasing failure rate) Weibull distribution to model wear-out failures. The combination of these two distributions, producing a bathtub-shaped hazard function, is the advantage of the competing risks model. The distribution of the failure time random variable (T) of the competing risks model is subject to k competing risks, which can be expressed as $T = \min\{X_1, X_2, \ldots, X_k\}$. The CR model is composed of the net and crude probabilities through the k competing failure risks. Suppose $X_1, X_2, ..., X_k$ represent k net lives and $Y_1, Y_2, ..., Y_k$ represent k crude lives for each cause. Actual observed lifetime $T = \min\{X_1, X_2, ..., X_k\}$. Set the k components in series.

The net probability of failure from risk k in $[a,b)$ is $q_k(a,b) = P[a \le X_k < b \mid X_k \ge a]$, whereas the crude probability of failure from risk k in $[a,b)$ is $Q_k(a,b) = P[a \le X_k < b, X_k < X_j, for \ \ all \ \ j \ne i \mid T \ge a]$. The net probability considers single dimension of k^{th} risk. However, the crude probability model considers X_k in the presence of other components simultaneously.

2.2 Recurring-event models for repairable systems

The purpose of maintenance is to restore a system into some previous state of reliability. Maintenance actions can be classified in two major categories: corrective and preventive. Corrective maintenance (CM) restores a failed system to operating condition, whereas preventive maintenance (PM) reduces the risk of operation system failure. The failure event process for a repairable system is typically modeled as a Non-homogeneous Poisson Process (NHPP). Let λ represent the degree by which the system reliability has been recovered to a reference state. The value of λ varies from 0 to 1 depending on the repair types, such as minimal repair (λ=0, as-bad-as-old), imperfect repair (0<λ<1), and replacement (λ=1, as-good-as-new) (Usher et al. (1998), Pham and Wang (1996), Kijima (1989), Kijima et al. (1988)). Imperfect maintenance restores the system to the status of somewhere between as-good-as-new (perfect maintenance or renewals) to as-bad-as-old (minimal repair). Pham and Wang (1996) provided a thorough survey of literature on imperfect maintenance.

This section reviews the two primary approaches to model the recurring events (NHPP data) for a repairable system: Parametric stochastic processes and proportional intensity (PI) models (including parametric Lawless and semi-parametric regression methods (PWP (Prentice-Williams-Peterson), AG (Andersen-Gill) and WLW (Wei-Lin-Weissfeld) models)).

2.2.1 Parametric stochastic processes

The parametric stochastic process section reviews the concepts of a counting process and NHPP. Failure data in a repairable system are commonly modeled as a stochastic process. Two NHPP PI functions are illustrated to demonstrate the parametric stochastic processes. A discussion session of a stationary process (HPP), instantaneous mean time between failures (IMTBF), and cumulative mean time between failures (CMTBF) follows.

A counting process $N(t)$, $t \geq 0$ shall satisfy the following criteria, according to Ross (1993).

1. $N(t) \geq 0$.
2. $N(t)$ belongs to integer set.
3. If $s < t$, then $N(s) \leq N(t)$.
4. For $s < t$, $N(t) - N(s)$ equals the number of events that have occurred in the interval (s,t).

A counting process satisfies the independent increments condition if the number of events in a certain interval (t_1,t_2) is independent of the number of events in interval (t_1+s, t_2+s). The Poisson process is the most common assumption in reliability studies of repairable systems. According to Ross (1993), a counting process is said to be a Poisson process with rate λ when

1. $N(t)=0$.

2. The counting process satisfies stationary and independent increments.

3. The number of events in any interval of length t is Poisson distributed with mean $\lambda \times t$. For all $s,t \geq 0$,

$$P\{N(t+s)-N(s)=n\}=e^{-\lambda t}\frac{(\lambda t)^n}{n!}, n=0,1,\ldots$$

4. $P\{N(h)=1\}=\lambda \times h+o(h).$

5. $P\{N(h)\geq 2\}=o(h).$

When the failure rate $\lambda(t)$ of a Poisson process is not constant but varies in terms of time, it becomes an NHPP with intensity function $\lambda(t)$, $t \geq 0$. A definition of NHPP is described in the following.

1. $N(0)=0$.

2. $N(t)$ has independent increments.

3. $P\{N(t+h)-N(t)\geq 2\}=o(h).$

4. $P\{N(t+h)-N(t)=1\}=\lambda(t)h+o(h).$

To simulate an NHPP with intensity function $\lambda(t)$ is to generate a sequence of random variables. Ross (1993) presented a method to generate the first t time units of a Poisson process with intensity function $\lambda(t)$, $t \geq 0$, where interarrival times $X_1, X_2, \ldots, X_n$ from distribution $F \xrightarrow{i.i.d.}$ independent increments stopping at $N=\min\{n: X_1+\ldots+X_n>t\}$. The conditional distribution of X_i is conditioned on $X_1, X_2, ..., X_n$. This conditional probability F_x observes the independent increment property with

$\overline{F}_x(t) = p\{$ 0 events in $(x, x+t)$ | event at x $\}$

$= p\{$ 0 events in $(x, x+t)$ by independent increments $\}$

$$= \exp\left\{-\int_0^t \lambda(x+y)dy\right\}$$

$$f_x(t) = \lambda(x+t)\exp\left\{-\int_0^t \lambda(x+y)dy\right\}$$

$$h_x(t) = \frac{f_x(t)}{\overline{F}_x(t)} = \lambda(x+t).$$

The event times $X_1, X_2, ..., X_{i-1}$ can be simulated by (1) generating X_1 from F_0 and (2) $\hat{X}_1$ to a value generated from F_{x_1} is equal to $\hat{X}_2$, and then (3) adding $\hat{X}_2$ to a value generated from $F_{\hat{x}_2}$ is equal to $\hat{X}_3$, and so on.

Two illustrations of performing the parametric stochastic process follow. In these illustrations, the underlying distribution is assumed to follow a power-law or a log-linear form. Assume the failure process follows an NHPP, where the baseline intensity function is specified as a power-law form; the PI function can be expressed as

$$\lambda(t;\tilde{z}) = \delta \times t^{\delta-1}\exp(\tilde{z}\tilde{\beta}),$$

where δ is the shape parameter, $\tilde{z}$ is the covariate vector, and $\tilde{\beta}$ is the regression coefficient vector that measures the covariate effect. When δ <1, it is a DROCOF (decreasing rate of occurrence of failures); when δ >1, it is an IROCOF (increasing rate of occurrence of failures); and when $\delta = 1$, it is a constant ROCOF (rate of occurrence of failures).

In an NHPP, the theoretical IMTBF in a power-law intensity function is described by the equation

$$IMTBF_{n+1}(t_n) = \left(\upsilon_{z_i} \times \delta \times t_n^{\delta-1}\right)^{-1},$$

where z_i represents a covariate variable, δ and υ are the shape and scale parameters, and n denotes the failure count. Likewise, assume the failure process follows an NHPP, where the baseline intensity function is specified as a log-linear form. The PI function can be expressed as

$$\lambda(t;\tilde{z}) = \exp(\mu + \theta \times t)\exp(\tilde{z}\tilde{\beta}),$$

where μ, θ are the parameters in the log-linear intensity function, $\tilde{z}$ is the covariate vector, and $\tilde{\beta}$ is the regression coefficient vector. An NHPP becomes an HPP when the intensity rate is a constant. For instance, in the case when the baseline intensity function follows a power-law form, the counting process becomes an HPP when the shape parameter δ is equal to 1. Likewise, in the case of a log-linear form, parameter $\theta = 0$ leads to an HPP.

Qureshi (1991) reported the relationship of instantaneous mean time between failures and intensity functions. IMTBF is defined as the derivative of failure time with respect to the expected number of failures (Patrick (1991) and Ascher and Feingold (1984)):

$$IMTBF = \left(\frac{dE[N(T)]}{dt}\right)^{-1},$$

where
$N(t)$ = the number of failures in $(0,t]$,

$E[N(t)]$ = the expected number of failures in $(0,t]$.

A related measurement, ROCOF (rate of occurrence of failures) is defined as the reciprocal of IMTBF, which is

$$ROCOF = \frac{1}{IMTBF} = \frac{dE[N(t)]}{dt}.$$

Likewise, Qureshi (1991) reported the relationship of cumulative mean time between failures and intensity functions. CMTBF is defined as the mean time between failures per event, which is calculated as (Patrick (1991) and Ascher and Feingold (1984))

$$CMTBF = \frac{t}{E[N(t)]},$$

where

$N(t)$ = the number of failures in $(0,t]$,

$E[N(t)]$ = the expected number of failures in $(0,t]$.

The event count stratifies the population into strata from event to event, and creates possibly different intensity functions associated with the event count. The Andersen and Gill (1982), AG method, considers recurring events as a counting process in which each occurrence is independent from other occurrences. As a result, event count does not play a role in the AG method since recurring events are assumed to be independent. However, the PWP method utilizes an event count and stratifies recurring events by blocking effect. For instance, samples remain in stratum i until event $(i+1)$, then move into stratum $(i+1)$.

2.2.2 Proportional intensity models

In discussing the PI models, two methods are reviewed: parametric Lawless and semi-parametric. There are various ways to estimate the parameters of the

NHPP-PI model. The more accurate method is the parametric Lawless (1982). Soroudi (1990), Landers and Soroudi (1991), Qureshi (1991), and Qureshi et al. (1994) summarized the case of a single constant covariate in implementing the parametric Lawless method, where the underlying baseline hazard function follows the power-law "Weibull" form. Likewise, Vithala (1994) and Landers et al. (2001) dealt with an exponential log-linear baseline hazard function. Both works are highly relevant to this research characterizing other semi-parametric PI models. When the underlying distribution is unknown, the semi-parametric PI model is preferred. Thus, the relevant formulas are summarized.

When the baseline intensity function is specified as power-law, the parametric PI function can be expressed as follows.

$$\lambda(t;\widetilde{\mathbf{z}}) = \delta \times t^{\delta-1} \exp(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}}),$$

where δ is the shape parameter.

$z_0 = 1$, $\upsilon = \exp(\beta_0)$ has been included into the exponential link function $\exp(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}})$.

The log likelihood function for the parametric model is

$$L(\delta,\widetilde{\boldsymbol{\beta}}) = N \times U \times \log\delta + (\delta-1)\sum_{i=1}^{U}\sum_{j=1}^{N}\log t_{ij} + \sum_{i=1}^{U} N \times \widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}} - \sum_{i=1}^{U} t_i^{\delta} \times N \times \exp(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}}).$$

The maximum likelihood estimate of δ is adapted by Qureshi (1991) from Lawless (1987), as shown below.

$$\hat{\delta} = \frac{-(N \times U)}{\sum_{i=1}^{U}\sum_{j=1}^{N}\log\left(\frac{t_{ij}}{t_{iN}}\right)},$$

$t_{iN} = \sum_{j=1}^{N} t_{ij}$, where N represents the observed number of failures.

The special case of single constant covariate adapted by Qureshi (1991) from Lawless (1987) yields the score vector U_0, U_1 and information matrix with elements $G_{00}, G_{01}, G_{10}, G_{11}$ as follows.

$$U_0 = \sum_{i=i}^{U} N - \sum_{i=1}^{U} t_{iN}^{\delta} \exp(\beta_0 + z_1\beta_1)$$

$$U_1 = \sum_{i=1}^{U} N \times z_1 - \sum_{i=1}^{U} t_{iN}^{\delta} \times z_1 \times \exp(\beta_0 + z_1\beta_1)$$

$$G_{00} = -\sum_{i=1}^{U} t_{iN}^{\delta} \exp(\beta_0 + z_1\beta_1)$$

$$G_{01} = G_{10} = -\sum_{i=1}^{U} t_{iN}^{\delta} \times z_1 \times \exp(\beta_0 + z_1\beta_1)$$

$$G_{11} = -\sum t_{iN}^{\delta} \times z_1^2 \times \exp(\beta_0 + z_1\beta_1).$$

The formula to derive $IM\hat{T}BF$ is in terms of $\hat{\delta}$ and $\hat{\upsilon}_0, \hat{\upsilon}_1$, where $\hat{\delta}$ is recursively derived by the Newton-Raphson method. The formula to calculate $IM\hat{T}BF$ is (Lawless (1987)):

$$\hat{\upsilon}_0 = \exp(\hat{\beta}_0), \hat{\upsilon}_1 = \exp(\hat{\beta}_0 + \hat{\beta}_1).$$

For the power-law NHPP,

$$IM\hat{T}BF(t_n) = \left(\hat{\upsilon} \times \hat{\delta} \times t_n^{\hat{\delta}-1}\right)^{-1}.$$

Likewise, given the baseline intensity function specified as log-linear, the parametric proportional intensity function follows (Cox and Lewis (1966)):

$$\lambda(t;\widetilde{\mathbf{z}}) = \exp(\mu + \theta \times t)\exp(\widetilde{\mathbf{z}\boldsymbol{\beta}}).$$

Let $z_0 = 1$; then $\exp(\mu) = \exp(\beta_0)$. The parameter μ will be included into the exponential function as $\exp(\widetilde{\mathbf{z}\boldsymbol{\beta}}) = \exp(z_0\beta_0 + z_1\beta_1 + \ldots z_k\beta_k)$, and the proportional intensity function becomes $\lambda(t;\widetilde{\mathbf{z}}) = \exp(\theta \times t)\exp(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}})$. The log likelihood function for the Lawless parametric model is

$$L(\theta,\boldsymbol{\beta}) = \theta\sum_{i=1}^{m}\sum_{j=1}^{n_i} t_{ij} + \sum_{i=1}^{m} n_i \times \mathbf{z}_i \times \boldsymbol{\beta} - \frac{1}{\theta}\sum_{i=1}^{m}(e^{\theta \times T_i} - 1)e^{\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}}},$$

where t_{ij} is the observed failure time corresponding to sample unit i and failure count j.

$$T_i = \sum_{j=1}^{n_i} t_{ij},$$

where

m represents the number of sample units,
n_i is the total failure count corresponding to unit i.

The Maximum likelihood estimator of θ can be obtained by setting

$\frac{\partial \log L}{\theta} = 0, \frac{\partial \log L}{\beta_0} = 0$ (Vithala (1994), adapted from Lawless (1987)):

$$\hat{\theta} = \frac{n - e^{\beta_0} \times \sum_{i=1}^{m} T_i}{\sum_{i=1}^{m}\sum_{j=i}^{n_i}(T_i - t_{ij})}.$$

In the special case of single constant covariate, which Vithala presented, the score vector $(U_0, U_1)'$ and information matrix with elements $G_{00}, G_{01}, G_{10}, G_{11}$, (Vithala (1994), adapted from (Lawless (1987)):

$$U_0 = \sum_{i=1}^{U} n_i - \frac{1}{\theta}\sum_{i=1}^{U}(e^{\theta T_i} - 1)\exp(\beta_0 + z_{i1}\beta_1)$$
$$U_1 = \sum_{i=1}^{U} n_i z_{i1} - \frac{1}{\theta}\sum_{i=1}^{U}(e^{\theta T_i} - 1)z_{i1}\exp(\beta_0 + z_{i1}\beta_1)$$

$$G_{00} = -\frac{1}{\theta}\sum_{i=1}^{U}(e^{\theta T_i} - 1)\exp(\beta_0 + z_{i1}\beta_1)$$

$$G_{01} = G_{10} = -\frac{1}{\theta}\sum_{i=1}^{U}(e^{\theta T_i} - 1)z_{i1}\exp(\beta_0 + z_{i1}\beta_1)$$

$$G_{11} = -\frac{1}{\theta}\sum_{i=1}^{U}(e^{\theta T_i} - 1)z_{i1}^2\exp(\beta_0 + z_{i1}\beta_1).$$

Utilizing the score vector and information matrix, $\hat{\beta}_0, \hat{\beta}_1$ are calculated, for the scale estimates $\hat{\mu}_0, \hat{\mu}_1$ in the two strata defined by the single covariate

$\hat{\mu}_0 = \hat{\beta}_0, \hat{\mu}_1 = \hat{\beta}_0 + \hat{\beta}_1$.

The formula to derive $IM\hat{T}BF$ is in terms of $\hat{\theta}$ and $\hat{\mu}_0, \hat{\mu}_1$, where $\hat{\delta}$ is recursively derived by the Newton-Raphson method. The formula to calculate $IM\hat{T}BF$ is given by Vithala (1994), adopted from Lawless (1987):

$IM\hat{T}BF(t_n) = e^{-(\hat{\mu}_i + \hat{\theta} \times t_n)}, i = 0,1$.

Semi-parametric PI models have been widely cited since PWP (1981) proposed this model in the biomedical studies. Multivariate failure time analysis has been extensively applied in medical research to determine what factors are critical to the survival pattern for patients. Extending the Cox PH model for single event data, PWP created the semi-parametric PWP model to estimate the intensity function in each stratum corresponding to the recurring event count. Many researchers have utilized the PWP model and some have extended the PWP to similar models based on different assumptions. The AG model (1982) by Andersen and Gill and the WLW model (1989) by Wei, Lin, and Weissfeld are widely cited in the literature. In reliability and maintainability engineering applications, a number of authors have applied the semi-parametric proportional intensity (hazards) model, for example, Ansell and Phillips (1989), Ansell and Phillips (1990), Landers and Soroudi (1991), Qureshi et al. (1994), Ansell and Phillips (1997), Landers et al. (2001), Ansell et al. (2001), and Ansell et al. (2003). A collection of the PI model applied to different industries includes: marine gas turbine engines (Asher, 1983), semiconductor, electrical, and pipeline industries (Ansell and Phillips, 1997), U.S. Army main battle tank (Landers et al., 2001), water supply industry (Ansell et al., 2001, 2003), etc. Ascher (1983) illustrated the use of the PWP model for analysis of reliability for marine gas turbine engines. Ascher and Feingold (1984) suggested application of the PWP model in the field of reliability engineering. Dale (1985) applied the PWP approach to simulated data for a reliability growth program with design improvements implemented after each of the five stages, resulting in a DROCOF. Wightman and Bendall (1986) and Bendall et al. (1991) cited the PWP model and advised caution in application for engineering studies. An introduction of PWP, AG, and WLW models is as follows.

A risk set with event-specific baseline hazard is called a restricted risk set (Kelly and Lim (2000)). Since the PWP gap time (PWP-GT) and total time (PWP-TT) models both have event-specific baseline hazards, the intensity function of the first event is merely decided by these subjects that have recorded first events. Likewise, for both PWP-GT and PWP-TT, the intensity of the third event depends only on the subjects that have experienced the second event and then experience the third event.

The PWP-GT model has the gap time risk interval (Figure 2.7(a)), while the PWP-TT model has a counting process risk interval (Figure 2.7(b)). The counting process has the same length of elapsed time as does the gap time. The SAS code to program the total time model is written as $(start, end, status(I))$ for each observation, where $status(1)$ denotes failure time and $status(0)$ denotes censored time. For instance, the recurring events occur at time 4, 7, 12 with the follow-up time 15, the records in the SAS database will be shown as $(0,4,1)$, $(4,7,1)$, $(7,12,1)$, and $(12,15,0)$. Likewise, the SAS dataset for the PWP-GT model is written as $(gap\_time, status(0))$. Thus, the records in the SAS dataset will be shown as $(4,1)$, $(3,1)$, $(5,1)$, and $(3,0)$.

Qureshi (1991) applied the PWP-GT model to engineering reliability. Failure data simulated as an NHPP with a power-law intensity function were generated by the Blanks and Tordon (1987) algorithm based on the setting of 20 sample units divided into two groups by a single covariate (CLASS). Each sample unit generated complete data (no censoring) with10 failures. Thus, the sample size associated with failure counts remained the same among all event strata. Using the PWP-TT method was equivalent to using the WLW method, since they both shared the same failure data sets. Unrestricted baseline hazard was employed in this model, which yielded a common baseline hazard for each failure intensity function in each stratum defined by failure count. Using a gap time dependent variable in the PWP model, the failure time data were sorted by failure and then by gap time in descending order before implementing the PH regression analysis (as PHREG in SAS program).

Likewise, Vithala (1994) conducted a simulation experiment with failure data generated by the Law and Kelton (1991) method. The recurring failure events followed an NHPP with a log-linear form. The parameter settings are summarized as follows: 60 sample units, 30 sample units/CLASS=0, 30 sample units/CLASS=1, 10 failures/unit. Since the sample size with respect to failure count was equal among the 60 sample units, using the PWP-TT method was equivalent to using the WLW method. Vithala utilized event-specific baseline hazards in the PWP-GT model, which resulted in event-specific regression coefficients corresponding to each stratum defined by failure count in the estimating process. Like Qureshi (1991), gap time was employed as a dependent variable to perform the PH regression analysis. Vithala used the same procedures executed in Qureshi (1991): failure data were sorted by failure counts and next by gap times in a descending order.

Andersen and Gill [2] developed the AG method as an extension of the Cox PH model, to accommodate recurring events in a counting process. The AG method explains general covariate effects (common baseline intensity function in the concept of risk set), since each event count re-starts the failure process, and thus does not feature event-stratifying effects. The risk interval of an AG model follows a counting process associated with recurring events, where recurrences $(N_i^{(n)}, Y_i^{(n)}, Z_i^{(n)})$ are independent and identically distributed replicates of (N, Y, Z), and the probability of the occurrence of two events at a given time is zero. Thus, the risk set of the $(n-1)^{st}$ event is identical as the risk set of the $(n)^{th}$ event. The AG model is defined as

$$\lambda_i^{(n)}(t) = Y_i^{(n)}(\mathrm{t})\lambda_0(\mathrm{t})\exp\{\boldsymbol{\beta} \times \mathbf{z}_i^{(n)}(t)\}$$

, where $Y_i^{(n)}$ is an at-risk indicator and $Y_i^{(n)} = 1$ unless the subject is withdrawn from the study.

A study regarding admissions to psychiatric hospitals for pregnant woman was investigated by using the AG model. Two states in a Markov process are defined as admissions and discharge corresponding to two forces of transition $\alpha_i(t), \mu_i(t)$. Number of visits to psychiatric hospitals, $N_i(t)$, is a counting process with intensity function $\lambda_i(t) = \alpha_i(t)Y_i(t)$. Parity of the woman (number of children) and age are covariates in this study. Three covariates of parity status (parity0, parity2, parity≥ 3, with $\lambda_0(t)$ representing parity1) and two covariates of age range (age≤ 18 and age$\geq$34) are employed and defined as follows.

$$Z_{i1} = \begin{cases} 1 & parity0 \\ 0 & otherwise \end{cases}, Z_{i2} = \begin{cases} 1 & parity2 \\ 0 & otherwise \end{cases}, Z_{i3} = \begin{cases} 1 & parity \geq 3 \\ 0 & otherwise \end{cases}$$

$$Z_{i4} = \begin{cases} 1 & age \leq 18 \\ 0 & otherwise \end{cases}, Z_{i5} = \begin{cases} 1 & age > 34 \\ 0 & otherwise, \end{cases}$$

where i represents subject i.

A Markov process model is considered to analyze admissions to psychiatric hospitals. A time-dependent covariate Z_{i6} is introduced to form a semi-Markov process model, where the covariate is defined as

$$Z_{i6} = \begin{cases} 1 & re-admitted \\ 0 & otherwise. \end{cases}$$

AG (1982) concluded that the oldest women and women with higher parity have the highest intensity of admissions to psychiatric hospitals during pregnancy.

The risk for each recurrence remains the same throughout the entire study unaffected by earlier events in the AG method, while the intensity function is affected by earlier recurrences in the PWP method. However, when a subject is withdrawn from the study, the subject does not contribute any information to the latter intensity functions based on the AG method, whereas the subject in the WLW method remains in the analysis. Since the AG model has the counting process of risk interval (Figure 2.7(c)), the data representation is written as $(start, end, status(0))$ for each observation. For instance, the recurring events occur at times 4, 7, 12 with the follow-up time 15, so the records in the SAS database will be shown as $(0,4,1)$, $(4,7,1)$, $(7,12,1)$, and $(12,15,0)$. The program stops executing when the condition: $(start < end)$ fails. The AG method has common baseline intensity, while PWP-TT has event-specific baseline intensity.

In terms of risk intervals, the AG method utilizes a counting process, whereas Qureshi (1991) and Vithala (1994) both adopt a gap time formulation. The duration of the dependent variable in the AG method is collected by gap time but the risk interval is not affected until the end of the previous event (due to the property of a counting process), while Qureshi (1991) and Vithala (1994) renew the risk interval to time zero at occurrence of each event, resulting in an entry into each new stratum. The AG method and Qureshi (1991) both have common baseline hazard, while Vithala (1994) has event-specific baseline hazards.

WLW (1989) proposed a marginal method, expanded from the conditional PWP method, in dealing with recurrent failure data. Compared to the PWP method, the WLW method has greater or equal risk set, depending on the sample size associated with the failure count. The PWP method estimates the intensity function by considering the subjects having a complete history of previous recurring events, while the WLW method additionally considers the subjects that have been withdrawn from the observation. The subjects that have been censored are still in the risk set; thus, contributing influence on events that are followed after the censoring time. The risk set of each subject using the WLW method remains the same regardless of complete data or censoring events since a subject is still at risk when the subject has been withdrawn from the experiment.

WLW (1989) in a bladder cancer study examined treatment effects by using the PWP and WLW models about placebo and thiotepa therapies for tumor patients. This bladder cancer example collects four recurrence times of tumors $T_1 \sim T_4$ corresponding to four marginal proportional hazards models. Rather than fitting each T_i one model at a time, WLW fits four marginal models in one analysis, simultaneously. This example has two response variables {failure time and censoring status}, three covariates {treatment, tumour number, tumour size}, and four recurrences of time. For the k^{th} failure type and the i^{th} failure event count, the hazard function $\lambda_{ki}(t)$ in WLW is assumed to take the form below:

$$\lambda_{ki}(t) = \lambda_{k0}(t)\exp\left\{\boldsymbol{\beta}_k{}' \times \mathbf{z}_{ki}(t)\right\} t \geq 0 \text{ ,}$$

where $\lambda_{k0}(t)$ is an unspecified baseline hazard function and $\boldsymbol{\beta}'_k = \left(\beta_{1k},\ldots,\beta_{pk}\right)'$ is a vector of failure-specific regression parameters. $\mathbf{z}_{ki}(t)$ denotes a $p \times 1$ vector of covariates for the i^{th} subject at time t with respect to the k^{th} type of failure, expressed as $\mathbf{z}_{ki}(t) = (z_{1ki}(t), z_{2ki}(t), \ldots z_{pki}(t))'$.

Let $\mathbf{X}_{ki}$ represent the failure time of the i^{th} subject for the k^{th} type of failure and let C_{ki} represent the censoring time. $\widetilde{X}_{ki}$ are observation values of $\mathbf{X}_{ki}$, where $\mathbf{X}_{ki} = \min\{\widetilde{X}_{ki}, C_{ki}\}$. The indicator variable Δ_i is utilized for determining the event as a failure or censoring. Let $\Delta_i = 1$, when $\mathbf{X}_{ki} = \widetilde{X}_{ki}$; otherwise $\Delta_i = 0$. Key assumptions for the WLW method are: (1) $\mathbf{X}_{ki} \perp C_{ki}$, i.e., the failure and censoring times are independent of each other; (2) $(\mathbf{X}_i, \boldsymbol{\Delta}_i, \mathbf{Z}_i)$ are i.i.d. random vectors, where $\mathbf{Z}_i$ represent covariates and i represents event count; and (3) the regression coefficients $\hat{\boldsymbol{\beta}}_i$ follow a normal distribution with mean $\overline{\boldsymbol{\beta}}_i$ denoted

$$(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \ldots, \hat{\beta}_k) \xrightarrow{iid} Normal\ (\overline{\beta}_1, \overline{\beta}_2, \overline{\beta}_3, \ldots, \overline{\beta}_k).$$

WLW (1989) examined a two-sample problem to compare the WLW method with two other approaches, the AG and PWP (both gap time and total time). Random variates $(t_1, t_2] = (start, end]$ are generated from a bivariate exponential distribution with a correlation parameter θ that governs the correlation between t_1 and t_2. Two gap times U and V from a bivariate exponential distribution represent two time endpoints $(U, U+V]$. The results indicate that the sizes of the Wald tests in the PWP (gap time) model significantly exceed the nominal level when the correlation θ is greater than 0.25, and the sizes in the PWP (total time) model significantly exceed the nominal level at all θ values. The results prove that PWP can be very sensitive to correlation coefficient θ and the assumption of failure time distribution (bivariate exponential in this case).

The WLW method has the total time of risk interval due to the usage of a marginal method, and the code is written as $(total, status(0))$ for each observation. For instance, if the recurring events occur at time 4, 7, 12 with the follow-up time 15, the records in the database will be shown as $(total\_time, status(0))$ $= (4,1)$, $(7,1)$, $(12,1)$, and $(15,0)$. The concepts of risk interval and risk set regarding the WLW method are as follows. The WLW method has total time of risk interval and event-specific baseline intensity of risk set. Total time carries the risk effect of earlier events to the later events for the WLW marginal method. The hazard ratio in the gap time is different from the total time or the counting process. The hazard ratio in the total time is equal to the ratio in the counting process, for they share the same time scale. The partial likelihood is defined as the following (Kelly and Lim (2002)):

$$L(\beta) = \prod_{j=1}^{d} \frac{\lambda(t_{(j)})}{\sum_{k \in R(t_{(j)})} \lambda(t_{(k)})},$$

where

d: uncensored events,

j: the j^{th} specific event,

$R(t(j))$: subjects on the risk interval at time $t(j)$.

In terms of risk interval formulation, the WLW method utilizes a total time formulation, while Qureshi (1991) and Vithala (1994) adopt a gap time formulation. The duration of the failure times in the WLW method is on a total-time basis, whereas in Qureshi (1991) and Vithala (1994) it is on a gap-time basis. The WLW method resets the risk interval (time clock) to zero at each stratum, and so do Qureshi (1991) and Vithala (1994). The WLW method and Vithala (1994) both have event-specific baseline intensity, while Qureshi (1991) has common baseline intensity.

The pattern of the recurrent data can be seen in many areas as remarked by Lin (1994), "Examples in biomedical research are the sequence of tumour recurrences or infection episodes, the development of physical symptoms or diseases in several organ systems, the occurrence of blindness in the left and right eyes, the onset of a genetic disease among family members, the initiation of cigarette smoking by classmates, and the appearance of tumors in littermates exposed to a carcinogen. Examples in other fields include the repeated breakdowns of equipment and systems in engineering reliability, the experiences of different life events by each person in sociological studies, and the purchases of various products by each consumer in marketing research."

PWP (1981) proposed a model that generalizes the proportional hazards model (PH model). This PWP model extends the case of single event to the case of multiple recurrent events (a stochastic process). Cox (1972) proposed the PH model by introducing explanatory variables to analyze the failure time data with censoring. The definition of a hazard rate based on the Cox model is the instantaneous failure rate between t and $t+\Delta t$ under the condition that this individual has survived after time point t, which can be expressed as follows:

$$\lambda(t) = \lim_{\Delta t \to 0+} \frac{P(t \le T < t+\Delta t \mid T \ge t)}{\Delta t}.$$

The main structure of a PH model is the product of an explanatory variables (covariates) $z_1, z_2, \ldots, z_p$ in the exponential form and a baseline intensity function:

$$\lambda(t; \widetilde{\mathbf{z}}) = \exp(\widetilde{\mathbf{z}}\widetilde{\boldsymbol{\beta}})\lambda_0(t).$$

The PWP method relaxes the assumption that the failure process follows a parametric form (e.g., NHPP power-law process). Since the PWP does not specify the baseline intensity function, it only estimates the covariate treatment effects. The two PWP models that represent the PWP-GT and PWP-TT are:

$$\lambda\{t|N(t), \mathbf{Z}(t)\} = \lambda_{0s}(t - t_{n(t)})\exp\{\mathbf{z}(t)\boldsymbol{\beta}_s\}$$

$$\lambda\{t|N(t), \mathbf{Z}(t)\} = \lambda_{0s}(t)\exp\{\mathbf{z}(t)\boldsymbol{\beta}_s\}.$$

The gap time measures elapsed time between any two consecutive events, whereas total time measures time from entry into the experiment (beginning of observation). PWP (1981) concluded that the gap time model usually tends to provide a more precise regression estimator at each failure count compared with the total time model.

Bowman (1996) surveyed and evaluated the AG, PWP, and WLW methods applied to needlestick incidents in veterinary practice. Bowman conducted a simulation based on a bivariate exponential distribution to generate bivariate recurrent events, in order to control the correlation (θ) between recurring events. Bowman utilized the bivariate exponential distribution (T_1, T_2) to generate the consecutive recurring event time $T(n) = T_1(n) + T_2(n)$, where n is the event count. The univariate event time $T(n)$ is composed of $T_1(n)$ and $T_2(n)$ with given correlation θ. The advantage of this type of simulation data makes it possible to manage the correlation of recurring events. In the model evaluation, Bowman evaluated performance of four methods: PWP-GT, PWP-TT, AG, and the WLW models, applied to the GT model as superior and then used it to analyze the needlestick injury data.

Lin (1994) also evaluated the PWP, AG, and WLW methods of Cox regression analysis in multivariate failure time data using a marginal approach. Lin let T_{ik} be the time when the k^{th} type of failure occurs on the i^{th} unit. Lin also let C_{ik} be the corresponding censoring time, $X_{ik} = \min(T_{ik}, C_{ik})$ with the resulting $\Delta_{ik} = I(T_{ik} \le C_{ik})$. The covariate vector for the i^{th} unit with respect to the k^{th} type of failure is $\widetilde{\mathbf{Z}}_{ik} = (Z_{1ik}, ..., Z_{pik})'$. The marginal approach can be expressed in the two forms below addressed by WLW and LWA (Lee-Wei-Amato) accordingly.

WLW : $\lambda_k(t; \mathbf{Z}_{ik}) = \lambda_{0k}(t)\exp\{\mathbf{Z}_{ik}(t)\boldsymbol{\beta}'\}$,

or

LWA: $\lambda_k(t; \mathbf{Z}_{ik}) = \lambda_0(t)\exp\{\mathbf{Z}_{ik}(t)\boldsymbol{\beta}'\}$.

The LWA assumes a common baseline intensity function across all strata defined by the failure type. The partial likelihood functions for β under WLW and LWA, corresponding score vector, and information matrix can be obtained from Lin (1994).

Wei and Glidden (1997) have reviewed the Cox-based methods designed to model recurrent data, and summarized the strengths and weaknesses for each method. In a commentary on the Wei and Glidden paper, Lipschutz and Snapinn (1997) stressed the two concepts of "event times" and "risk sets" as crucial to choosing the appropriate model. First, event elapsed times are related to the total time, gap time, and counting process. The PWP-TT and WLW are modeled by total time, while only PWP-GT is modeled by gap time. The risk interval of the AG model belongs to the counting process class. Intuitively, total (global) times within a subject are highly correlated, with similar indication on the first recurrence and subsequent events. The total time model may indicate large treatment effect throughout the entire study, although the gap time model has indicated little treatment effect beyond a certain recurrence. The counting process concept of the AG method implies each recurrence is not affected by previous events, and does not contribute to future events.

The risk set consists of the subjects at risk for a specified event (e.g., failure). There are three types of risk sets: conditional (e.g., PWP), counting process (e.g., AG), or marginal (e.g., WLW). As a marginal method, the WLW method assumes a subject is at risk regardless of event count until the observation for the subject terminates by censoring. The AG method also provides an index of a general covariate effect, which is expressed by the common baseline intensity (unrestricted risk set). However, a subject in the PWP method has event-specific baseline intensity (restricted risk set), in that the proportional intensity of event k only considers the subjects that have experienced $(k-1)$ events. Lipschutz and Snapinn (1997) suggested guidelines as follows in choosing the appropriate models:

- Use total time, common baseline hazard (unrestricted risk set) when the general effect is of interest.

- Use gap time, event-specific baseline hazards (restricted risk set) when the primary concern is how the treatment will affect the recurring events beyond the first occurrence.

Some Cox-based proportional hazards models are very sensitive to misspecification due to dependence structure that exists among recurring events. Examples include the AG and PWP models (Wei and Glidden (1997)). The misspecification problem causes parameter estimators to become overestimated or underestimated. Kelly and Lim (2000) addressed three ways to deal with misspecification problems: conditional, marginal, and random effects. The conditional method introduces time-variant covariates intended to capture the dependence structure. The marginal method utilizes a robust variance named a "sandwich estimator", where a robust variance is added to the variance of the estimator. The approach of random effects, also named the frailty method, includes a random covariate into the model aimed to induce the dependence structure among the failure events. Kelly and Lim (2000) applied the conditional and marginal approaches to childhood infectious disease cases, and concluded that applying the marginal method (robust variance) is not effective to resolve misspecification problems if any dependence exists.

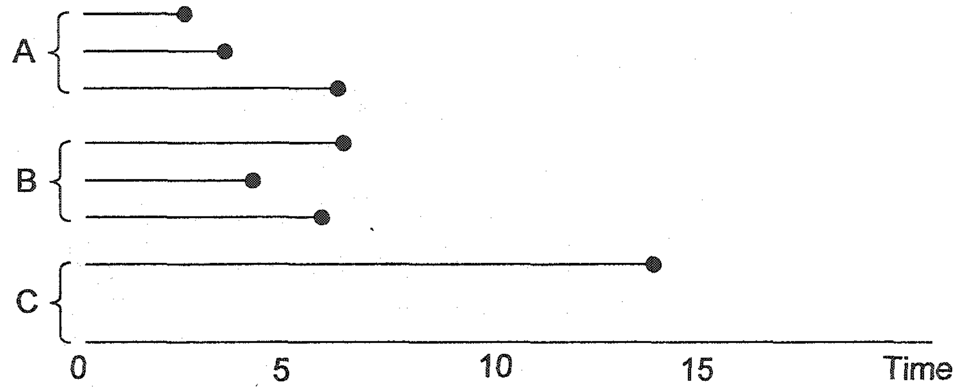
Jiang et al. (1999) investigated the misspecification problem and addressed three potential misspecification factors: (1) neglect of random effects, (2) omitted covariates, and (3) measurement error. They commented that in a special case where there is no measurement error, it will not affect the point estimator. However, the variance adjustment is needed, which can be attained through a sandwich formula. If both errors exist in the model, a double-sandwich formula is derived to adjust the variance. A naïve estimator requires the adjustment to reach a consistent estimator. As for the measurement error associated with covariates, Jiang et al. (1999) illustrated an example in a skin tumor study. Treatment assignment (Se or placebo) and baseline (plasma Se) status are chosen as two covariates. The treatment assignment is accurate without any error, whereas the Se status may result in a measurement error. Other researchers have worked on the robust variance model, such as Lin and Wei (1989), and Therneau and Hamilton (1997), to name a few.

Risk interval can be defined by three formulations: (1) gap time, (2) total time, and (3) counting process. Risk interval determines whether a model is marginal in the total time or conditional in the gap time. The risk interval of any event in total time is not influenced by any previous events, but measures time from entry into the experiment (beginning of observation). However, the risk interval of the gap time begins from the end of last event (Kelly and Lim (2000)). Counting processes use the total time scale and share the same elapsed time as does the gap time model. However, the risk interval starts from the previous event instead of the entry time. Kelly and Lim (2000) illustrated three risk interval formulations shown in Figure 2.7. Three subjects A, B, and C are in the experiment.

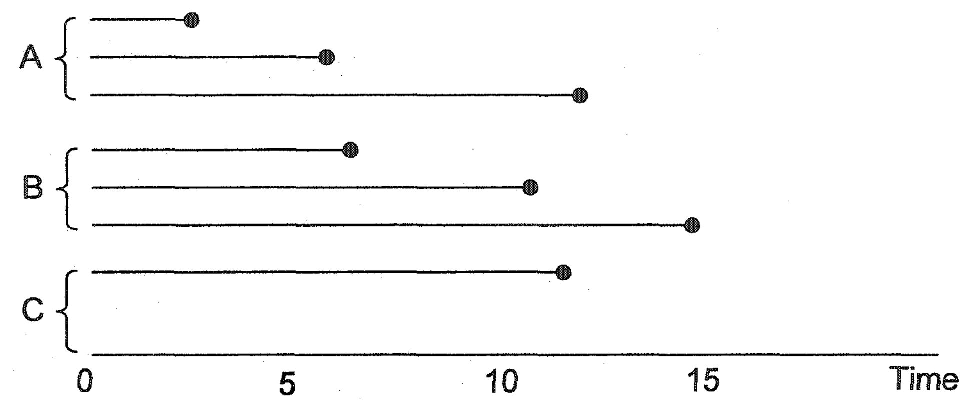
Based on the common or event-specific baseline intensities, the risk set is labeled as either unrestricted or restricted. Kelly and Lim (2000) defined three possible risk sets {(1) unrestricted, (2) restricted, and (3) semi-restricted} in deciding which sample units are at risk of contributing to event k. Table 2.1

summaries the methods Kelly and Lim categorized on a basis of risk set versus risk interval.

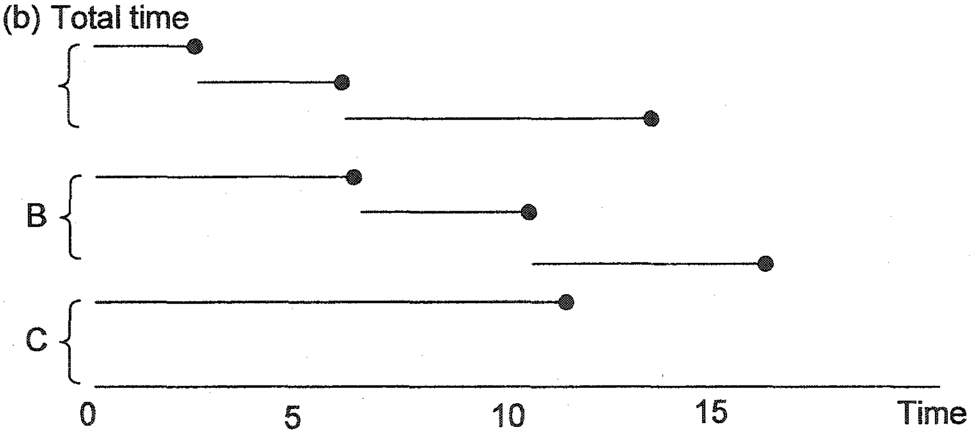
The LWA model, similar to the WLW model, has a common baseline intensity function (Lin (1994)). The model Qureshi (1991) employed in a robustness study of the PWP model can be classified as (risk interval, risk set) = (gap time, common). Likewise, Vithala (1994) may be classified as (risk interval, risk set) = (gap time, event-specific). The PWP-TT (termed as PWP-CP in Kelly and Lim (2000)) is specified as a counting process instead of a total time model, due to the conditionality. PWP-CP is a stratified AG model. A marginal approach, such as the WLW method, employs the total time concept since subjects are at risk since the entry of the experiment.



(a) Gap time



(b) Total time



(c) Counting process

Figure 2.7(a)-(c). Risk interval formulations (Kelly and Lim (2000))

Table 2.1. Taxonomy of risk interval and risk set for each model (Kelly and Lim (2000))

Models	Risk set/ baseline intensity		
Risk interval	Unrestricted/ common	Semi-restricted/ event specific	Restricted/ event specific
Gap time	Qureshi (1991)		PWP-GT (1981) Vithala (1994)
Total time	LWA (1994)	WLW (1989)	
Counting process	AG (1982)		PWP-TT (1981)

2.2.3 Robustness studies

The section reports the results of the researcher's attempts to replicate the work of Qureshi and Vithala. This exercise was beneficial in gaining thorough understanding of their studies and results. It also serves, on a small scale, to validate their results. Most importantly, the studies provided insights regarding extensions of the work that would contribute to the body of knowledge and potential for engineering applications.

The PWP method is an appealing method to model the recurrent failure processes since it does not require specification of a baseline intensity function. Qureshi et al. (1991) extended the research begun by Landers and Soroudi (1991) in a pilot study. Qureshi investigated the robustness of a PWP-GT model for the case of data from a true underlying process that is an NHPP with power-law intensity function. PWP-GT estimates were compared to the true underlying model and to the estimates obtained from the parametric Lawless (1987) method. They concluded that if the baseline intensity function is a power-law form, the Lawless method is preferred to model the recurrent failure processes for constant and moderately IROCOF.

If the true underlying process is NHPP, with power-law intensity function, a heuristic expression may be postulated to state the time to the n^{th} failure (Soroudi (1990)):

$$t_n = \left(\frac{n_t}{\upsilon_{z_i}} \right)^{1/\delta},$$

where z_i represents a covariate variable, and δ and υ are the shape and scale parameters of the power-law form. Note that $E[n(t)] = n_t$ holds since the following are equivalent in a counting process:

- The expected number of failures happened at time $t = E[n(t)]$.
- The n^{th} failure at time $t = n_t$.

Simulated datasets generated from an NHPP with the power-law intensity function are employed as the sample failure times in a recurrent failure process. The aim of this study is to model this recurrent process fitted by two main methods, a parametric Lawless method, and a semi-parametric PWP-GT method, which are discussed in the following. The Lawless method involves the estimation of shape parameter $\hat{\delta}$, intercept and slope regression coefficients $\hat{\beta}_0, \hat{\beta}_1$, and instantaneous mean time between failures $IM\hat{T}BF$. The PWP-GT method involves slope regression coefficient $\hat{\beta}_1$, survival function $\hat{S}$, and mean time to failures $M\hat{T}TF$. The performance measurements for the Lawless or PWP-GT method, compared with the theoretical value of instantaneous mean time between failures $IM\hat{T}BF$ are collected. BIAS, MAD, and MSE are employed as the performance measures to reveal the robustness of the estimating process.

The main purpose of the study is to investigate how well the PWP-GT can estimate the theoretical intensity function. In other words, when the time is specified as t_n (n failure times have been observed), the corresponding time to next failure ($n+1$), $MT\hat{T}F_{n+1}$ is estimated either by the Lawless or PWP methods and compared to the theoretical $MTTF_{n+1}$. The PWP method utilizes the nonparametric Product-Limit estimators, integrating the areas under its estimated survival function, to derive the mean time to next failure.

There is an update of the programming syntax in SAS 8.0 regarding the usage of PHGLM, which has been replaced in the SAS library by PHREG. A blocking option in PHGLM is utilized to allow all subjects to be stratified in each stratum and to obtain an intensity function representing the stratum. PHGLM syntax is replaced by "STRATA" or "BY" statement in PHREG. Since the PHGLM procedure employed in Qureshi's work is renewed to PHREG in Release 8.01 version, the blocking option in PHGLM is changed to the programming statement shown below.

```
PROC PHREG DATA=FILE_NAME;
MODEL FAILURE_TIME=COVARIATE;
STRATA FAILURE_COUNT;
```

FILE_NAME : the file which stores the recurrent data
FAILURE_TIME: recurrent failure time data
COVARIATE: the covariate variable
FAILURE_COUNT: the stratum is defined by failure count.

There is one clarification to Qureshi's work that the blocking option is supposed to employ failure count as a stratification variable. Thus, there is one regression estimate in each stratum, which means a covariate effect is estimated within each stratum, instead of one global covariate regression coefficient across all strata. The global coefficient estimate in Qureshi is actually the strata estimate; the results should be modified to report regression coefficient estimates in all strata defined by event count. Qureshi did in-depth examination of the $\hat{\beta}$ estimates for the special case of an HPP, for which the strata $\hat{\beta}$ are theoretically equal.

If the true underlying process is NHPP with a power-law baseline intensity function, the Lawless method is appropriate to estimate the intensity function $\lambda(t)$. The parametric Lawless method is utilized to obtain three estimators, $\hat{\delta}$, $\hat{\beta}_0$,and $\hat{\beta}_1$ in terms of the seed number, shape parameter δ, baseline scale parameter υ_0, alternate scale parameter υ_1, and sample size. The chosen values of δ, υ_0, and υ_1 in Qureshi's work are taken from the air-conditioning data set for plane #7908 (Proschan (1963)) to base the robustness study in parameter values for a realistic range. Two classes that can be distinguished by covariate effect divide all samples evenly. Experiments on eight combinations of the sample size= 20 and 60 and shape parameter=0.5, 1, 1.5, and 3 are investigated in this research to duplicate the Qureshi estimates of β and mean time to n^{th} failure $(IM\hat{T}BF(n))$, for $n = 1,2,...,10$. Note that each value of $IM\hat{T}BF(n)$ is determined by the average of three replicates in terms of seed numbers 539, 255, and 59. The results of the estimation on, $\hat{\delta}$, $\hat{\beta}_0$, and $\hat{\beta}_1$ are computed. However, only three examples of $\delta =$ 0.5, 1.0, and 1.5 are shown in Table 2.2 for verification purposes. To compare with the Qureshi results, the case of $\delta = 1.5$ is taken from

Qureshi (1991) and presented here. The other two cases of $\delta = 0.5$ and $\delta = 1$ are not listed in Qureshi (1991).

Table 2.2. The results of ($\hat{\delta}$, $\hat{\beta}_0$, $\hat{\beta}_1$) using the Lawless method

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 0.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$			
Seed number	$\hat{\delta}$	$\hat{\beta}_0$	$\hat{\beta}_1$
539	0.53729	-7.539608	2.2746569
255	0.48321	-6.532990	2.0593530
59	0.51619	-7.110656	2.3367614
Average	0.51223	-7.061080	2.2235900

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 1, \upsilon_0 = 0.001, \upsilon_1 = 0.01$			
Seed number	$\hat{\delta}$	$\hat{\beta}_0$	$\hat{\beta}_1$
539	1.07459	-7.539608	2.2746569
255	0.96642	-6.532990	2.0593530
59	1.03237	-7.110656	2.3367614
Average	1.02446	-7.061080	2.2235900

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 1.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$			
Seed number	$\hat{\delta}$	$\hat{\beta}_0$	$\hat{\beta}_1$
539	1.61188	-7.539608	2.2746569
255	1.44963	-6.532990	2.0593530
59	1.54856	-7.110656	2.3367614
Average	1.53669	-7.061080	2.2235900
Average(Qureshi)	1.53670	-7.061100	2.2236000

$IM\hat{T}BF$ can be obtained using the parametric Lawless method from the following formula, where the estimator $\hat{\delta}$ is recursively derived by the Newton-Raphson method and the formula to scale estimates in two classes υ_0, υ_1 are shown below.

$$\hat{\upsilon}_0 = \exp(\hat{\beta}_0),\ \ \hat{\upsilon}_1 = \exp(\hat{\beta}_0 + \hat{\beta}_1)$$

$$IM\hat{T}BF_{n+1}(t_n) = \left(\hat{\upsilon} \times \hat{\delta} \times t_n^{\hat{\delta}-1}\right)^{-1}$$

$$t_n = \left(\frac{n_t}{\upsilon_{z_i}}\right)^{1/\delta}.$$

Table 2.3 presents the results of utilizing the Lawless method to estimate this NHPP in the combination of 20 sample units, $\delta = 1.5$, and $\upsilon_0 = 0.001$, $\upsilon_1 = 0.01$ in two classes.

Table 2.3. The parametric Lawless method

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 1.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$						
			CLASS = 0			
Failure	E(t$_n$)	IMTBF(t$_n$)	$\hat{IMTBF}$ (t$_n$)	e_n	$\lvert e_n \rvert$	e_n^2
1	100.00	66.67	64.07	-0.03902	0.039023	0.001523
2	158.74	52.91	49.99	-0.05518	0.055178	0.003045
3	208.01	46.22	43.24	-0.06450	0.064502	0.004161
4	251.98	42.00	39.01	-0.07106	0.071062	0.005050
5	292.40	38.99	36.02	-0.07612	0.076119	0.005794
6	330.19	36.69	33.74	-0.08023	0.080229	0.006437
7	365.93	34.85	31.93	-0.08369	0.083691	0.007004
8	400.00	33.33	30.44	-0.08668	0.086679	0.007513
9	432.67	32.05	29.19	-0.08931	0.089306	0.007976
10	464.16	30.94	28.11	-0.09165	0.091650	0.008400
BIAS = -0.071793367						
		MAD = 0.071793367				
				MSE = 0.006069119		
BIAS (Qureshi) = -0.0718						
		MAD (Qureshi) = 0.0718				
				MSE (Qureshi) = 0.0061		
			CLASS = 1			
Failure	E(t$_n$)	IMTBF(t$_n$)	$\hat{IMTBF}$ (t$_n$)	e_n	$\lvert e_n \rvert$	e_n^2
1	21.54	14.36	15.80	0.100222	0.100222	0.010044
2	34.20	11.40	12.33	0.081725	0.081725	0.006679
3	44.81	9.96	10.67	0.071050	0.071050	0.005048
4	54.29	9.05	9.62	0.063540	0.063540	0.004037
5	63.00	8.40	8.88	0.057751	0.057751	0.003335
6	71.14	7.90	8.32	0.053044	0.053044	0.002814
7	78.84	7.51	7.88	0.049081	0.049081	0.002409
8	86.18	7.18	7.51	0.045660	0.045660	0.002085
9	93.22	6.90	7.20	0.042652	0.042652	0.001819
10	100.00	6.67	6.93	0.039969	0.039969	0.001597
BIAS = 0.062708426						
		MAD = 0.062708426				
				MSE = 0.004778524		
BIAS (Qureshi) = 0.0627						
		MAD (Qureshi) = 0.0627				
				MSE (Qureshi) = 0.0048		

If the true underlying process is known to be NHPP with a power-law baseline intensity function, the Lawless method is a common method to estimate the intensity function $\lambda(t)$. Under this circumstance, the PWP-GT method provides a way to approach the theoretical intensity function without requiring knowledge of the true underlying process. Utilizing the semi-parametric PWP-GT method to obtain $M\hat{T}TF$ in terms of the seed number, shape parameter δ, and sample size, the results of $M\hat{T}TF$ in the example of $\delta = 1.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$ are summarized in Table 2.4. Note that two classes of the covariate effect divide all samples evenly.

Table 2.4. Average $M\hat{T}TF$ obtained from the PWP-GT method

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 1.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$					
CLASS = 0					
	$M\hat{T}TF$			Average $M\hat{T}TF$	Qureshi results
Failure	539	255	59		
1	56.41	61.19	59.47	59.47	59.47
2	67.22	59.13	62.73	62.73	62.73
3	55.85	40.92	41.99	41.99	41.99
4	46.15	53.53	48.89	48.89	48.89
5	31.18	61.51	46.98	46.98	46.98
6	46.57	37.30	39.41	39.41	39.41
7	37.13	24.17	38.43	38.43	38.43
8	34.66	21.24	29.72	29.72	29.72
9	46.14	35.09	41.18	41.18	41.18
10	25.08	39.14	31.34	31.34	31.34
CLASS = 1					
	$M\hat{T}TF$			Average $M\hat{T}TF$	Qureshi results
Failure	539	255	59		
1	23.92	15.96	20.78	20.22	20.22
2	17.53	13.25	18.38	16.39	16.39
3	6.85	10.79	5.31	7.65	7.65
4	12.86	12.54	8.61	11.34	11.34
5	5.81	13.08	11.63	10.17	10.17
6	16.48	8.46	7.68	10.87	10.87
7	6.30	13.97	7.89	9.39	9.39
8	6.20	9.28	8.32	7.93	7.93
9	10.47	6.67	7.84	8.33	8.33
10	6.04	10.04	5.89	7.32	7.32

$M\hat{T}TF$ can be obtained using the semi-parametric PWP-GT method by implementing the Product-Limit method, which integrates the area under its survival function. Experiments on eight combinations of the sample size= 20, 60 and shape parameter=0.5, 1, 1.5, 3 are investigated in the estimating of mean time to n^{th} failure $(M\hat{T}TF(n))$, for $n = 1,2,...,10$. Note that each value of $M\hat{T}TF(n)$ is determined by the average of three replicates in terms of seed numbers 539, 255, and 59. The results of one example $\delta = 1.5$ and $\upsilon_0 = 0.001,\ \upsilon_1 = 0.01$ are summarized in Table 2.5.

Table 2.5. The semi-parametric PWP method

20 sample units, 10 failures/unit, 10units/class, 2 classes, $\delta = 1.5, \upsilon_0 = 0.001, \upsilon_1 = 0.01$

CLASS = 0

Failure n	$E(t_n)$	$IMTBF(t_n)$	$M\hat{T}TF$	e_n	$\lvert e_n \rvert$	e_n^2
1	100.00	66.67	62.73	-0.05905	0.05905	0.003487
2	158.74	52.91	41.99	-0.20644	0.206439	0.042617
3	208.01	46.22	48.89	0.057674	0.057674	0.003326
4	251.98	42.00	46.98	0.118642	0.118642	0.014076
5	292.40	38.99	39.41	0.010852	0.010852	0.000118
6	330.19	36.69	38.43	0.047479	0.047479	0.002254
7	365.93	34.85	29.72	-0.14722	0.147215	0.021672
8	400.00	33.33	41.18	0.2354	0.2354	0.055413
9	432.67	32.05	31.34	-0.02215	0.022153	0.000491
10	464.16	30.94				

BIAS = 0.00391001

MAD = 0.10054481

MSE = 0.017931775

BIAS (Qureshi) = 0.0039

MAD (Qureshi) = 0.1005

MSE (Qureshi) = 0.0179

CLASS = 1

Failure n	$E(t_n)$	$IMTBF(t_n)$	$M\hat{T}TF$	e_n	$\lvert e_n \rvert$	e_n^2
1	21.54	14.36	16.39	0.141135	0.141135	0.019919
2	34.20	11.40	7.65	-0.32894	0.328938	0.1082
3	44.81	9.96	11.34	0.138705	0.138705	0.019239
4	54.29	9.05	10.17	0.123998	0.123998	0.015376
5	63.00	8.40	10.87	0.294129	0.294129	0.086512
6	71.14	7.90	9.39	0.187975	0.187975	0.035335
7	78.84	7.51	7.93	0.056162	0.056162	0.003154
8	86.18	7.18	8.33	0.159933	0.159933	0.025579
9	93.22	6.90	7.32	0.060107	0.060107	0.003613
10	100.00	6.67				

BIAS = 0.092578441

MAD = 0.165675758

MSE = 0.039615717

BIAS (Qureshi) = 0.0926

MAD (Qureshi) = 0.1657

MSE (Qureshi) = 0.0396

The robustness test is aimed to evaluate the performance of two methods in estimating the instantaneous mean time between failures that come from an

NHPP with underlying power-law intensity function. Three performance measurements are utilized for the comparison of the Lawless and PWP-GT methods in estimating the mean time to the n^{th} failure. The definitions of three performance measurements BIAS, MAD (mean absolute deviation), and MSE (mean squared error) are written in the following.

$$BIAS = \frac{\sum_{n=1}^{n-1} e_n}{n-1}$$

$$MAD = \frac{\sum_{n=1}^{n-1} |e_n|}{n-1}$$

$$MSE = \frac{\sum_{n=1}^{n-1} e_n^2}{n-2},$$

where $e_n = \dfrac{IM\hat{T}BF(t_n) - IMTBF(t_n)}{IMTBF(t_n)}$ in the Lawless method

and $e_n = \dfrac{M\hat{T}TF(t_n) - IMTBF(t_n)}{IMTBF(t_n)}$ in the PWP method.

All combinations of the sample size= 20, 60 and shape parameter=0.5, 1, 1.5, 3 are investigated in this duplication of results to estimate mean time to n^{th} failure, $n = 1,2,...,10$ implemented by the Lawless and PWP methods. Each value of either $IM\hat{T}BF(n)$ or $M\hat{T}TF(n)$ is determined by the average of three replicates of different seed numbers in order to decrease the bias effect. The robustness tests of the estimating methods are summarized in Table 2.6 and Table 2.8. To compare with the Qureshi results, Table 2.7 and Table 2.9 are taken from Qureshi (1991) and listed here.

Table 2.6. The summary of the robustness test using the lawless method

10 failures/unit, 10units/class, 2 classes, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$

		CLASS = 0			CLASS = 1		
U[a]	δ	BIAS	MAD	MSE	BIAS	MAD	MSE
20	0.5	-0.071754	0.071754	0.006063	0.062747	0.062747	0.004784
20	1.0	-0.071754	0.071754	0.006063	0.062747	0.062747	0.004784
20	1.5	-0.071793	0.071793	0.006069	0.062708	0.062708	0.004779
20	3.0	-0.071763	0.071763	0.006064	0.062740	0.062740	0.004783
60	0.5	0.019899	0.040084	0.003308	-0.036847	0.054280	0.004081
60	1.0	0.019899	0.040084	0.003308	-0.036847	0.054280	0.004081
60	1.5	0.019804	0.040075	0.003305	-0.036914	0.054320	0.004087
60	3.0	0.019804	0.040075	0.003305	-0.036914	0.054320	0.004087

[a] U represents the sample size in terms of the number of sample units.

Table 2.7. Qureshi results

10 failures/unit, 10units/class, 2 classes, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$

		CLASS = 0			CLASS = 1		
U[a]	δ	BIAS	MAD	MSE	BIAS	MAD	MSE
20	0.5	-0.0712	0.0712	0.0060	0.0632	0.0632	0.0048
20	1.0	-0.0721	0.0721	0.0061	0.0625	0.0625	0.0047
20	1.5	-0.0718	0.0718	0.0061	0.0627	0.0627	0.0048
20	3.0	-0.0718	0.0718	0.0061	0.0627	0.0627	0.0048
60	0.5	0.0206	0.0401	0.0033	-0.0363	0.0540	0.0040
60	1.0	0.0197	0.0401	0.0033	-0.0370	0.0544	0.0041
60	1.5	0.0193	0.0400	0.0033	-0.0372	0.0545	0.0041
60	3.0	0.0197	0.0401	0.0033	-0.0370	0.0544	0.0041

[a] U represents the sample size in terms of the number of sample units.

Table 2.8. The summary of the robustness test using the PWP method

10 failures/unit, 10units/class, 2 classes, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$

		CLASS = 0			CLASS = 1		
U[a]	δ	BIAS	MAD	MSE	BIAS	MAD	MSE
20	0.5	0.059314	0.180375	0.059257	0.590499	0.619370	0.662950
20	1.0	-0.032945	0.119595	0.023036	0.134369	0.209770	0.059830
20	1.5	0.003910	0.100545	0.017932	0.092578	0.165680	0.039620
20	3.0	0.092401	0.147581	0.030666	0.105328	0.177390	0.040060
60	0.5	0.325973	0.325973	0.279226	0.331310	0.341870	0.236930
60	1.0	0.041010	0.091598	0.012567	0.015725	0.078300	0.008040
60	1.5	0.033088	0.063391	0.007110	-0.026681	0.069790	0.007570
60	3.0	0.082385	0.087137	0.012807	-0.029658	0.050720	0.005210

[a] U represents the sample size associated with sample units.

Table 2.9. Qureshi results

10 failures/unit, 10units/class, 2 classes, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$

		CLASS = 0			CLASS = 1		
U[a]	δ	BIAS	MAD	MSE	BIAS	MAD	MSE
20	0.5	0.0593	0.1804	0.0593	0.5905	0.6194	0.6629
20	1.0	-0.0329	0.1196	0.0230	0.1343	0.2097	0.0598
20	1.5	0.0039	0.1005	0.0179	0.0926	0.1657	0.0396
20	3.0	0.0936	0.1475	0.0310	0.1053	0.1774	0.0401
60	0.5	0.3260	0.3260	0.2792	0.3313	0.3419	0.2369
60	1.0	0.0410	0.0916	0.0126	0.0157	0.0783	0.0080
60	1.5	0.0331	0.0634	0.0071	-0.0267	0.0698	0.0076
60	3.0	0.0821	0.0863	0.0127	-0.0297	0.0507	0.0052

[a] U represents the sample size associated with sample units.

Another crucial contribution from Qureshi (1991) comes from the model confirmation. The case of $\delta = 1$ was assumed and a regression analysis ANOVA was performed by the GLM procedure in the SAS software. The intension of the study is to find the slope and intercept estimates $(\hat{\delta}, \hat{\upsilon}_0, \hat{\upsilon}_1)$ in the case of $\delta = 1$, $\upsilon_0 = \ln(0.001) = 6.9078$, $\upsilon_1 = \ln(0.01) = 4.6051$ as theoretical values. Theoretically, for an HPP ($\delta = 1$), the slope is equal to zero and the failure intensity equation is simplified to

$$\lambda(t) = \upsilon \times \delta \times t^{\delta - 1} = \upsilon .$$

The results taken from Qureshi (1991) are listed in Table 2.10.

Table 2.10. GLM summary (Qureshi (1991))

Units	$\upsilon_0 = \lambda$, $\upsilon_1 = \lambda$	Slope estimate	Intercept estimate	$\hat{\lambda}$	t statistics[b]
	0.001	0.0298(0.0917)[a]	6.5895(0.7665)	0.0014	-0.4153
20	0.010	0.0525(0.0917)	4.3560(0.5563)	0.0128	-0.448
	0.001	0.0206(0.0438)	6.7645(0.3658)	0.0012	-0.3917
60	0.010	-0.0232(0.0438)	4.7514(0.2655)	0.0086	0.5507
	0.001	-0.0118(0.0362)	7.0079(0.3025)	0.0009	0.3309
120	0.010	0.0289(0.0362)	4.4562(0.2196)	0.0116	-0.6785

[a] Estimated standard errors in parentheses
[b] t statistics for H_0 : intercept = $\ln(1/\lambda)$

Vithala (1994) extended the work of Qureshi by investigating the baseline intensity function in a log-linear form. Vithala reached the same conclusions that the PWP-GT model performs well in the case of constant and moderately increasing ROCOF. The research of Qureshi et al. (1991) and Vithala (1994) both confirm the PWP-GT model is a robust method for many important applications, in which the baseline intensity function is unknown.

One correction to the Vithala code needs to be made in order to run the PWP-GT method. On page 157, the code is

```
R=(LOG(THETA*T+ EXP(MU))-MU)/THETA;
```

It should be modified as follows, to allow the case $\theta = 0$ without a divided-by-zero error.

```
IF THETA=0 THEN R=T/EXP (MU);
ELSE R=(LOG(THETA*T+EXP(MU))-MU)/THETA;
```

There are three corrections to the SAS code that Vithala wrote to implement the parametric Lawless method.

1. on page 188, Vithala wrote

```
DATA INSERT;
SET THETA;
RETAIN XY 0;
DROP ITEM T FAILURE CLASS MU Y G;
```
→ delete this line since the variable mentioned here does not exist in the data file theta.

2. on page 189, the code reads

```
DATA MIX;
MERGE MULTIPLY PURGE;
PROC PRINT DATA=MIX;
DROP A H HH V K E;
```
→ delete this line since the variable mentioned here does not exist in the data files multiply or purge.

3. on page 192, it says

```
DATA ALIGNED;
SET ALIGNED;
BETAI0=BETAI;
DROP E BETAI;
```

The code is corrected by two data generating statements shown below.

```
DATA  ALIGNED;
SET IREGRESS;
RETAIN E 0;
E=E+1;
IF E>1 THEN DELETE;

DATA REDUCE;
SET ALIGNED;
BETAI0=BETAI;
DROP E BETAI;
```

If the true underlying process is an NHPP, with log-linear baseline intensity function, a heuristic expression may be postulated to state the time to n^{th} failure (Vithala (1994)):

$$t_n = \frac{1}{\theta} \times Ln\left(\frac{\theta \times n}{e^{\mu_i}} + 1\right), n = 1,2,\ldots,n.$$

where $\mu_i, i = class0, class1$ represents a covariate variable. θ, μ are two parameters in the log-linear form as $e^{\mu+\theta\times t}$.

Note that $E[n(t)] = n_t$ holds since the following are equivalent in a counting process.

- The expected number of failures at time $t = E[n(t)]$.
- The n^{th} failure at time $t = n_t$.

However, for the HPP case (when $\theta = 0$), t_n is not defined in the proceeding equation due to division by zero. The heuristic value of t_n can be obtained from the formula utilized in the case of HPP of Qureshi work when $\delta = 1$ below:

$$t_n = \left(\frac{n_t}{\upsilon_{z_i}}\right).$$

The estimation of intensity functions is approached by two methods: the parametric Lawless and semi-parametric PWP-GT. The Lawless method involves the estimation of the shape parameter $\hat{\theta}$, regression coefficients $\hat{\beta}_0, \hat{\beta}_1$, and instantaneous mean time between failures $IM\hat{T}BF$. The shape parameter θ produces an IROCOF when positive, a constant ROCOF when zero, and a DROCOF when negative. The PWP-GT method involves regression coefficient $\hat{\beta}_1$, survival function $\hat{S}$, and mean time to failure $M\hat{T}TF$.

In an NHPP, the theoretical IMTBF in a log-linear intensity function is derived by the following equations

$IMTBF(t_n) = e^{-(\mu_{z_i} + \theta \times t_n)}$

$IMTBF(t_n) = e^{-(\mu_{z_i})}$, when $\theta = 0$,

where z_i represents a covariate variable and θ and μ_{z_i} are the parameters of the log-linear intensity function.

BIAS, MAD, and MSE are employed as the performance measures to perform robustness tests. The main purpose of the study is to investigate how well the PWP-GT can estimate the theoretical intensity function. In other words, when the time is specified as t_n (n failure times have been observed), the corresponding expected time to next failure ($n+1$), $MT\hat{T}F_{n+1}$ is derived either by the Lawless or PWP-GT methods for comparison with the theoretical $MTTF_{n+1}$.

If the true underlying process is NHPP with a log-linear baseline intensity function, the Lawless method is appropriate to estimate the intensity function $\lambda(t)$. The parametric Lawless method is utilized to obtain three estimators, $\hat{\theta}$, $\hat{\beta}_0$, and $\hat{\beta}_1$ in terms of the seed number, parameter θ, baseline scale parameter μ_0, alternate scale parameter μ_1, and sample size. Two classes distinguished by a covariate divide all samples evenly. Experiments on four combinations of the sample size $U =$ 60, 120 and parameter $\theta =$ 1.2, 2.0 are investigated in this research to duplicate the Vithala estimates of mean time to n^{th} failure ($IM\hat{T}BF(n)$), for $n = 1,2,...,10$. Note that each value of $IM\hat{T}BF(n)$ is determined by the average of three replicates of seed number 539, 255, and 59. The results of the estimation on, $\hat{\theta}$, $\hat{\beta}_0$, and $\hat{\beta}_1$ are shown in Table 2.11. To compare with the Vithala results, the case of $\theta = 1.2, U = 60$ is taken from Vithala (1994) and presented here. The other three cases of $\theta = 2.0, U = 60$, $\theta = 1.2, U = 120$, and $\theta = 2.0, U = 120$ are not listed in Vithala (1994).

Table 2.11. ($\hat{\theta}$, $\hat{\beta}_0$, $\hat{\beta}_1$) obtained from the Lawless method

60 sample units, 10 failures/unit, 30units/class, 2 classes, $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$			
Seed number	$\hat{\theta}$	$\hat{\beta}_0$	$\hat{\beta}_1$
539	1.241880	-7.282879	2.5032100
255	1.275460	-7.476738	2.4971794
59	1.350040	-7.891322	2.5818163
Average	1.289127	-7.550310	2.5274020
Average(Vithala)	1.289130	-7.550890	2.5277650

$IM\hat{T}BF$ can be obtained using the parametric Lawless method from the following formula, where the estimators μ_0, μ_1 and $\hat{\theta}$ are obtained by the Newton-Raphson method and using the following formulas:

$$\hat{\mu}_0 = \hat{\beta}_0, \hat{\mu}_1 = \hat{\beta}_0 + \hat{\beta}_1$$
$$IM\hat{T}BF(t_n) = e^{-(\hat{\mu}+\hat{\theta}\times t_n)}.$$

t_n is heuristically derived from NHPP with a log-linear form, by the expression

$$t_n = \frac{1}{\theta} \times Ln\left(\frac{\theta \times n}{e^{\mu}} + 1\right), n = 1,2,\ldots,n.$$

Table 2.12 has the results of utilizing the Lawless method to model this NHPP in the combination of 60 sample units, $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$ in two classes.

Table 2.12. The parametric Lawless method

60 sample units, 10 failures/unit, 30units/class, 2 classes, $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$

CLASS = 0

n	t_n	IMTBF(t_n)	$I\hat{MTBF}$ (t_n)	e_n	$\lvert e_n \rvert$	e_n^2
1	5.902634188	0.83263407	0.942768	0.1322716	0.132272	0.017496
2	6.479907133	0.416491777	0.4479323	0.0754889	0.075489	0.005699
3	6.817678122	0.277700038	0.2898062	0.0435943	0.043594	0.0019
4	7.057354876	0.208289602	0.2127757	0.0215378	0.021538	0.000464
5	7.243272849	0.166638677	0.1674302	0.0047501	0.00475	2.26E-05
6	7.395184155	0.138869451	0.1376527	-0.008762	0.008762	7.68E-05
7	7.523626394	0.119033338	0.1166474	-0.020045	0.020045	0.000402
8	7.634890059	0.104155733	0.1010608	-0.029714	0.029714	0.000883
9	7.733032869	0.092583953	0.0890505	-0.038165	0.038165	0.001457
10	7.820825524	0.083326335	0.0795215	-0.045661	0.045661	0.002085

BIAS = 0.020106
MAD =0.041592
MSE = 0.00355
BIAS (Vithala) = 0.020698
MAD (Vithala) =0.041681
MSE (Vithala) = 0.003581

CLASS = 1

n	t_n	IMTBF(t_n)	$I\hat{MTBF}$ (t_n)	e_n	$\lvert e_n \rvert$	e_n^2
1	3.992219332	0.826410878	0.88376	0.0693954	0.069395	0.004816
2	4.566373547	0.414928835	0.421588	0.0160482	0.016048	0.000258
3	4.903101776	0.277004333	0.273129	-0.013992	0.013992	0.000196
4	5.142256551	0.207897967	0.200666	-0.034786	0.034786	0.00121
5	5.327861145	0.166387916	0.157965	-0.050622	0.050622	0.002563
6	5.479563451	0.138695258	0.129906	-0.063372	0.063372	0.004016
7	5.607856365	0.118905331	0.110104	-0.074021	0.074021	0.005479
8	5.719008014	0.104057711	0.095405	-0.083149	0.083149	0.006914
9	5.817063689	0.092506494	0.084077	-0.091127	0.091127	0.008304
10	5.904786627	0.083263587	0.075087	-0.098205	0.098205	0.009644

BIAS = -0.0362
MAD = 0.05517
MSE = 0.00422
BIAS (Vithala) = -0.03597
MAD (Vithala) = 0.055057
MSE (Vithala) = 0.004203

Utilizing the semi-parametric PWP-GT method to obtain $M\hat{T}TF$ in terms of the seed number, parameter θ, and sample size, the results of $M\hat{T}TF$ in the example of $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$ are summarized in Table 2.13. Note that two classes that can be distinguished by the covariate effect divide all samples evenly. There is a calculating error in Vithala table regarding the average $M\hat{T}TF$ in failure numbers 7, 8, and 10 of CLASS=1, which should be corrected as

0.1103333, 0.139, and 0.11, accordingly, based on the three $M\hat{T}TF$ values in Vithala table listed in Table 2.14. After the correction, the duplicative results are very close to the Vithala results, which prove that the duplicative work is reliable.

Table 2.13. Average $M\hat{T}TF$ obtained from the PWP method

60 sample units, 10 failures/unit, 30units/class, 2 classes, $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$					
CLASS = 0					
	$M\hat{T}TF$			Average $M\hat{T}TF$	Vithala results
Failure	539	255	59		
1	5.349641982	5.198709123	5.231432122	5.259927742	5.26
2	1.121911863	1.045134844	1.005004705	1.057350471	1.0573
3	0.357002717	0.499915438	0.327856446	0.394924867	0.395
4	0.257464462	0.297334521	0.342227573	0.299008852	0.29867
5	0.15745239	0.298393934	0.261153878	0.239000067	0.23867
6	0.207883454	0.171514117	0.156506356	0.178634642	0.179
7	0.158173277	0.150163123	0.192077086	0.166804495	0.16667
8	0.11112206	0.126900585	0.128013973	0.122012206	0.122
9	0.111719866	0.104901648	0.11821875	0.111613421	0.11167
10	0.115923406	0.115139303	0.089179423	0.106747377	0.10667
CLASS = 1					
	$M\hat{T}TF$			Average $M\hat{T}TF$	Vithala results
Failure	539	255	59		
1	3.153395177	3.61694346	3.686388826	3.48557582	3.48533
2	1.146859147	0.66078219	0.823795448	0.87714559	0.87733
3	0.412534866	0.516007395	0.433962555	0.45416827	0.45533
4	0.303167926	0.294818079	0.220211824	0.27273261	0.27267
5	0.276920782	0.209555661	0.200803662	0.22909337	0.22933
6	0.204151742	0.165247898	0.187858335	0.18575266	0.18567
7	0.112479919	0.112624617	0.105528102	0.11021088	**0.12841**
8	0.14722054	0.129605016	0.139804935	0.13887683	**0.1139**
9	0.138303359	0.107650349	0.071416031	0.10578991	0.10567
10	0.123336704	0.105547047	0.100557877	0.10981388	**0.0911**

Table 2.14. Average $\hat{MTTF}$ obtained from the PWP method (corrected)

60 sample units, 10 failures/unit, 30units/class, 2 classes,

$\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$

	CLASS = 0			
	$\hat{MTTF}$			
Failure	539	255	59	Average $\hat{MTTF}$
1	5.35	5.199	5.231	5.26
2	1.122	1.045	1.005	1.0573
3	0.357	0.5	0.328	0.395
4	0.257	0.297	0.342	0.29867
5	0.157	0.298	0.261	0.23867
6	0.208	0.172	0.157	0.179
7	0.158	0.15	0.192	0.16667
8	0.111	0.127	0.128	0.122
9	0.112	0.105	0.118	0.11167
10	0.116	0.115	0.089	0.10667
	CLASS = 1			
	$\hat{MTTF}$			
Failure	539	255	59	Average $\hat{MTTF}$
1	3.153	3.617	3.686	3.48533
2	1.147	0.661	0.824	0.87733
3	0.413	0.516	0.437	0.45533
4	0.303	0.295	0.22	0.27267
5	0.277	0.21	0.201	0.22933
6	0.204	0.165	0.188	0.18567
7	0.112	0.113	0.106	$0.12841 \rightarrow 0.110$
8	0.147	0.13	0.14	$0.1139 \rightarrow 0.139$
9	0.138	0.108	0.071	0.10567
10	0.123	0.106	0.101	$0.0911 \rightarrow 0.11$

$\hat{MTTF}$ can be obtained using the semi-parametric PWP method by implementing the Product-Limit method, which integrates the area under its survival function. Experiments on eight combinations of the sample size= 60, 120 and shape parameter=1.2, 2.0 are investigated in the estimating of mean time to n^{th} failure ($\hat{MTTF}(n)$), for $n = 1,2,\ldots,10$. Note that each value of $\hat{MTTF}(n)$ is determined by the average of three replicates in terms of seed numbers 539, 255, and 59. The results of one example $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$ are summarized in Table 2.15.

Table 2.15. The semi-parametric PWP method

60 sample units, 10 failures/unit, 30units/class, 2 classes, $\theta = 1.2, \mu_0 = -6.9, \mu_1 = -4.6$

CLASS = 0

Failure	t_n	IMTBF(t_n)	$M\hat{T}TF$	e_n	$\lvert e_n \rvert$	e_n^2
1	5.902634188	0.83263407	1.06	0.2698861	0.269886	0.072839
2	6.479907133	0.416491777	0.39	-0.051782	0.051782	0.002681
3	6.817678122	0.277700038	0.30	0.0767332	0.076733	0.005888
4	7.057354876	0.208289602	0.24	0.1474412	0.147441	0.021739
5	7.243272849	0.166638677	0.18	0.0719879	0.071988	0.005182
6	7.395184155	0.138869451	0.17	0.2011605	0.20116	0.040466
7	7.523626394	0.119033338	0.12	0.0250255	0.025025	0.000626
8	7.634890059	0.104155733	0.11	0.0716013	0.071601	0.005127
9	7.733032869	0.092583953	0.11	0.1529793	0.152979	0.023403
10	7.820825524	0.083326335				

BIAS = 0.107226

MAD = 0.118733

MSE = 0.022244

BIAS (Vithala) = 0.107011

MAD (Vithala) = 0.118478

MSE (Vithala) = 0.022124

CLASS = 1

Failure	t_n	IMTBF(t_n)	$M\hat{T}TF$	e_n	$\lvert e_n \rvert$	e_n^2
1	3.992219332	0.826410878	0.88	0.0613916	0.061392	0.003769
2	4.566373547	0.414928835	0.45	0.0945691	0.094569	0.008943
3	4.903101776	0.277004333	0.27	-0.015421	0.015421	0.000238
4	5.142256551	0.207897967	0.23	0.101951	0.101951	0.010394
5	5.327861145	0.166387916	0.19	0.1163831	0.116383	0.013545
6	5.479563451	0.138695258	0.11	-0.205374	0.205374	0.042178
7	5.607856365	0.118905331	0.14	0.1679613	0.167961	0.028211
8	5.719008014	0.104057711	0.11	0.0166466	0.016647	0.000277
9	5.817063689	0.092506494	0.11	0.1870937	0.187094	0.035004
10	5.904786627	0.083263587				

BIAS = 0.05836

MAD = 0.10742

MSE = 0.01782

BIAS (Vithala) = 0.059042

MAD (Vithala) = 0.107965

MSE (Vithala) = 0.017995

The robustness test is aimed to evaluate the performance of two methods in estimating the instantaneous mean time between failures from an NHPP with underlying log-linear intensity function. Three performance measurements are utilized for the comparison of the Lawless and PWP-GT methods in estimating the mean time to the n^{th} failure. Three performance measurements are BIAS, MAD (mean absolute deviation), and MSE (mean squared error).

All combinations of the sample size =60, 120 and shape parameter=1.2 and 2.0 are investigated in this duplicating process of Vithala results to estimate mean time to n^{th} failure, $n = 1,2,...,10$ implemented by the Lawless and PWP-GT methods. Each value of either $IM\hat{T}BF(n)$ or $M\hat{T}TF(n)$ is determined by the average of three replicates of different seed numbers in order to decrease the bias effect. The robustness tests of the estimating methods are summarized in Table 2.16 and Table 2.18. To compare with Vithala results, Table 2.17 and Table 2.19 are taken from Vithala (1994) and listed here.

Table 2.16. The summary of the robustness test using the Lawless method

10 failures/unit, 10units/class, 2 classes, $\mu_0 = -6.9, \mu_1 = -4.6$							
		CLASS = 0			CLASS = 1		
U[a]	θ	BIAS	MAD	MSE	BIAS	MAD	MSE
60	1.2	0.020106	0.041592	0.00355	-0.0362	0.05517	0.00422
60	2.0	0.019785	0.04098	0.003444	-0.0363	0.05475	0.00415
120	1.2	-0.00546	0.056711	0.005217	-0.0004	0.05536	0.00521
120	2.0	-0.00556	0.055985	0.005076	-0.0005	0.05467	0.00507

[a] U represents the sample size in terms of the number of sample units.

Table 2.17. Vithala results

10 failures/unit, 10units/class, 2 classes, $\mu_0 = -6.9, \mu_1 = -4.6$							
		CLASS = 0			CLASS = 1		
U[a]	θ	BIAS	MAD	MSE	BIAS	MAD	MSE
60	1.2	0.020697	0.041681	0.003581	-0.03597	0.058057	0.005483
60	2.0	0.013391	0.040388	0.003168	-0.03618	0.054701	0.004147
120	1.2	-0.00546	0.056711	0.005217	-0.00139	0.055424	0.005197
120	2.0	-0.00556	0.055985	0.005076	-0.00055	0.054672	0.005074

[a] U represents the sample size in terms of the number of sample units.

Table 2.18.The summary of the robustness test using the PWP method

10 failures/unit, 10units/class, 2 classes, $\mu_0 = -6.9, \mu_1 = -4.6$							
		CLASS = 0			CLASS = 1		
U[a]	θ	BIAS	MAD	MSE	BIAS	MAD	MSE
60	1.2	0.107226	0.118733	0.022244	0.05836	0.10742	0.01782
60	2.0	0.106745	0.120436	0.022704	0.06133	0.11036	0.0184
120	1.2	0.061103	0.061406	0.007691	0.01809	0.04043	0.0025
120	2.0	0.061068	0.061251	0.007658	0.02001	0.03923	0.00249

[a] U represents the sample size associated with sample units.

Table 2.19. Vithala results

		10 failures/unit, 10units/class, 2 classes, $\mu_0 = -6.9, \mu_1 = -4.6$					
		CLASS = 0			CLASS = 1		
U[a]	θ	BIAS	MAD	MSE	BIAS	MAD	MSE
60	1.2	0.107011	0.118478	0.022124	0.059042	0.107879	0.017886
60	2.0	0.104911	0.122469	0.021196	0.060492	0.108437	0.016935
120	1.2	0.061839	0.062102	0.007844	0.017854	0.040589	0.002506
120	2.0	0.062393	0.062633	0.007916	0.019233	0.03865	0.002428

[a] U represents the sample size associated with sample units.

Another important contribution from Vithala (1994) comes from the statistical analysis to confirm the model adequacy. The case of $\theta = 0$ in a log-linear intensity function was assumed and a regression analysis ANOVA was performed by the GLM procedure in the SAS software. The intension of the study is to find the slope and intercept estimates $(\hat{\delta}, \hat{\upsilon}_0, \hat{\upsilon}_1)$ in the case of $\theta = 0$, $\upsilon_0 = \ln(0.001) = 6.9078$, $\upsilon_1 = \ln(0.01) = 4.6051$ as theoretical values. Theoretically, for an HPP ($\theta = 0$), the slope is equal to zero and the failure intensity equation is simplified to

$\lambda(t) = \exp^{\mu + \theta \times t} = \exp^{\mu}$.

The results taken from Vithala (1994) are listed in Table 2.20.

Table 2.20. GLM summary (Vithala (1994))

Units	$\upsilon_0 = \lambda$ $\upsilon_1 = \lambda$	Slope estimate	Intercept estimate	$\hat{\lambda}$	t statistics[a]
	0.001	0.000001	6.849663	0.001052	49.97
20	0.010	0.000281	4.518974	0.010902	39.97
	0.001	0.000003	6.888710	0.001021	109.31
60	0.010	-0.000005	4.623224	0.009821	73.36
	0.001	0.000004	6.878877	0.001029	109.31
120	0.010	-0.000005	4.623615	0.009817	73.64

[a] : t statistics for H_0 : intercept = $\ln(1/\lambda)$

2.2.4 Censoring of recurrent events

A common phenomenon in data collection is the existence of censoring data. Engelhardt et al. (1993) reviewed and explained the fundamentals of a censoring experiment, where the failure data in a repairable system are often modeled as a counting process. There are two types of right-censoring mechanisms depending on what criteria by which the data collection is terminated: fixed time length or fixed failure number. According to Engelhardt et al., a process is said to be failure truncated if it is observed until a fixed number of failures have occurred, and it is said to be time truncated if it is observed for a fixed length of time.

Qureshi (1991) and Vithala (1994) simulated the recurring failure data by the manner of failure censoring, since ten failure events were generated from an NHPP-power-law and NHPP-log-linear processes, respectively, for each sample unit. In the circumstances when historical data are not available, data may be considered left-truncated. Moreover, if the number of failures is known even though the failure times are not recorded, the data are termed as left-censored. Engelhardt et al. (1993) derived maximum likelihood estimation formulas for left-truncated data, in which the likelihood function $L(\upsilon,\delta)$ was defined as follows (υ and δ denoted the scale parameter and the shape parameter, accordingly).

If a power-law process has been truncated from the left at time τ_1 and time truncated from the right at time τ_2, with $R = r$ observed failure times, $t_1 < \ldots < t_r$, in the interval $[\tau_1, \tau_2]$, the likelihood function $L(\lambda,\delta)$ is given by Engelhardt et al. (1993) as

$$L(\upsilon,\delta) = \begin{cases} (\upsilon\delta)^r \left[\prod_{i=1}^{r} t_i\right]^{\delta-1} \exp\left[-\upsilon\left(\tau_2^{\delta} - \tau_1^{\delta}\right)\right] & r \geq 1 \\ \exp\left[-\upsilon\left(\tau_2^{\delta} - \tau_1^{\delta}\right)\right] & r = 0. \end{cases}$$

To involve covariate effects by dividing sample units in a population, Hu and Lawless (1996) have developed estimation procedures regarding a censoring experiment on automobile failure data.

Cai and Prentice (1995) created a way to simulate censorship with probability (P_i,P_j) for two streams from a bivariate distribution, where (P_i,P_j) was a given set of fixed probabilities. For instance, if (P_i,P_j) = (0.5,0.9), the censoring procedures are executed as follows. First, since $P_2 = 0.9$, then 10% of the uncensored failure data T_{k2} is reserved theoretically. Second, since $P_1 = 0.5$, then $10\% \times 50\% = 5\%$ of the paired failure times (T_{k1},T_{k2}) is remained in the second step. $P_i \in Uniform(0,1)$ is defined as the censoring probability. Bowman (1996) generated recurring failure times (X_1,X_2) from a bivariate exponential distribution and independent censoring times (C_1,C_2) from an exponential distribution.

The probability of an observation being a censored time is $P[X > C] = \frac{\lambda_c}{\lambda_c + \lambda_x}$, where λ_x, λ_c are parameters of the failure and censoring distributions.

The failure event time X from the exponential distribution and the corresponding censoring time can be expressed as

$$X = -\frac{1}{\lambda_x}\ln(U_x)\exp(-\beta \times z)$$

$$C = -\frac{1}{\lambda_c}\ln(U_c)$$

$$\lambda_c = \frac{\lambda_x e^{\beta \times z} P[X > C]}{1 - P[X > C]}.$$

Thus, the observed event times (T_1,T_2) are the minimal of (X_1,X_2) and (C_1,C_2). In other words,

$$T_1 = \min\{X_1, C_1\}$$
$$T_2 = \min\{X_2, C_2\}.$$

The Indicator function (I_1,I_2) can be defined as $I = 1$ when $T = X$; otherwise $I = 0$.

Various sample sizes (number of recurrences) among sample units leads to a censoring experiment. In the setting of a right-censoring, the higher numbered failures are removed from the experiment. To select the censored units, a random number ranging between (0,1) is generated. If the random number is less than censoring probability C_p, the unit is treated as a censored unit; otherwise, it is an uncensored unit. The censoring time for those censored units is assumed as the last failure time in Qureshi (1991) and Vithala (1994). However, the censoring time is determined by the follow-up time in most medical studies, such as WLW (1989).

The maximum partial likelihood function determines the estimate of regression coefficient $\hat{\beta}$ for the Cox-based methods. The concern is on how the censoring (unequal sample size of failure times) affects the partial likelihood function, and how different it is compared with equal sample size. PWP (1981) addressed a partial likelihood function (with no censorship or equal sample size in failure counts) based on the PWP-TT method written as follows.

$$L(\beta)=\prod_{s\geq 1}\prod_{i=1}^{d_s}\left[\frac{\exp\{z_{si}(t_{si})\beta_s\}}{\sum_{l\in R(t_{si},s)}\exp\{z_l(t_{si})\beta_s\}}\right].$$

In the case of censorship allowed, WLW (1989) developed a partial likelihood function with respect to failure event-specific stratum s based on the WLW method.

$$L_s(\beta)=\prod_{i=1}^{d_i}\left[\frac{\exp\{z_{si}(t_{si})\beta_s\}}{\sum_{l\in R_s(t_{si})}\exp\{z_{sl}(t_{si})\beta_s\}}\right]^{\Delta_{si}},\ t_{si}=\min\{\tilde{t}_{si},c_{si}\},$$

where

s: event-specific stratum s,
i: i^{th} subject ,
R: i^{th} subject s^{th} stratum in the risk set at t_{si},
Δ_{si}: censoring indicator,
where $\Delta_{si}=\begin{cases}1 & t_{si}=\tilde{t}_{si}\\ 0 & t_{si}=c_{si}\end{cases}.$

To illustrate the partial likelihood function employed in the PWP method, five subjects are assumed with equal sample size d_i associated with failure counts. In other words, $d_1=d_2=d_3=d_4=d_5=4$. The procedures of how the partial likelihood is formed can be illustrated below.

1. Assume the sorted failure times in the first stratum ($s=1$) from the five subjects are $t_{s4}<t_{s5}<t_{s1}<t_{s2}<t_{s3}, s=1$.

2. $L_{s=1}(\beta)=\prod_{i=1}^{5}\left[\frac{\exp\{z_{si}(t_{si})\beta_s\}}{\sum_{l\in R(t_{si},s)}\exp\{z_l(t_{si})\beta_s\}}\right]$

$$= \frac{\exp\{z_{s4}(t_{s4})\beta_s\}}{\exp\{z_{s4}(t_{s4})\beta_s\}+\exp\{z_{s5}(t_{s4})\beta_s\}+\exp\{z_{s1}(t_{s4})\beta_s\}+\exp\{z_{s2}(t_{s4})\beta_s\}+\exp\{z_{s3}(t_{s4})\beta_s\}} \times$$

$$\frac{\exp\{z_{s5}(t_{s5})\beta_s\}}{\exp\{z_{s5}(t_{s5})\beta_s\}+\exp\{z_{s1}(t_{s5})\beta_s\}+\exp\{z_{s2}(t_{s5})\beta_s\}+\exp\{z_{s3}(t_{s5})\beta_s\}} \times$$

$$\frac{\exp\{z_{s1}(t_{s1})\beta_s\}}{\exp\{z_{s1}(t_{s1})\beta_s\}+\exp\{z_{s2}(t_{s1})\beta_s\}+\exp\{z_{s3}(t_{s1})\beta_s\}} \times$$

$$\frac{\exp\{z_{s2}(t_{s2})\beta_s\}}{\exp\{z_{s2}(t_{s2})\beta_s\}+\exp\{z_{s1}(t_{s2})\beta_s\}} \times \frac{\exp\{z_{s1}(t_{s1})\beta_s\}}{\exp\{z_{s1}(t_{s1})\beta_s\}}.$$

Likewise, in the other strata $s = 2,3,4$, $L_{s=2}(\beta), L_{s=3}(\beta), L_{s=4}(\beta)$ can be produced in the same manner depending on the ordered failure times t_{si}. In the PWP method, to measure the general covariate effect β,

$$L(\beta) = \prod_{s=1}^{4} L_s(\beta) = L_{s=1}(\beta) \times L_{s=2}(\beta) \times L_{s=3}(\beta) \times L_{s=4}(\beta).$$

3. The maximum likelihood equation, $\frac{\partial \log L(\beta)}{\partial \beta} = 0$ provides the m.l.e. $\hat{\beta}$.

For the censoring data with different sample size in failure counts, the WLW method has provided a way to produce the partial likelihood function. Assume the data are taken from five subjects, which contain unequal sample size of failures, (i.e., $d_1 = 4, d_2 = 1, d_3 = 2, d_4 = 0, d_5 = 4$, in a left-censoring experiment). Assume also that the sorted failure times in the first stratum ($s = 1$) from the five subjects are $t_{s4} < t_{s5} < t_{s1} < t_{s2} < t_{s3}$, $s = 1$, where the order is not necessarily the same as in complete data depending on censoring times.

$L_{s=1}(\beta) =$

$$\left[\frac{\exp\{z_{s4}(t_{s4})\beta_s\}}{\exp\{z_{s4}(t_{s4})\beta_s\}+\exp\{z_{s5}(t_{s4})\beta_s\}+\exp\{z_{s1}(t_{s4})\beta_s\}+\exp\{z_{s3}(t_{s4})\beta_s\}+\exp\{z_{s2}(t_{s4})\beta_s\}}\right]^0\times$$
$$\left[\frac{\exp\{z_{s5}(t_{s5}(\beta_s\}}{\exp\{z_{s5}(t_{s5})\beta_s\}+\exp\{z_{s1}(t_{s5})\beta_s\}+\exp\{z_{s3}(t_{s5})\beta_s\}+\exp\{z_{s2}(t_{s5})\beta_s\}}\right]^1\times$$
$$\left[\frac{\exp\{z_{s1}(t_{s1})\beta_s\}}{\exp\{z_{s1}(t_{s1})\beta_s\}+\exp\{z_{s3}(t_{s1})\beta_s\}+\exp\{z_{s2}(t_{s1})\beta_s\}}\right]^1\times$$
$$\left[\frac{\exp\{z_{s3}(t_{s3})\beta_s\}}{\exp\{z_{s3}(t_{s3})\beta_s\}+\exp\{z_{s2}(t_{s3})\beta_s\}}\right]^0\times\left[\frac{\exp\{z_{s2}(t_{s2})\beta_s\}}{\exp\{z_{s2}(t_{s2})\beta_s\}}\right]^0.$$

As the formula reveals, a censored unit does not have an impact upon the likelihood function through Δ_{si}=0. However, the censored unit does contribute through the risk set $R(t_{si})$ at time t_{si}. Thus, as far as the censoring time or failure time is involved in the risk set, it will contribute to the likelihood function.

As for the PWP method, the censoring time is not allowed in developing the likelihood function. Thus, the censoring times in implementing the PWP method are excluded from the dataset. Note that the censoring failure on the border will be kept in the dataset due to the conditionality approach. The partial likelihood function of the PWP method, where the data present a censoring pattern as below, has been proposed by WLW (1989).

$$L_s(\beta)=\prod_{i=1}^{d_i}\left[\frac{\exp\{z_{si}(t_{si})\beta_s\}}{\sum_{l\in R_s(t_{si})}\exp\{z_{sl}(t_{si})\beta_s\}}\right]^{\Delta_{si}},t_{si}=\min\{\tilde{t}_{si},c_{si}\},$$

where

s: event-specific stratum s,
i: i^{th} subject,
R: i^{th} subject s^{th} stratum in the risk set at t_{si},
Δ_{si}: censoring indicator,

where $\Delta_{si} = \begin{cases} 1 & t_{si} = \widetilde{t}_{si} \\ 0 & t_{si} = c_{si} \end{cases}$.

The only adjustment to the PWP partial likelihood function is the removal of the data not in the risk set $R(t_{si})$ at time t_{si}. Thus, the dataset needs to be modified in the manner that the likelihood function is not underestimated, since there will be fewer components in the denominator after adjustment. That is, the data that are not in the risk set should be removed. The determinant of data removal is when current and previous data are both censoring times.

The Andersen and Gill (1982), or AG method, employs the marginal martingale theory to manage the partial likelihood function in order to represent the general covariate effect by $\hat{\beta}$. The counting process $N^{(n)}$ has intensity process $\lambda^{(n)}$. The local martingales on the time interval $t \in [0,1]$ are defined as (Andersen and Gill (1982)):

$$M_{si}(t) = N_{si}(t) - \int_0^t \lambda_{si}(x)dx.$$

The local square integrable martingales (Andersen and Gill (1982))

$$\left\langle M_i^{(n)}, M_i^{(n)} \right\rangle(t) = \int_0^t \lambda_i^{(n)}(u)du$$

$$\left\langle M_i^{(n)}, M_j^{(n)} \right\rangle(t) = 0 \qquad i \neq j.$$

The partial likelihood function in using the AG method with censorship can be expressed as below proposed by Cai and Prentice (1995).

$$L(\beta) = \prod_{i=1}^{d_1} L_{s=1}(\beta) \times \prod_{i=1}^{d_2} L_{s=2}(\beta) \times \ldots \prod_{i=1}^{d_s} L_{s=s}(\beta) = \prod_{s\geq 1} \prod_{i=1}^{d_s} \left[\frac{\exp\{z_{si}(t_{si})\beta_s\}}{\sum_{l \in R_s(t_{si})} Y_{sl}(t_{si}) \exp\{z_{sl}(t_{si})\beta_s\}} \right]^{\Delta_{si}},$$

where $Y_{sl}(t)$ is an at-risk indicator for subjects.

The solution $\hat{\beta}$ to the partial likelihood score equation $\frac{\partial \log L(\beta)}{\partial \beta} = 0$ is expressed as (Cai and Prentice (1995)).

$$\sum_{s\geq 1}\sum_{i=1}^{n}\int_{0}^{\infty} z_{si}(u)U_{si}(u)du = 0,$$

where $U_{si}(t) = \hat{M}_{si}(t) = N_{si}(t) - \int_{0}^{t} Y_{si}(h)\exp\{z_{si}^{T}(h)\beta\}\hat{\Lambda}_{0i}(dh)$,

$U_{si}(t)$ is the estimated marginal martingale corresponding to T_{si},

$$\hat{\Lambda}_{0i}(t) = \int_{0}^{t}\left[\sum_{l=1}^{s} Y_{li}(s)\exp\{z_{li}^{T}(h)\beta\}\right]^{-1}\sum_{l=1}^{s} N_{li}(dh).$$

Kelly and Lim (2000) have categorized each possibility of the Cox-based PI models based on the risk set and the risk interval, and used a hypothetical example in Figure 2.8 to illustrate each partial likelihood function. There are three subjects A, B, and C under observation, and each box represents a failure event and each dot represents a censoring event.

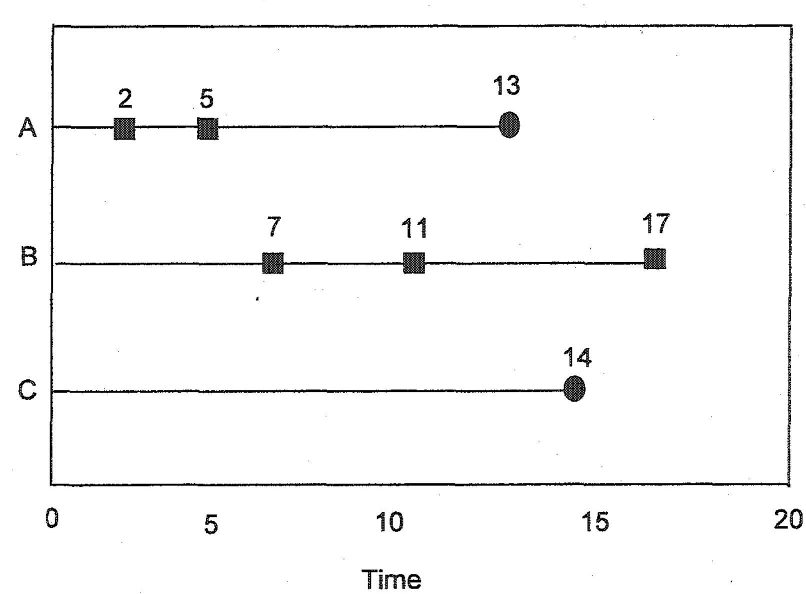


Figure 2.8. A hypothetical example from Kelly and Lim (2000)

Based on each partial likelihood function in each category, the PWP-GT, PWP-TT, AG, and WLW are selected and presented according to the hypothetical example. Kelly and Lim (2000) demonstrated the partial likelihood functions based on two primary categories: (1) Common baseline intensity model (2) Event-specific baseline intensity model.

(1) Common baseline intensity model

$$L(\beta)=\prod_{i=1}^{n}\prod_{k=1}^{K}\left(\frac{e^{z(X_{ik})\beta}}{\sum_{j=1}^{n}\sum_{l=1}^{K}Y_{jl}(X_{ik})e^{z(X_{ik})\beta}}\right)^{\Delta_{ik}}.$$

(2) Event-specific baseline intensity model

$$L(\beta)=\prod_{i=1}^{n}\prod_{k=1}^{K}\left(\frac{e^{z(X_{ik})\beta}}{\sum_{j=1}^{n}Y_{jk}(X_{ik})e^{z(X_{ik})\beta}}\right)^{\Delta_{ik}}.$$

The AG model belongs to the former category, while the PWP-GT, PWP-TT, and WLW models belong to the latter category.

According to the Chebychev inequality below, the magnitude of the variance of $\hat{\beta}$ due to censoring plans or sample units can be explained in the law of large numbers (Guttman et al. (1982)) for a distribution in a general form. If x is a random variable having finite mean μ and variance σ^2, then

$$P(|x-\mu|>\lambda\sigma)\leq\frac{1}{\lambda^2}.$$

Guttman et al. (1982) have derived the distribution of sample mean $\bar{x}$ with its variance σ^2 / n, where the probability that $\bar{x}$ falls outside the interval $[\mu - \varepsilon, \mu + \varepsilon]$ follows the chebychev inequality.

Let $\lambda = \dfrac{\varepsilon\sqrt{n}}{\sigma}$,

The Chebychev inequality becomes

$$P\left(|\bar{x} - \mu| > \frac{\varepsilon\sqrt{n}}{\sigma}\frac{\sigma}{\sqrt{n}}\right) \le \frac{\sigma^2}{\varepsilon^2 n}.$$

Equivalently,

$$P(|\bar{x} - \mu| > \varepsilon) \le \frac{\sigma^2}{\varepsilon^2 n},$$

where ε is an arbitrarily small positive number.

Comparing the $Var_A(\hat{\beta})$ utilizing the complete data (equal sample size (failure events)) with $Var_B(\hat{\beta})$ utilizing the censored data, $Var_B(\hat{\beta}) > Var_A(\hat{\beta})$ due to the concern of the sample size, according to the Chebychev inequality. However, censored data can still provide sufficient information if the sample size is chosen wisely. The experimental design (Sections 3.1.1-3.1.2) will explain how to decide the appropriate sample size.

2.2.5 Multiple event types

Lin (1994) has dealt with two failure types of recurring events using the Cox-based regression methods to analyze the effectiveness of the treatments. The paired failure times were generated from a bivariate exponential distribution (Gumbel (1960)) with correlation coefficient $\theta = 0.25$. Case (1): Lin reported a study of multiple failure types in colon cancer, where cancer recurrences and death were two types of failures. Obviously, the recurring events are all prior to death if a patient dies. As a result, two baseline intensity functions $\lambda_{01}(t), \lambda_{02}(t)$ were employed in this colon cancer study. Lin concluded that the common treatment effect was not asymptotically equivalent since the correlation between two failure types was strong. Moreover, high correlation of two failure types of events also contributed to different results on naïve and robust variances. Case (2): In another study of reducing infection rate by taking gamma interferon, the first three infections were selected to do the analysis while the three infections were treated as three different failure types in order to capture the dependence of infections. A time-varying covariate was utilized, where the covariate was equal to one when the patient had an infection within the past 60 days; otherwise, the covariate equaled zero. Lin observed that the WLW method is always valid, whereas the PWP and AG methods are valid when the dependence structure is correctly specified. Case (3): In a study of the diabetic retinopathy causing the occurrence of blindness, two failure types were defined as the blindness occurrences on left eye and the blindness occurrences on right eyes. Using the WLW method was valid, although the correlation between two eyes had been anticipated. The results indicated that the robust standard error estimates were smaller than the naïve estimates in this study. Case (4): in a genetic epidemiologic study of schizophrenia, the number of the relatives ranging from 1 to 12 was selected to represent the multiple failure types. Age at onset of the illness and gender (two covariates) were suspected to affect the age at diagnosis of schizophrenia for a relative (failure time).

Lin (1993) developed MULCOX2 software in implementing Cox-based methods, such as the AG, PWP-total time, PWP-gap time, and WLW models. This code is capable of analyzing multiple failure types of recurring events. MULCOX2 required two types of data entries: control parameters (problem titles, file names, dimension number, and other relevant information) and data files (associated with identification numbers, failure times, status, covariates, and other indicator variables). Lin utilized the data from a schizophrenia study (twelve failure types: the number of relatives in a family ranging from one to twelve) and a chronic granulomatous disease, a CGD study (three failure types).

In the schizophrenia study (Lin (1993)), 487 first-degree relatives of 93 female schizophrenic patients were enrolled. The covariate vector is $\mathbf{Z}_{ik}$ = $(Z_{i1}, Z_{i2}, ..., Z_{ik})$, where k represents the failure type and i denotes the subject. In this case, twelve failure types were involved representing the number of relatives ranging between (1,12). Each record in the data set has the form of $(start, end, status, Z_{i1}, Z_{i2}, ..., Z_{i12})$, where $(start, end)$ denotes the failure interval, and status equals to one (1) when the subject is under observation, otherwise zero (0). Furthermore, two types of covariates $p =$ age and gender, are considered in the covariate vector. Thus, covariate vector $\mathbf{Z}_{pik} = (Z_{1ik}, Z_{2ik})$ follows this form as a record in the dataset.

The CGD study has 128 patients involved with using gamma interferon treatment to reduce the granulomatous disease. To estimate the treatment effect for three failure types individually, three covariates $Z_{i1} = (R_i, 0, 0)$, $Z_{i2} = (0, R_i, 0)$, and $Z_{i3} = (0, 0, R_i)$ representing three failure types are employed in the analysis, where i indicates the subject ($R_i = 1$, for gamma interferon, $R_i = 0$, for placebo). Since three failure types are employed in the analysis, three regression coefficients are yielded in representing the treatment effect for each failure type. To estimate a general treatment effect for three failure types, $Z_{ik} = R_i$ is employed as a single covariate in the estimating process. Two AG methods, semi-Markov and Markov processes, are implemented for comparison purposes. Likewise, two PWP methods, total time and gap time, are introduced in the analysis. The marginal approach gives a larger estimate of the common treatment parameter along with a larger standard error estimate (Lin (1994)). In general, the results from different methods all conclude that the treatment effect (by taking gamma interferon) reduces the infection sufficiently. One interesting fact in the analysis is that the treatment effect analyzed by the PWP method is not significant in the second and third infections.

In medical studies, the issue of treatment effects is the main concern, while engineering reliability has additionally emphasized the effects of multiple failure types, such as the major and minor failure types in machinery. Covariate modeling is approached to examine the covariate effects for major and for minor events. Let the treatment factor be defined as $R_i = 0$, for class=0; $R_i = 1$, for class=1, and covariate vector $\mathbf{Z}_{ik} = (Z_{i1}, Z_{i2})$ represents major and minor failure types, where i represents the sample unit and k represents the failure type. That is, class0 and class1 for major events can be expressed in the forms of

$(Z_1, Z_2) = (0,0)$ and $(1,0)$. Likewise, class0 and class1 for minor events are in the forms of $(Z_1, Z_2) = (0,0)$ and $(0,1)$. Two regression coefficients represent the covariate effects corresponding to major and minor events. To estimate the general covariate effect for both major and minor event types altogether, a single covariate $Z_{ik} = R_i$, where $R_i = 0$, for class = 0 and $R_i = 1$, for class = 1 is introduced in the regression model (i represents the sample unit and k represents the failure type). In this case, only one regression coefficient will be obtained, which represents the covariate effect based on the major or minor event type.

Hansen and Ascher (2002) examined an automobile for intermittent failures, which often lead to a series of unsuccessful repair attempts, and reported that repair times for intermittent failures cannot be assumed negligible and the model must be designed to account for machine downtimes. Kobbacy and Jeon (2002) considered both failure times and machine downtimes in the PI model for preventive maintenance (PM) in a deteriorating repairable system. Therneau and Hamilton (1997) proposed an alternative method of handling two types of recurrent events, and introduced the concept of discontinuous risk-free-intervals that may be applied in reliability engineering as the duration of performing major overhauls. A study of rhDNase in patients with cystic fibrosis has involved discontinuous risk-free-intervals due to the intervals of receiving an intravenous (IV) antibiotics and a seven-day risk-free period following the IV antibiotics. The two event types are (1) at risk of infection with stochastic interval of recurrence and (2) risk-free-intervals with deterministic interval.

2.3 References

[1] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 1982;10(4):1100-20.

[2] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Assessing the maintenance in a process using a semi-parametric approach. Quality and Reliability Engineering International, 2001;17:163-7.

[3] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Analysing maintenance data to gain insight into systems performance. Journal of Operational Research Society, 2003;54:343-9.

[4] Ansell JI, Phillips MJ. Practical problems in the statistical analysis of reliability data (with Discussion). Applied Statistics, 1989;38(2):205-47.

[5] Ansell JI, Phillips MJ. Practical reliability data analysis. Reliability Engineering and System Safety, 1990;28(3):337-56.

[6] Ansell JI, Phillips MJ. Practical aspects of modeling of repairable systems data using proportional hazards models. Reliability Engineering and System Safety, 1997;58(2):165-71.

[7] Asher H. Regression analysis of repairable systems reliability. In: Skwirzynski JK, editor. Electronic systems effectiveness and life cycle costing. Berlin: Springer, 1983. p.119-133.

[8] Ascher H, Feingold H. Repairable systems reliability--modeling inference, misconceptions and their causes, Lecture notes in statistics, New York: Marcel Dekker Inc 1984;7:p.22-4.

[9] Ascher H, Feingold H. Repairable systems reliability--modeling inference, misconceptions and their causes, Lecture notes in statistics, New York: Marcel Dekker Inc 1984;7.

[10] Bendell A, Wightman DW, Walker EV. Applying proportional hazards modeling in reliability. Reliability Engineering and System Safety, 1991;34:35-53.

[11] Blanks HS, Tordon MJ. Laplace and Mann aging trend test effectiveness. Proceedings of the Annual Reliability and Maintainability Symposium, 1987:165-70.

[12] Bowman ME. An evaluation of statistical models for the analysis of recurrent events data: with application to needlestick injuries among a cohort of female veterinarians. Doctoral Dissertation, Ohio State University, 1996.

[13] Cai J, Prentice RL. Estimating equations for hazard ratio parameters based on correlated failure time data. Biometrika, 1995;82(1):151-64.

[14] Cox DR. Regression models and life tables. Journal of the Royal Statistical Society, Series B, 1972;34:187-202.

[15] Cox DR, Lewis PAW. The statistical analysis of series of events, London: Methuen, 1966.

[16] Dale C. Application of the proportional hazards model in the reliability

field. Reliability Engineering, 1985;10:1-14.
[17] Engelhardt M, Williams DH, Bain LJ. Statistical analysis of a power-law process with left-truncated data. Edited by Basu, Asit P. Advances in Reliability Elsevier Science Publishers B. V., Amsterdam 1993;105-21.
[18] Gumbel EJ. Bivariate exponential distributions. Journal of the American Statistical Association, 1960;55: 698-707.
[19] Guttman I, Wilks SS, Hunter, JS. Introductory Engineering Statistics, Third Edition: Wiley Series in Probability and Mathematical Statistics-Applied, John Wiley & Sons, Inc 1982;p.105.
[20] Hansen CK, Ascher HE. Modeling repair events under intermittent failures and failures subject to unsuccessful repair. Quality and Reliability Engineering International, 2002;18:453-65.
[21] Hu XJ, Lawless JF. Estimation from truncated lifetime data with supplementary information on covariates and censoring times. Biometrika, 1996;83(4):747-61.
[22] Jiang W, Turnbull BW, Clark LC. Semiparametric regression models for repeated events with random effects and measurement error. Journal of the American Statistical Association: Theory and Methods, Mar. 1999;94(445): 111-24.
[23] Kelly PJ, Lim LL-Y. Survival analysis for recurrent data: An application to childhood infectious diseases. Statistics in Medicine, 2000;19:13-33.
[24] Kijima M. Some results for repairable systems with general repair. Journal of Applied Probability 1989; 26: 89-102.
[25] Kijima M. Morimura, H., and Suzuki, Y. Periodical replacement problem without assuming minimal repair. European Journal of Operation Research, 1988;37:194-203.
[26] Kobbacy KAH, Jeon J. Generalized non-stationary preventive maintenance model for deteriorating repairable systems. Quality and Reliability Engineering International, 2002;18:363-72.
[27] Landers TL, Jiang ST, Peek JR. Semi-parametric PWP model robustness for log-linear increasing rates of occurrence of failures. Reliability Engineering and System Safety, 2001;73:145-53.
[28] Landers TL, Soroudi HE. Robustness of a Semi-parametric Proportional Intensity Model. IEEE transactions on Reliability, Jun.1991;40(2):161-4.
[29] Law AM, Kelton MD. Simulation Modeling and Analysis, New York: McGraw-Hill Inc. 1991.
[30] Lawless JF. Statistical models and methods for lifetime data. New York: John Wiley & Sons, Inc 1982.
[31] Lawless JF. Regression methods for Poisson process data. Journal of the American Statistical Association, 1987; 82(399):808-815.
[32] Leemis LM. Reliability: Probabilistic Models and Statistical Methods. Prentice Hall 1995.
[33] Lin DY. MULCOX2: A general computer program for the Cox

regression analysis of multivariate failure time data. Computer Methods and Programs in Biomedicine 1993; 40:279-93.
[34] Lin DY. Cox's regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine, 1994; 13: 2233-47.
[35] Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. Journal of the American Statistical Association: Theory and Methods, Dec. 1989;84(408):1074-8.
[36] Lipschutz KH, Snapinn SM. Discussion of paper by Wei and Glidden. Statistics in Medicine, 1997;16:841-51.
[37] Patrick DT. O'Connor, Practical Reliability Engineering, Third Edition, New York: John Wiley & Sons, Inc. 1991;p. 293.
[38] Pham H, Wang H. Imperfect maintenance. European Journal of Operation Research, 1996; 94:425-38.
[39] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika, 1981; 68(2):373-9.
[40] Proschan F. Theoretical explanation of observed decreasing failure rate. Technometrics, Aug. 1963;5(3):375-83.
[41] Qureshi WM. Assessment of a proportional intensity model for repairable-system reliability. Doctoral Dissertation, The University of Arkansas Dec.1991.
[42] Qureshi WM, Landers TL, Gbur EE. Robustness evaluation of a semi-parametric proportional intensity model. Reliability Engineering and System Safety, 1994;44:103-9.
[43] Ross SM. Introduction to Probability Models, Fifth Edition, Academic Press, Inc. 1993.
[44] Soroudi HE. Reliability estimation for repairable systems with covariates. MS thesis, The University of Arkansas May 1990.
[45] Therneau TM, Hamilton SA. rhDNase as an example of recurrent event analysis. Statistics in Medicine, 1997;16:2029-47.
[46] Usher JS, Kamal AH, Syed WH. Cost optimal preventive maintenance and replacement scheduling. IIE Transactions, 1998;30:1121-8.
[47] Vithala S. Robustness of Semi-parametric proportional Intensity Model for the Case of a Log-linear Nonhomogeneous Poisson Process. MS thesis, The University of Arkansas Dec.1994.
[48] Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Statistics in Medicine, 1997; 16:833-9.
[49] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, Dec.1989;84(408):1065-73.
[50] Wightman DW, Bendell A. The practical application of proportional hazards modeling. Reliability Engineering, 1986;15:29-53.

3 Research methodology

As an extension of the robustness study of the PWP-gap time (PWP-GT) model in Qureshi (1991) and Vithala (1994), other models are introduced to handle recurrent event reliability problems, namely the PWP-total time (PWP-TT), Andersen-Gill (AG), and Wei-Lin-Weissfeld (WLW) methods. In addition, more questions commonly encountered in the industry will be raised and investigated, such as incidence of right-censoring and multiple event types. Sections 3.1 and 3.2 are designated to investigate the two research questions: (1) How do the PWP-GT, PWP-TT, AG, and WLW methods compare in performance under right-censoring? (2) How do the multi-dimensional covariate modeling and discontinuous risk-free-intervals methods perform in estimating the regression coefficients for two failure types (major and minor)?

The first research objective examines the PWP-GT model robustness as a function of right-censoring severity measured by BIAS, MAD (mean absolute deviation), and MSE (mean squared error). The special case of common baseline intensity function (WLW and PWP-TT models) is investigated to compare with the AG model. The second research objective examines the robustness of the four reliability estimates (PWP-GT, PWP-TT, AG, and WLW) as a function of right-censoring severity for the special case of a stationary counting process. BIAS, MAD, and MSE are employed to measure the robustness of the three event-specific models (PWP-GT, PWP-TT, and WLW), while the common baseline model (AG) estimates the general covariate effect. The third research objective examines multi-dimensional covariate modeling as an alternate method to deal with two types of complete (uncensored) recurrent events. In this study, two types of recurrent events are generated from the same data stream with common shape parameter forming proportional intensities. The fourth research objective examines risk-free-intervals within an NHPP process where there are two event types (major and minor) and the time interval following a major failure is substantial. The robustness study of the four methods (PWP-GT, PWP-TT, AG, and WLW) is conducted in terms of sample size, power-law shape parameter, censoring probability, and gap time ratio (discontinuous risk-free-intervals).

3.1 Robustness of semi-parametric methods under right censoring

Depending on the selection of the baseline intensity functions (common and event-specific) on PWP-TT and WLW models, there are two studies (NHPP (Section 3.1.1) and HPP (Sections 3.1.2)) conducted in each specified model (common or event-specific). Essentially, the PWP-TT and WLW models are designated as event-specific models. However, due to the model restriction, in an NHPP case and 10 failure events (for each sample unit), a common baseline model is required in order to have a robust model performance. In the case of an HPP and 4 failure events (for each sample unit), an event-specific model can perform properly without the assumption of a common baseline.

3.1.1 NHPP

Using the Cox-based regression methods (PWP-GT, PWP-TT, AG, and WLW methods), model recurring failure events (with right-censorship), which follow an NHPP with power-law intensity function, and examine the robustness of the four methods. Selecting appropriate baseline hazards and risk interval is the key to an adequate model. Regarding the baseline hazard, there are two options to choose from: common baseline hazard function and event-specific hazard function. For risk interval, there are three options: total time model, gap time model, and counting process.

Censored data is generally present in field data. The left-censored case arises when the historic event times are not available but the number of missing events is known. The right-censored case arises when the subject or sample is withdrawn from observation (such as machines retired from service). Censoring from the right is chosen for examination in this research. The censoring probability controls the number of censored sample units in the experiment, and the number of censored events in this study is designed to follow a random pattern. The four Cox-based regression methods are compared based on the theoretical values of regression coefficients, which measure the covariate effects.

The experiment is conducted based on the following settings. Sample units (U) are evenly divided into two groups defined by a single covariate named CLASS. Each sample unit produces 10 failure times ($N = 10$) generated from an NHPP with a power-law form, by the Blanks & Tordon (1987) simulation algorithm as follows:

$$t_n = \left[t_{n-1}^{\delta} - \frac{\ln(X_i)}{\upsilon} \right]^{1/\delta},$$

where

X_i: a random variate generated from a $(0,1)$ uniform distribution ,
δ, υ: the shape parameter and the scale parameter of the power-law form,
n : failure count,

t: recurring failure times.

Two covariate levels (CLASS=0 and CLASS=1) are defined by setting the scale parameter to $\upsilon_0 = 0.001$ or $\upsilon_1 = 0.01$.

Simulation data generated from the Blanks & Tordon algorithm provide complete data, where each sample unit contains an equal number of failures (N). In order to have various numbers of failure counts (i.e., right-censored recurrent data), two groups of sample units were classified, in which one group contains the sample units with complete data and the other group contains the sample units with right-censored data. The ratio (probability) of the sample units that have censored times to total sample units is defined as censored probability (P_c). In the group of censored units, the right-censored pattern is set to be random. A random probability (P_1) is generated to compare with P_c. The sample unit is specified as a censored unit if $P_1 < P_c$; otherwise, it is not a censored unit. To form recurring data (failure times in a sequence) with right-censoring in a random pattern, another random probability (P_2) is generated. In the censored group, it is a censored time if both of the logic rules are met: 1) the sample unit is a censored unit and, 2) the failure count is greater than F, where F = floor ($N \times$(ranuni (seed))) +1, and the floor (argument) function is to return the largest integer that is less than or equal to the argument. The data in the non-censored group are all complete data.

To implement the four Cox-Based regression methods (PWP-GT, PWP-TT, AG, and WLW) requires formulation of three types of datasets (i.e., three formats for the same set of failure events, according to the theory underlying each methodology). First, for the AG method, the data set is.formed from the time interval (T_1, T_2) with respect to the following counting process formulation:

$$\lim_{h \to 0} \frac{1}{h} P[N(t+h) - N(t) = 1 \mid T > t] = \lambda(t).$$

Thus, the logic rule to form the dataset is: $T_2 > T_1$. As a result, all the censored failure times are removed from the dataset since $T_2 = T_1$ when it is a censored event as stipulated for the AG method.

The concept of forming the dataset for the PWP method originates from the probability theory of conditionality. The later failure times after the n^{th} failure count cannot be included into the dataset when the intensity function at the n^{th} failure count is estimated. That is, for each censored unit, the censored times are removed from the dataset except for the first censored event count. The logic in generating the dataset for the PWP method is to remove the record if both of the following conditions hold: (1) the current record is marked censored and (2) the previous record is marked censored.

Due to the marginal probability theory of the WLW method, the dataset contains full records including all censored events, such that censored units remain in the risk set. The Lee-Wei-Amato (LWA) model (1994) is a special case of the WLW with common baseline intensity function, and is used in the case of a total time model. Likewise, the PWP-TT model in this study is a special case, where the baseline intensity function is set to the common baseline intensity function leading to a regression coefficient to explain general covariate effects.

Three factors are chosen in the experimental design: number of the sample units (U), shape parameter (δ), and censoring probability (P_c). I_0 and I_1 represent the number of units in each class. Table 3.1 contains the experimental design for the right-censoring experiments. The selection of the U, δ, and P_c levels has taken the following considerations: (1) The parameter settings in the previous relevant works (Proschan (1963), Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001)) (2) Severe right-censorship may cause the small sample size (U=20) to have insufficient data. The selection of P_c levels takes into account the light, moderate, and heavy censoring. The selection of U and δ levels is taken from the parameter settings in the previous research works, and it has also considered the small, median, and large sample sizes for U.

Table 3.1. Three-factor experimental design: (U,δ,P_c)

N=10 failure events/unit, υ_0=0.001, υ_1=0.01														
U	δ	P_c	I_0	I_1	U	δ	P_c	I_0	I_1	U	δ	P_c	I_0	I_1
60	0.5	0.4	30	30	120	0.5	0.4	60	60	180	0.5	0.4	90	90
60	0.5	0.6	30	30	120	0.5	0.6	60	60	180	0.5	0.6	90	90
60	0.5	0.8	30	30	120	0.5	0.8	60	60	180	0.5	0.8	90	90
60	0.5	1.0	30	30	120	0.5	1.0	60	60	180	0.5	1.0	90	90
60	0.8	0.4	30	30	120	0.8	0.4	60	60	180	0.8	0.4	90	90
60	0.8	0.6	30	30	120	0.8	0.6	60	60	180	0.8	0.6	90	90
60	0.8	0.8	30	30	120	0.8	0.8	60	60	180	0.8	0.8	90	90
60	0.8	1.0	30	30	120	0.8	1.0	60	60	180	0.8	1.0	90	90
60	1.0	0.4	30	30	120	1.0	0.4	60	60	180	1.0	0.4	90	90
60	1.0	0.6	30	30	120	1.0	0.6	60	60	180	1.0	0.6	90	90
60	1.0	0.8	30	30	120	1.0	0.8	60	60	180	1.0	0.8	90	90
60	1.0	1.0	30	30	120	1.0	1.0	60	60	180	1.0	1.0	90	90
60	1.2	0.4	30	30	120	1.2	0.4	60	60	180	1.2	0.4	90	90
60	1.2	0.6	30	30	120	1.2	0.6	60	60	180	1.2	0.6	90	90
60	1.2	0.8	30	30	120	1.2	0.8	60	60	180	1.2	0.8	90	90
60	1.2	1.0	30	30	120	1.2	1.0	60	60	180	1.2	1.0	90	90
60	1.5	0.4	30	30	120	1.5	0.4	60	60	180	1.5	0.4	90	90
60	1.5	0.6	30	30	120	1.5	0.6	60	60	180	1.5	0.6	90	90
60	1.5	0.8	30	30	120	1.5	0.8	60	60	180	1.5	0.8	90	90
60	1.5	1.0	30	30	120	1.5	1.0	60	60	180	1.5	1.0	90	90
60	1.8	0.4	30	30	120	1.8	0.4	60	60	180	1.8	0.4	90	90
60	1.8	0.6	30	30	120	1.8	0.6	60	60	180	1.8	0.6	90	90
60	1.8	0.8	30	30	120	1.8	0.8	60	60	180	1.8	0.8	90	90
60	1.8	1.0	30	30	120	1.8	1.0	60	60	180	1.8	1.0	90	90
60	2.0	0.4	30	30	120	2.0	0.4	60	60	180	2.0	0.4	90	90
60	2.0	0.6	30	30	120	2.0	0.6	60	60	180	2.0	0.6	90	90
60	2.0	0.8	30	30	120	2.0	0.8	60	60	180	2.0	0.8	90	90
60	2.0	1.0	30	30	120	2.0	1.0	60	60	180	2.0	1.0	90	90

3.1.2 HPP

To relax the common baseline function applied on PWP-TT and WLW utilized in Section 3.1.1 (NHPP), four failure events are generated from an HPP with a right-censoring pattern and thus the event-specific baseline PWP-TT and WLW models can be employed. Unlike the gap time scale (PWP-GT), the total time scale (PWP-TT and WLW) has a misspecification problem. The gap time scale has been considered a better model to capture the dependence structure existing among failure times than has the total time scale. Thus, in any rate of occurrence of failures, utilizing the gap time scale can capture the trend and give a sound estimate of covariate effects, while the total time scale appears to overestimate covariate effects as the event count progresses. Besides, the total time scale is invariant to the shape parameter (δ), because δ does not influence the likelihood function in the total-time model. The counting process (AG) adopts the total time scale, and thus becomes an estimator invariant to shape parameter.

Simulation data with right-censored patterns (the underlying distribution follows power-law NHPP intensity function) were generated by a modified Blanks & Tordon (1987) simulation algorithm. Since stationary data are specified, $\delta = 1$ is set to convert a power-law NHPP into an HPP. There are two experimental factors (Table 3.2): experimental units (U) and censoring probability (P_c). The levels for each factor are selected as follows: (1) U = 60, 120, and 180 and (2) P_c= 0, 0.4, 0.8, and 1.0. Note that $P_c = 0$ represents complete data, which provides the comparison with censored data. The selection of the U and P_c levels has taken the following considerations: (1) the parameter settings from the previous relevant works (Proschan (1963), Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001)) (2) Severe right-censorship may cause the small sample size (U =20) to have insufficient data to perform the model analysis. The selection of P_c levels takes into account the light, moderate, and heavy censoring. Likewise, the selection of sample units (U) levels considers the small, median, and large sample sizes.

Table 3.2. Two-factor experimental design: (U, P_c)

N=4 failure events/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$			
Number of units U	Censoring probability P_c	Units per class (I)	
		I_0	I_1
60	0.0	30	30
60	0.4	30	30
60	0.8	30	30
60	1.0	30	30
120	0.0	60	60
120	0.4	60	60
120	0.8	60	60
120	1.0	60	60
180	0.0	90	90
180	0.4	90	90
180	0.8	90	90
180	1.0	90	90

3.2 Modeling of multiple failure types in recurrent events

Multiple failure types are often observed in reliability failure data. The scope of this research is to investigate two situations involving recurrent failure processes composed of two failure types (major and minor). A multi-dimensional covariate may be used to model multiple failure types having common shape parameter forming proportional intensities. A major overhaul period may be defined as a risk-free interval to perform the maintenance/repair. In the aircraft industry, for example, a major overhaul of a substantial time interval is performed when a major failure or a fixed interval inspection is scheduled, whichever occurs first.

3.2.1 Multi-dimensional covariate modeling

Lin (1993, 1994) studied chronic granulomatous disease and employed a multiple dimensional covariate method to handle the recurrent data with multiple failure types. Lin considered three types of failure outcomes by defining three covariates with three dimensions. For the special case of two failure types, let two covariates Z_1, Z_2 represent the major and minor failure types in two dimensions. That is, the major failure type is coded as $Z_1 = [R_{1i}, 0]$, while the minor failure type is coded as $Z_2 = [0, R_{2i}]$, where R_{1i} =1, if class=1; R_{1i} =0, if class=0; R_{2i} =1, if class=1; R_{2i} =0, if class=0. The corresponding regression coefficient estimates are interpreted as the covariate effect applied to the major failure and minor failure types.

In industry, minor failure rate is typically higher than major failure rate. Most researchers have formulated this problem as univariate. The Lin method of multi-dimensional covariates permits consolidation of major and minor failures in a single, stratified model so long as the proportional intensity rule holds. The simulation method of Blanks & Tordon (1987) is modified to generate an NHPP with two failure types, where the underlying distribution follows a power-law intensity function. Most of the parameters remain unchanged except that the sample unit size has been increased due to the insufficient sample size of major events. In the process of generating the simulation data, there is a major failure event out of these ten recurring events (i.e., nine minor events) for each sample unit. To perform the event-specific intensity estimation, the dataset may not have any major event for a certain event count if the sample size is not large enough.

This research suggests the minimal sample size of 120 for this study. The parameter setting is illustrated as follows: $U = 120$, $F = 10$, $\mu_0 = 0.001$, $\mu_1 = 0.01$. The fixed time-invariant covariate vector $Z_i, i = 1,2$ is defined as follows:

Major event, Class=0: $Z_1 = (0,0)$
Major event, Class=1: $Z_1 = (1,0)$
Minor event, Class=0: $Z_2 = (0,0)$
Minor event, Class=1: $Z_2 = (0,1)$.

Ten failure events are generated for each sample unit. To determine the time of major failure event in the counting process, a uniformly distributed variate $(U(0,1))$ is introduced to decide the event number (F) for occurrence of the major failure. As a consequence, the F^{th} event time to have a major failure is generated as:

$F = FLOOR(10 \times RANUNI(SEED)) + 1$.

The remaining nine events are minor failure events. In this way, a counting process contains major and minor failure events, where the one major failure is inserted randomly among the $N-1$ minor failures. The event number for the major failure is randomly selected depending on the F value. Large enough sample size is generated in order to obtain sufficient data for each failure count in a PWP-GT model. Two factors are chosen in the experimental design: number of the sample units (U) and shape parameter (δ). Table 3.3 provides the experimental design for the covariate modeling. The selection of the δ level has taken the parameter settings from the previous relevant works (Proschan (1963), Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001)).

Table 3.3. Two-factor experimental design: (U,δ)

N=10 failure events/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$			
Number of units U	Shape parameter δ	Units per class (I)	
		I_0	I_1
120	0.5	60	60
120	0.8	60	60
120	1.0	60	60
120	1.2	60	60
120	1.5	60	60
120	1.8	60	60
120	2.0	60	60
180	0.5	90	90
180	0.8	90	90
180	1.0	90	90
180	1.2	90	90
180	1.5	90	90
180	1.8	90	90
180	2.0	90	90
240	0.5	120	120
240	0.8	120	120
240	1.0	120	120
240	1.2	120	120
240	1.5	120	120
240	1.8	120	120
240	2.0	120	120

3.2.2 Discontinuous risk-free-intervals modeling

The second method to approach multiple failure types is applying the concept of discontinuous risk-free-intervals, proposed by Therneau and Hamilton (1997). A study of rhDNase in patients with cystic fibrosis involved a seven-day discontinuous risk-free-interval, initiated by intravenous (IV) administration of antibiotics (Therneau and Hamilton (1997)). For instance, suppose three failures have taken place at days 25, 60, and 90, where two days of performing a major overhaul are required after the second failure. The data records, expressed as $(n,t_1,t_2,status)$ for the three failure times in the PWP-GT model, can be written as $(1,0,25,1)$, $(2,25,60,1)$, and $(3,62,90,1)$, where $(n,t_1,t_2,status)$ denotes (failure count, start time, stop time, (0,1) indicator variable for censor (0) event or failure (1) event). The value $t_2 = 90$ of the third failure with a major overhaul records global time to failure with the third failure coinciding with a risk-free-interval. However, the consideration of major overhaul of duration D requires a change from interval (t_1, t_2) to interval $(t_1 + D, t_2)$. In the aircraft industry, the duration D could be as long as one year after flying for 3000 hours for a major overhaul or as short as a few hours for a minor repair. A robustness study examines how the magnitude of D affects the PWP-GT model, as measured by the regression estimates ($\hat{\beta}_i$).

In this study, simulated recurring data are generated from a modified Blanks & Tordon algorithm (1987). To determine the time to perform major overhauls in the counting process, a uniformly distributed $U(0,1)$ random variate is introduced to select the event number F, where the major overhaul is performed. The major overhaul is arranged after the F^{th} event, and we assume that a period D is required to perform a major overhaul. As a consequence, the next event time, which belongs to the $(F+1)^{st}$ event, occurs depending on the F^{th} event time plus the major overhaul duration. For instance, if ten failure events (failure count F_c=1 to 10) are generated for each sample unit in the database, then the time point to perform a major overhaul occurs at the F^{th} event time. The SAS statement syntax to derive F is

$F = FLOOR(10 \times RANUNI(SEED)) + 1$.

As a result, three FOR loops in the SAS program are created as follows: (1) FOR F_c=1 to F, (2) FOR F_c=$F+1$, and (3) FOR F_c=$F+2$ to N, where N

denotes the last event number. To assure the adequacy of event numbers, two extreme values of F equal to 1 and N are examined. First, when $F = 1$, three loops are derived as (1) FOR F_c=1, (2) FOR F_c=2, and (3) FOR F_c= 3 to N. Second, when $F = N$, three loops become (1) FOR F_c=1 to N, (2) FOR F_c=N+1, and (3) FOR F_c= N+2 to N. The logic in the third loop (i.e. FOR F_c= N+2 to N) requires the following statement:

IF $F_c > N$ THEN DELETE.

The duration to perform a major overhaul is inserted into the interval (t_F, t_{F+1}), which makes the interval of risk become $(t_F + D, t'_{F+1})$, where new event time t'_{F+1} is determined by $t_F + D$ in the Blanks & Tordon formula. As a consequence, the gap time and (t_F, t_{F+1}) have been altered compared to the recurrent data without the interruption of a major overhaul interval. However, the discontinuous risk-free-intervals concept in Therneau and Hamilton (1997) is different in terms of (t_F, t_{F+1}), while the gap time remains unchanged. "For instance, in a study of patients with hip fracture, a subject who fractured at day 100, followed by a 15 day hospital stay and then 300 more days of uneventful follow-up would be represented as two at-risk intervals: $(0,100]$ and $(115,415]$" (Therneau and Hamilton (1997)). The gap times of 100 days and 300 days remain the same, while the risk interval has been shifted forward from $\{(0,100], (100,400]\}$ to $\{(0,100], (115,415]\}$.

The magnitude of D is determined based on the previous gap time Y_{F-1}, where F is a random variate indicating the F^{th} event is a major failure event;

otherwise, a minor failure event. In other words, the relationship between D and Y_{F-1} is:

$$D = R \times Y_{F-1} \longrightarrow R = D / Y_{F-1},$$

where

R is the gap time ratio that controls the magnitude of D,
Y_{F-1} represents gap time associated with a minor event prior to a major event,
F is the event number that represents the major failure.

The concept of utilizing the gap time ratio R in determining the major overhaul duration strengthens the model, since there are three types of power-law intensity functions (increasing rate of occurrence of failures (IROCOF), constant ROCOF, and DROCOF). The recurrent failure interval can vary from one time unit to a large value depending on the shape parameter.

The parameter settings are as follows when a discontinuous risk interval model is associated with the repair time: (1) scale parameters in CLASS0 and CLASS1 are set to 0.001 and 0.01; (2) number of failures N = 10; (3) F^{th} event represents a major failure, followed by a major overhaul; and (4) seed numbers for three replicates are 539, 255, and 59. The magnitude of D is examined as the primary factor that affects the performance of the semi-parametric PI models. In the experimental design for the discontinuous risk interval model, there are three experimental factors: (1) Number of the experimental units (U), (2) Shape parameter (δ), and (3) Gap time ratio (R) that controls the major overhaul duration (D). I_0 and I_1 represent the number of units in each class.

The simulation data in the form of discontinuous risk intervals is illustrated by the following numerical example: Let $N = 10$ failure events/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$, $(U, \delta, R) = (120, 1.5, 0.50)$, and seed = 539. Two units are chosen for demonstration purposes and each unit has 10 failure times in the form of risk interval $=(T_1, T)$. Also let $Y = T - T_1$. Note that F represents the major failure event. Thus, the major overhaul takes place immediately after the F^{th} event. As is shown in Table 3.4, the first item has generated F=7 resulting in a discontinuous risk interval starting from the end of 7^{th} failure, at time 211.71966. Since the major overhaul duration $D = R \times Y = 0.5 \times 25.60324$ =12.80162, the risk-free-interval ends after 211.71966+D=224.52128, which is the beginning of the risk interval for the 8^{th} failure. Likewise, for the second unit, F=4. T_1 for the 5^{th} failure is changed to $227.10957 + 0.5 \times 81.95362 = 268.08637$.

Three factors are chosen in the experimental design: number of the sample units (U), shape parameter (δ), and gap time ratio (R). Table 3.5 contains the experimental design for discontinuous risk interval experiments. Small sample size $U = 20$ is introduced to reflect the poor performance of the PWP-GT model as the gap time ratio increases. The selection of the U, δ, and R levels has taken the following considerations: (1) the parameter settings from the previous relevant works (Proschan (1963), Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001)) (2) Gap time ratio reflects the repair/overhaul duration that starts from an immediate repair (zero time) to a five times of the previous interarrival failure time (Y).

Table 3.4. Simulation dataset example

n	Item	T_1	T	Y	F	n	Item	T_1	T	Y	F
1	1	0.00000	25.70391	25.70391	7	1	2	0.00000	7.53385	7.53385	4
2	1	25.70391	87.10256	61.39865	7	2	2	7.53385	135.05711	127.52327	4
3	1	87.10256	101.09154	13.98898	7	3	2	135.05711	145.15595	10.09884	4
4	1	101.09154	119.87100	18.77946	7	4	2	145.15595	**227.10957**	**81.95362**	4
5	1	119.87100	165.66447	45.79347	7	5	2	**268.08637**	270.69678	43.58721	4
6	1	165.66447	186.11641	20.45194	7	6	2	270.69678	276.21527	5.51849	4
7	1	186.11641	**211.71966**	**25.60324**	7	7	2	276.21527	305.22196	29.00669	4
8	1	**224.52128**	232.26621	20.54656	7	8	2	305.22196	392.22638	87.00443	4
9	1	232.26621	237.85589	5.58967	7	9	2	392.22638	393.61682	1.39044	4
10	1	237.85589	258.58578	20.72989	7	10	2	393.61682	417.92776	24.31095	4

Table 3.5. Three-factor experimental design: (U,δ,R) .

N=10 failure events/unit, υ_0=0.001,υ_1=0.01														
U	δ	R	I_0	I_1	U	δ	R	I_0	I_1	U	δ	R	I_0	I_1
20	0.5	0.001	10	10	60	0.5	0.001	30	30	120	0.5	0.001	60	60
20	0.5	0.1	10	10	60	0.5	0.1	30	30	120	0.5	0.1	60	60
20	0.5	0.3	10	10	60	0.5	0.3	30	30	120	0.5	0.3	60	60
20	0.5	0.5	10	10	60	0.5	0.5	30	30	120	0.5	0.5	60	60
20	0.5	3.0	10	10	60	0.5	3.0	30	30	120	0.5	3.0	60	60
20	0.5	5.0	10	10	60	0.5	5.0	30	30	120	0.5	5.0	60	60
20	0.8	0.001	10	10	60	0.8	0.001	30	30	120	0.8	0.001	60	60
20	0.8	0.1	10	10	60	0.8	0.1	30	30	120	0.8	0.1	60	60
20	0.8	0.3	10	10	60	0.8	0.3	30	30	120	0.8	0.3	60	60
20	0.8	0.5	10	10	60	0.8	0.5	30	30	120	0.8	0.5	60	60
20	0.8	3.0	10	10	60	0.8	3.0	30	30	120	0.8	3.0	60	60
20	0.8	5.0	10	10	60	0.8	5.0	30	30	120	0.8	5.0	60	60
20	1.0	0.001	10	10	60	1.0	0.001	30	30	120	1.0	0.001	60	60
20	1.0	0.1	10	10	60	1.0	0.1	30	30	120	1.0	0.1	60	60
20	1.0	0.3	10	10	60	1.0	0.3	30	30	120	1.0	0.3	60	60
20	1.0	0.5	10	10	60	1.0	0.5	30	30	120	1.0	0.5	60	60
20	1.0	3.0	10	10	60	1.0	3.0	30	30	120	1.0	3.0	60	60
20	1.0	5.0	10	10	60	1.0	5.0	30	30	120	1.0	5.0	60	60
20	1.2	0.001	10	10	60	1.2	0.001	30	30	120	1.2	0.001	60	60
20	1.2	0.1	10	10	60	1.2	0.1	30	30	120	1.2	0.1	60	60
20	1.2	0.3	10	10	60	1.2	0.3	30	30	120	1.2	0.3	60	60
20	1.2	0.5	10	10	60	1.2	0.5	30	30	120	1.2	0.5	60	60
20	1.2	3.0	10	10	60	1.2	3.0	30	30	120	1.2	3.0	60	60
20	1.2	5.0	10	10	60	1.2	5.0	30	30	120	1.2	5.0	60	60
20	1.5	0.001	10	10	60	1.5	0.001	30	30	120	1.5	0.001	60	60
20	1.5	0.1	10	10	60	1.5	0.1	30	30	120	1.5	0.1	60	60
20	1.5	0.3	10	10	60	1.5	0.3	30	30	120	1.5	0.3	60	60
20	1.5	0.5	10	10	60	1.5	0.5	30	30	120	1.5	0.5	60	60
20	1.5	3.0	10	10	60	1.5	3.0	30	30	120	1.5	3.0	60	60
20	1.5	5.0	10	10	60	1.5	5.0	30	30	120	1.5	5.0	60	60
20	1.8	0.001	10	10	60	1.8	0.001	30	30	120	1.8	0.001	60	60
20	1.8	0.1	10	10	60	1.8	0.1	30	30	120	1.8	0.1	60	60
20	1.8	0.3	10	10	60	1.8	0.3	30	30	120	1.8	0.3	60	60
20	1.8	0.5	10	10	60	1.8	0.5	30	30	120	1.8	0.5	60	60
20	1.8	3.0	10	10	60	1.8	3.0	30	30	120	1.8	3.0	60	60
20	1.8	5.0	10	10	60	1.8	5.0	30	30	120	1.8	5.0	60	60
20	2.0	0.001	10	10	60	2.0	0.001	30	30	120	2.0	0.001	60	60
20	2.0	0.1	10	10	60	2.0	0.1	30	30	120	2.0	0.1	60	60
20	2.0	0.3	10	10	60	2.0	0.3	30	30	120	2.0	0.3	60	60
20	2.0	0.5	10	10	60	2.0	0.5	30	30	120	2.0	0.5	60	60
20	2.0	3.0	10	10	60	2.0	3.0	30	30	120	2.0	3.0	60	60
20	2.0	5.0	10	10	60	2.0	5.0	30	30	120	2.0	5.0	60	60

4. Semi-parametric proportional intensity models robustness for right-censored recurrent failure data

4.0 Abstract

This paper reports the robustness of the four proportional intensity (PI) models: PWP-gap time (PWP-GT), PWP-total time (PWP-TT), Andersen-Gill (AG), and Wei-Lin-Weissfeld (WLW), for right-censored recurrent failure event data that follow a Non-homogeneous Poisson Process (NHPP). The results are beneficial to practitioners in anticipating the more favorable applications domains and selecting appropriate PI models in applying to right-censored recurrent failure data. This experimental design has incorporated three levels of censorship severity (light, moderate, and severe) to evaluate these four proposed PI models. The PWP-GT and AG prove to be models of choice, evaluated in terms of the bias, mean absolute deviation, and mean squared error of covariate regression coefficients over ranges of sample sizes, shape parameters, and censoring severity encountered in engineering applications. The more favorable engineering applications ranges are recommended. At the smaller sample size (U=60), the PWP-GT proves to perform well for moderate right-censoring ($0.0 \le P_c \le 0.8$) and moderately decreasing, constant, and moderately increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.8 \le \delta \le 1.8$). For the large sample size (U=180), the PWP-GT performs well for severe right-censoring ($0.0 \le P_c \le 1.0$) and moderately decreasing, constant, and moderately increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.8 \le \delta \le 2.0$). The AG model proves to outperform the

PWP-TT and WLW for stationary process (HPP) across a wide range of right-censorship ($0.0 \le P_c \le 1.0$) and for sample sizes of 60 (30 per class) or more.

Keywords: repairable systems reliability, right-censoring, recurrent events, proportional intensity models

Nomenclature

Acronyms

AG	Andersen and Gill model
C.I.	Confidence interval
DROCOF	Decreasing rate of occurrence of failures
HPP	Homogeneous Poisson Process
IROCOF	Increasing rate of occurrence of failures
i.i.d	Independent and identically distributed
LWA	Lee, Wei, and Amato model
MTTF	Mean time to failure
MAD	Mean absolute deviation
MSE	Mean squared error
NHPP	Non-homogeneous Poisson Process
PH	Proportional hazards
PI	Proportional intensity
PWP	Prentice, Williams, and Peterson model
PWP-GT	Prentice, Williams, and Peterson-gap time model
PWP-TT	Prentice, Williams, and Peterson-total time model
WLW	Wei, Lin, and Weissfeld model

Notation

C_{ki}	Censoring time for the i^{th} subject of the k^{th} type of failures
$h(t;z)$	Proportional hazard function
$h_0(t)$	Baseline hazard function
I_0	Number of sample units in class ϕ

I_1	Number of sample units in class 1
i.i.d.	Independent and identically distributed
N	Successive failure count
$N(t)$	Random variable for the number of failures in $(0,t]$; a counting process
n	An integer counting successive failure times; a stratification indicator subscript
P_c	Censoring probability
s.d.	Standard deviation
T_1, T_2	The beginning and end of an event; bivariate exponential variables
T_n	Random variable for cumulative time of occurrence of the n^{th} failure
t_n	Cumulative time of occurrence of the n^{th} failure; a realization of T_n
U	Sample size (number of units)
$\widetilde{X}$	Observation time
$Y_i^{(n)}$	An at-risk indicator in the AG model
$\mathbf{Z}(t)$	Covariate process up to time t
$\mathbf{z}$	$(k \times 1)$ vector of covariates, $\mathbf{z} = (z_1, z_2, \ldots z_k)'$
$\boldsymbol{\beta}_n$	$(k \times 1)$ vector of stratum-specific regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$
δ	Shape parameter of a power-law NHPP
Δ	Indicator of a failure or censored time; limit to time zero
λ_0	Baseline value of λ for power-law NHPP
$\lambda_0(t)$	Baseline intensity function
$\lambda_{0n}(t)$	Stratum-specific baseline intensity function
$\lambda(t;\mathbf{z})$	Proportional intensity function
υ	Scale parameter of a power-law NHPP
υ_0	Baseline value of υ, the scale parameter of a power-law NHPP
υ_1	Alternate value of υ, the scale parameter of a power-law NHPP

σ Standard deviation

^ Denotes an estimator

' Denotes the transpose of a vector

4.1 Introduction

Failure time data on a repairable system are realizations of a stochastic point process, in which the instantaneous rate of occurrence of failures (ROCOF) is $\lambda(t)$. Prentice, Williams, and Peterson (PWP) [1] proposed a semi-parametric approach to model recurrent failure event data from a repairable system using two methods: PWP-GT (gap time) and PWP-TT (total time). Several researchers have subsequently proposed alternate modeling methods by modifying the risk set (common or event-specific baseline intensity function) and the risk interval (gap time, total time, or counting process). These include the AG (Andersen-Gill) [2] and WLW (Wei-Lin-Weissfeld) [3] models.

Cox proposed the distribution-free (semi-parametric) proportional hazards (PH) model in 1972 [4]. The Cox-based regression models (PWP-GT, PWP-TT, AG, and WLW) have been applied to recurring events in medical studies (biostatistics field), such as recurrent infections of a patient. For engineering applications, Landers and Soroudi [5], Qureshi et al. [6], Vithala [7], and Landers et al. [8] have investigated robustness of the PWP-GT model, where the underlying recurrent failure time data are from a Non-homogeneous Poisson Process (NHPP) with a power-law or a log-linear intensity function. These references also report the engineering applications of the PWP-GT model cited in the literature. Qureshi et al. [6] found that the PWP-GT model performs best for constant and moderately increasing rate of occurrence of failures (IROCOF) and decreasing rate of occurrence of failures (DROCOF) and for larger sample sizes from power-law NHPPs. Vithala [7] considered the case of log-linear increasing ROCOF, and concluded the PWP-GT model performs best for moderately increasing ROCOF and for larger sample sizes. Both Qureshi et al. [6] and Vithala [7] have examined robustness of the PWP-GT model for the complete (uncensored) data. However, the phenomenon of censoring data is generally present in field data. This research has extended their work to the important case of right-censorship and has examined other semi-parametric PI models (PWP-TT, AG, and WLW).

Compared to the extensive literature on applications of the Cox-based regression models in the biostatistics field, there have been few reported engineering applications. Abundant federal funding received in biostatistics / medical research has advanced the PI models to become well developed and widely referenced. PI models for medical applications could also apply to recurring failure/repair data in engineering problems. The AG, PWP-GT, PWP-TT, and WLW models are potentially powerful analytical tools for engineering practitioners as they become better recognized and understood. This paper reports the robustness of the PWP-GT, PWP-TT, AG, and WLW models for right-censored recurrent failure event data. The results are beneficial to practitioners in anticipating the more favorable applications domains and selecting appropriate PI models.

4.2 Semi-parametric Proportional Intensity models

Cox [4] proposed a PH formulation to include explanatory variables (covariates) in survival models. PWP proposed an extension of the Cox model to stochastic processes and applied the approach to model recurring infections in aplastic anemia and leukemia patients having received bone-marrow transplants. This application involved several subjects and a small number of events (up to five) for each subject. The paper by PWP did not address the baseline intensity function but rather reported the relative risks for the test and control groups. In reliability and maintainability engineering applications, a number of authors have applied the semi-parametric PI (PH) model, for example, Ansell and Phillips [9], Ansell and Phillips [10], Landers and Soroudi [5], Qureshi et al. [6], Ansell and Phillips [11], Landers et al. [8], Ansell et al. [12], and Ansell et al. [13]. A collection of the PI model applied to different industries includes: marine gas turbine engines (Asher [14]), semiconductor, electrical, and pipeline industries (Ansell and Phillips [11]), U.S. Army main battle tank (Landers et al. [8]), water supply industry (Ansell et al. [12], [13]), etc. Ascher [14] illustrated the use of the PWP model for analysis of reliability for marine gas turbine engines. Ascher and Feingold [15] suggested application of the PWP model in the field of reliability engineering. Dale [16] applied the PWP approach to simulated data for a reliability growth program with design improvements implemented after each of the five stages, resulting in a DROCOF. Wightman and Bendell [17] and Bendell et al. [18] cited the PWP model and advised caution in application for engineering studies.

Qureshi et al. [6] performed a robustness study to determine how well the PWP-GT method performed when applied to data from a failure process that was actually parametric (specifically the NHPP with power-law intensity function). The 2 σ bounds of the PWP-GT estimates can cover the true values with few exceptions. The PWP-GT method performed well, except at small values of shape parameter ($\delta < 0.6$). The PWP-GT method was best for larger sample size and for moderately decreasing, constant, and moderately increasing ROCOF. The validation process for the case of an HPP in Section 2.2.3 (also refer to Table 2.10) indicated that the estimated $MTTF$ (mean time to failure) differences between the PWP-GT model and theoretical values were not statistically significant. As for the PWP-GT estimates of the covariate regression coefficient, the true value of coefficient β lies within the 2 σ confidence bounds on the estimate $\hat{\beta}$ for $1.0 \le \delta \le 1.4$. The PWP-GT method tends to underestimate β for a DROCOF (e.g., BIAS= -26% at $\delta = 0.5$) and overestimate β for an IROCOF (e.g., BIAS= 19% at $\delta = 3.0$).

The AG model (Andersen and Gill [2]) and the WLW model (Wei et al. [3]) are widely cited in the literature. Bowman [19] and Lin [20] surveyed and evaluated the PWP-GT, PWP-TT, AG, and WLW methods. Bowman identified the PWP-GT model as superior and then used it to analyze needle-stick injury data. Wei and Glidden [21] have reviewed the Cox-based methods designed to model recurrent data, and summarized the strengths and weaknesses for each method. In a commentary on the Wei and Glidden paper, Lipschutz and Snapinn [22] stressed the two concepts of “event times” and “risk sets” as crucial to choosing the appropriate model. Event elapsed times are related to the total time, gap time, and counting process. The PWP-TT and WLW are modeled by total time, while only PWP-GT is modeled by gap time. The risk interval of the AG model belongs to the counting process class. Intuitively, total (global) times within a subject are highly correlated. The total time model may indicate large treatment effect throughout the entire study, even though the gap time model has indicated little treatment effect beyond a certain recurrence. The counting process concept of the AG method implies each recurrence is not affected by previous events, and does not contribute to future events.

The risk set consists of the subjects at risk for a specified event (e.g., failure). There are three types of risk sets: conditional (e.g., PWP), counting process (e.g., AG), or marginal (e.g., WLW). As a marginal method, the WLW method assumes a subject is at risk regardless of event count until the observation for the subject terminates by censoring. The AG method also provides an index of a general covariate effect, which is expressed by the common baseline hazard (unrestricted risk set). However, a subject in the PWP method has event-specific baseline hazards (restricted risk set), in that the proportional intensity of event k only considers the subjects that have experienced $(k-1)$ events. Lipschutz and Snapinn [22] suggested guidelines as follows in choosing the appropriate models:

- Use total time, common baseline hazard (unrestricted risk set) when the general effect is of interest.

- Use gap time, event-specific baseline hazards (restricted risk set) when the primary concern is how the treatment will affect the recurring events beyond the first occurrence.

Kelly and Lim [23] noted that risk interval can be defined by three formulations {(1) gap time, (2) total time, and (3) counting process} demonstrated in Fig. 4.1(a)-(c). Risk interval determines whether a model is marginal in the total time or conditional in the gap time. The risk interval of any event in total time is not influenced by any previous events, but measures time from entry into the experiment (beginning of observation). However, the risk interval of the gap time begins from the end of last event (Kelly and Lim [23]). Counting processes use the total time scale and share the same elapsed time as does the gap time model. However, the risk interval starts from the previous event instead of the entry time. Based on the common or event-specific baseline intensities, the risk set is labeled as either unrestricted or restricted. Kelly and Lim [23] defined three possible risk sets {(1) unrestricted, (2) restricted, and (3) semi-restricted} in deciding which sample units are at risk of contributing to event k. Kelly and Lim [23] employed the concepts of the risk interval and risk set and categorized the AG, PWP-GP, PWP-TT, WLW, LWA (Lee-Wei-Amato), and other methods.

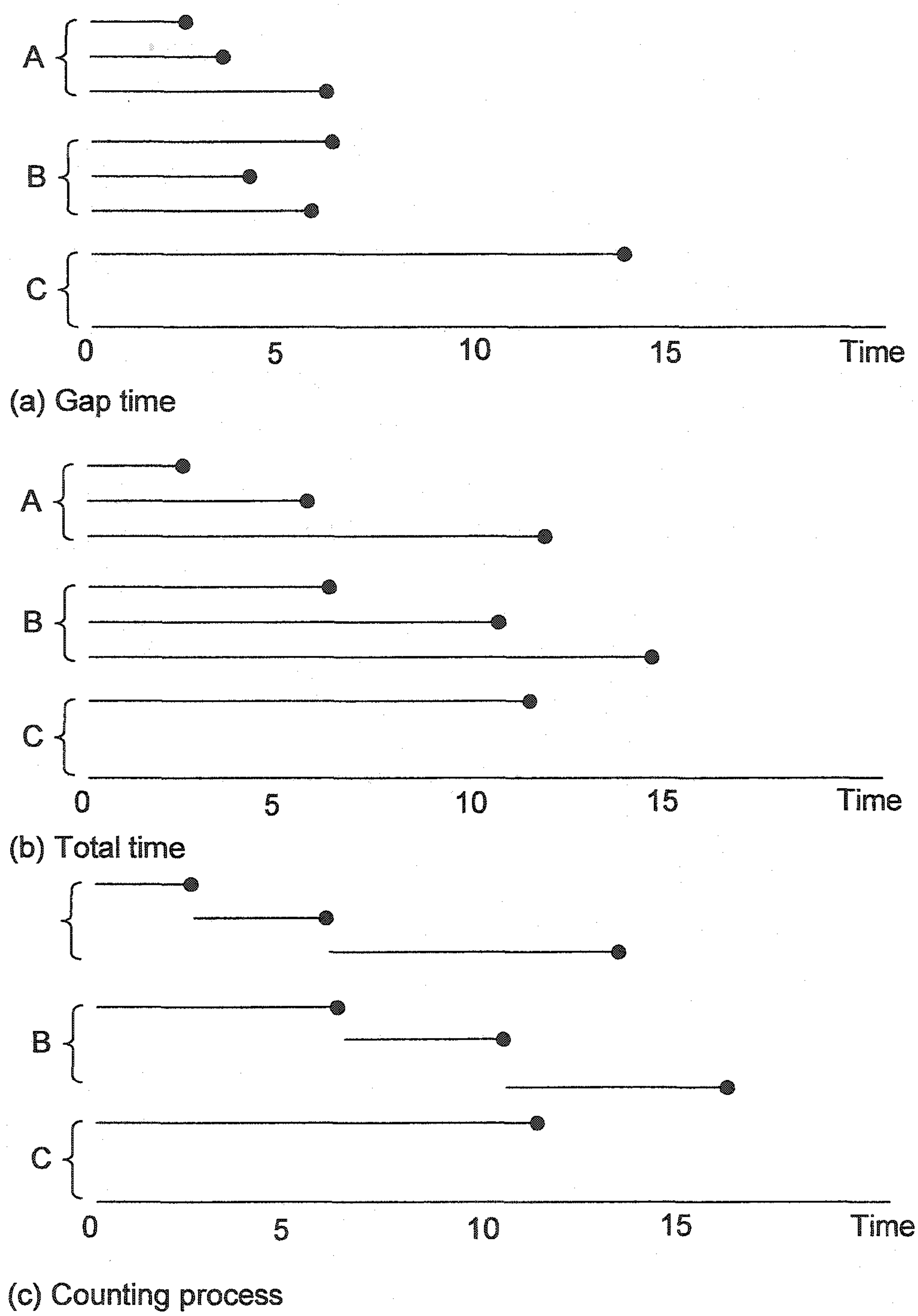


Fig. 4.1(a)-(c) Risk interval formulations (Kelly and Lim [23])

4.3 Models and methods

Sections 4.3.1-4.3.4 review the semi-parametric Cox regression model for single event and the related regression models for recurrent events. Section 4.3.5 reviews the NHPP with power-law intensity function. Section 4.3.6 describes the method used to assess the robustness of the four semi-parametric PI models for the case of censored data from a true but unknown power-law NHPP.

4.3.1 Cox regression model

For the case of a time-to-failure random variable, Cox [4] proposed a PH regression model of the form:

$$h(t;\mathbf{z}) = h_0(\mathrm{t})\exp(\boldsymbol{\beta}'\mathbf{z}), \tag{1}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The PH model is composed of two parts: baseline hazards function $h_0(t)$ and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect. The Cox model can be used to describe the semi-parametric distribution of time-to-failure for non-repairable items with covariates. Under proportional hazards, the ratio of the hazard functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline hazard function. When the baseline hazard function is fully specified (e.g.,Weibull) the analytical procedure is termed a parametric method. Alternatively, $h_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

4.3.2 Semi-parametric PWP model

The PWP model is a generalization of the semi-parametric Cox proportional hazard function to a proportional intensity function $\lambda(t;\mathbf{z})$ for the case of repeated failure events. Under proportional intensities, the ratio of the intensity functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline intensity function. When the baseline intensity function is fully specified (e.g., power-law or log-linear) the analytical procedure is termed a parametric method. Alternatively, $\lambda_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

Given the counting and covariate processes at time t, the general semi-parametric intensity function was defined by PWP as follows:

$$\lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lim \ \Pr\{t \le T_{n(t)+1} < t + \Delta \mid N(t), \mathbf{Z}(t)\} \,/\, \Delta\,, \qquad (2)$$

where $N(t)$ represents a random variable for the number of failures in $(0,t]$, $\mathbf{Z}(t)$ denotes the covariate process up to time t, and Δ limits the time span to zero. Among the semi-parametric regression models specified by PWP were the following:

$$PWP-GT: \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t - t_{n-1})\exp[\boldsymbol{\beta}_n'\mathbf{z}(t)] \qquad (3)$$

$$PWP-TT: \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t)\exp[\boldsymbol{\beta}_n'\mathbf{z}(t)]. \qquad (4)$$

In the PWP-GT model of Eq. (3), the time metric is the interval between times of failure t_{n-1} and t_n, defined as gap time. The PWP model stratifies a failure data set based on the failure event count. When a unit is placed into operation it has experienced no failures and so resides in stratum 1 ($n = 1$), and when the first failure occurs the unit moves to the second stratum ($n = 2$). In general, the unit moves to stratum n immediately following the $(n-1)^{st}$ failure and remains there until the n^{th} failure. Unlike the gap time model, the limitation of the event-specific total time model restricts the number of recurring events. Ten recurring failure events generated from a power-law NHPP in this study have shown a highly correlated relationship. Thus, the PWP-TT model is modified to a special case of Eq. (4), where the baseline intensity function is set to a common baseline intensity function denoted as $\lambda_{on}(t) = \lambda_0(t)$.

4.3.3 Semi-parametric AG model

Andersen and Gill [2] developed the AG method as an extension of the Cox PH model, to accommodate recurring events in a counting process. The AG method explains general covariate effects (common baseline intensity function in the concept of risk set), since each event count re-starts the failure process, and thus does not feature event-stratifying effects. The risk interval of an AG model follows a counting process associated with recurring events, where recurrences $(N_i^{(n)}, Y_i^{(n)}, Z_i^{(n)})$ are independent and identically distributed (i.i.d.) replicates of (N, Y, Z), and the probability of the occurrence of two events at a given time is zero. Symbols: (N, Y, Z) represent the successive failure count, an at-risk indicator, and covariates. Thus, the risk set of the $(n-1)^{st}$ event is identical to the risk set of the n^{th} event. The AG model is defined as

$$\lambda_i^{(n)}(t) = Y_i^{(n)}(\mathrm{t})\lambda_0(\mathrm{t})\exp\{\boldsymbol{\beta} \times \mathbf{z}_i^{(n)}(t)\}, \tag{5}$$

where $Y_i^{(n)}$ is an at-risk indicator and $Y_i^{(n)} = 1$ unless the subject is withdrawn from the study.

4.3.4 Semi-parametric WLW model

WLW proposed a marginal method, expanded from the conditional PWP method, in dealing with recurrent failure data. Compared to the PWP method, the WLW method has greater or equal risk set, depending on the sample size associated with the failure count. The PWP method estimates the intensity function by considering the subjects having a complete history of previous recurring events, while the WLW method additionally considers the subjects that have been withdrawn from observation. The subjects that have been censored are still in the risk set; thus, contributing influence on events that are followed after the censoring time. The risk set of each subject using the WLW method remains the same regardless of complete data or censoring events since a subject is still at risk when the subject has been withdrawn from the experiment.

Wei et al. [3] in a bladder cancer study examined treatment effects by using the PWP and WLW models about placebo and thiotepa therapies for tumor patients. This bladder cancer example collects four recurrence times of tumors $T_1 \sim T_4$ corresponding to four marginal proportional hazards models. Rather than fitting each T_i one model at a time, WLW fits four marginal models in one analysis, simultaneously. This example has two response variables {failure time and censoring status}, three covariates {treatment, tumour number, tumour size}, and four recurrent events over time.

For the k^{th} failure type and the i^{th} failure event count, the hazard function

$\lambda_{ki}(t)$ in WLW is assumed to take the form:

$$\lambda_{ki}(t) = \lambda_{k0}(\mathrm{t})\exp\{\boldsymbol{\beta}'_k \times \mathbf{z}_{ki}(t)\}, \mathrm{t} \geq 0, \tag{6}$$

where $\lambda_{k0}(t)$ is an unspecified baseline hazard function and $\boldsymbol{\beta}'_k = \left(\beta_{1k},\ldots.,\beta_{pk}\right)'$ is a vector of failure-specific regression parameters. $\mathbf{z}_{ki}(t)$ denotes a $p \times 1$ vector of covariates for the i^{th} subject at time t with respect to the k^{th} type of failure, expressed as $\mathbf{z}_{ki}(t) = (z_{1ki}(t), z_{2ki}(t), \ldots z_{pki}(t))'$.

Let X_{ki} represent the failure time of the i^{th} subject for the k^{th} type of failure and let C_{ki} represent the censoring time. $\widetilde{X}_{ki}$ are observation values of X_{ki}, where $X_{ki} = \min\{\widetilde{X}_{ki}, C_{ki}\}$. The indicator variable Δ_i is utilized for determining the event as a failure or censoring. Let $\Delta_i = 1$, when $X_{ki} = \widetilde{X}_{ki}$; otherwise $\Delta_i = 0$. Key assumptions for the WLW method are: (1) $X_{ki} \perp C_{ki}$, i.e., the failure and censoring times are independent of each other; (2) (X_i, Δ_i, Z_i) are i.i.d. random vectors, where Z_i represent covariates and i represents event count; and (3) the regression coefficients $\hat{\boldsymbol{\beta}}_i$ follow a normal distribution with mean $\overline{\boldsymbol{\beta}}_i$ denoted $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \ldots, \hat{\beta}_k) \xrightarrow{iid} Normal \ \ (\overline{\beta}_1, \overline{\beta}_2, \overline{\beta}_3, \ldots, \overline{\beta}_k)$.

Unlike the gap time model, the limitation of the event-specific total time model restricts the number of recurring events. Ten recurring failure events generated from a power-law NHPP in this study have shown a highly correlated relationship. Thus, the baseline intensity function of Eq. (6) is set to a common baseline intensity function denoted as $\lambda_{k0}(t) = \lambda_0(t)$. This simplified model is then termed as Lee-Wei-Amato (LWA) model designated to measure general covariate effects.

4.3.5 Power-law intensity function

For a power-law NHPP, the baseline intensity function is

$$\lambda_0(t) = v_0 \delta \times t^{\delta - 1}\,, \tag{7}$$

where δ is the shape parameter and υ is the scale parameter of the power-law form. If we define $\upsilon_0 = \exp(\beta_0 \times z_0)$ and $z_0 = 1$, then the power-law PI model becomes

$$\lambda(t;\mathbf{z}) = \delta \times t^{\delta - 1} \exp(\boldsymbol{\beta}'\mathbf{z})\,, \tag{8}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The power-law intensity function is composed of two parts: baseline intensity function that follows a power-law form and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect.

This process could model the reliability of a repairable system with rapid deterioration, since the failure intensity is increasing at an exponential rate with time. The analogous case for maintainability is a rapid learning process. The intensity function $\lambda(t)$ is strictly decreasing for δ < 1, constant for δ = 1, and strictly increasing for δ > 1. Thus, we have a DROCOF for δ < 1, an HPP for δ = 1, and an IROCOF for δ > 1.

4.3.6 Method

Simulation data with right-censored patterns, where the underlying distribution follows a power-law NHPP, is generated by a modified Blanks & Tordon [24] simulation algorithm. In order to simulate right-censored recurrent data, two groups of sample units were generated, in which one group contains the sample units with complete data and the other group contains the sample units with right-censored data. In the group of censored units, the right-censored pattern is set to be random. The ratio (probability) of the sample units that have censored times to total sample units is defined as censored probability (P_c).

A discrete indicator covariate $\mathbf{z}_1$ was used to separate the data into two strata for an arbitrary treatment effect. For consistency with the work of Qureshi et al. [6], simulated data was generated from a power-law NHPP with like parameter values. A proportional intensity function dataset was created using two different values for the scale parameter ($\upsilon_0 = 0.001, \upsilon_1 = 0.01$) corresponding to the two values of the indicator covariate $\mathbf{z}_1$ ($\mathbf{z}_1 = 0, \mathbf{z}_1 = 1$).

There are three experimental factors: experimental units (U), shape parameter (δ), and censoring probability (P_c). The levels for each factor are selected as follows: (1) U = 60, 120, and 180 (2) δ=0.5, 0.8, 1.0, 1.2, 1.5, 1.8, and 2.0 (3) P_c= 0.0, 0.4, 0.6, 0.8, and 1.0. The selection of the U, δ, and P_c levels has taken the following considerations: (1) the parameter settings in the previous relevant works (Proschan [25], Landers and Soroudi [5], Qureshi et al. [6], and Landers et al. [8]) (2) Severe right-censorship may cause the small sample size (U=20) to have insufficient data. The selection of P_c levels takes into account the light, moderate, and severe censoring. The selection of U and δ levels is taken from the parameter settings in the previous research works, and it has also considered the small, median, and large sample sizes for U.

To implement the four Cox-Based regression methods (PWP-GT, PWP-TT, AG, and WLW), requires formulation of three types of datasets (i.e. three formats for the same set of failure events, according to the theory underlying each methodology). For the AG method, the data set is formed from the time interval (T_1, T_2) defined as starting and ending times of an event with respect to the following counting process formulation:

$$\lim_{\Delta \to 0} \frac{1}{\Delta} p\left[N(t+\Delta) - N(t) = 1 \mid T > t\right] = \lambda(t), \tag{9}$$

where

$\lambda(t)$: proportional intensity function of failure process,
$N(t)$: random variable for number of failures in $(0,t]$.

Eq. (9) defines the instantaneous failure rate between t and $t+\Delta$ under the condition that this individual has survived after time t. Thus, the logic rule to form the dataset is: $T_2 > T_1$. As a result, all the censored failure times are removed from the dataset since $T_2 = T_1$ when it is a censored event as stipulated for the AG method. The concept of forming the dataset for the PWP method originates from the probability theory of conditionality. The later failure times after the n^{th} failure count cannot be included into the dataset when the intensity function at the n^{th} failure count is estimated. That is, for each censored unit, the censored times are removed from the dataset except for the first censored event count. Due to the marginal probability theory of the WLW method, the dataset contains full records including all censored events, such that censored units remain in the risk set.

The four semi-parametric methods were implemented using the SAS™ Users Group (SUGI) software code PHREG [26], which performs the semi-parametric Cox regression method with a blocking option to stratify for a covariate, such as failure event count, not satisfying the proportional hazards conditions. SAS PHREG includes a blocking option that stratifies events in strata defined by the event count and thereby provides event-specific intensity functions. PHREG applies the product-limit method to estimate the reliability function within all strata defined by the failure count and for all values of the covariate. PHREG also applies the Cox method to estimate the vector of regression coefficients $\boldsymbol{\beta}$ and the covariance matrix. Appendix III provides the programming code to perform the four semi-parametric methods.

To measure and compare model performance, three robustness metrics were compiled:

- relative signed error (BIAS);
- relative mean absolute deviation (MAD) and
- relative mean squared error (MSE).

The estimates of the PWP-GT regression coefficients were also compared to the theoretical value based on ten failures per unit. Additionally, 95% confidence intervals were constructed on the estimates of $\boldsymbol{\beta}_i$. In the special case of an HPP, the other three models (i.e., PWP-TT, AG, and WLW), which have common baseline intensity function in the concept of risk set, were compared and their 95% confidence bounds were constructed about the PWP-TT, AG, and WLW estimates.

4.4 Results

4.4.1 PWP-GT model results

This section examines the PWP-GT model robustness in estimating the covariate effect denoted as $\hat{\boldsymbol{\beta}}_i$. Three experimental factors are experimental units (U), shape parameter (δ), and censoring probability (P_c). Table 4.1 summarizes the robustness across strata defined by ordered failures. In the case of $U = 60$, results for censoring probability P_c from 0.4 to 1.0 are as follows. For the range of the shape parameter, $0.8 \le \delta \le 2.0$, with censoring probability $P_c = 0.4$, the PWP-GT estimates have relative MSE in the range of (1.1%, 17.5%), relative BIAS in the range of (-0.6%, 18.0%), and relative MAD in the range of (8.6%, 27.0%). As the value of P_c is increased to 0.6, the PWP-GT estimates have relative MSE in the range of (1.7%, 18.7%), relative BIAS in the range of (-1.6%, 17.0%), and relative MAD in the range of (11.2%, 29.3%). Likewise, when P_c is increased to 0.8, the PWP-GT estimates have relative MSE in the range of (1.6%, 19.3%), relative BIAS in the range of (-5.9%, 16.0%), and relative MAD in the range of (8.8%, 29.8%). However, when P_c is increased to 1.0, the PWP-GT estimates deteriorate substantially, with relative MSE in the range of (9.5%, 247.9%), relative BIAS in the range of (-22.1%, 65.6%), and relative MAD in the range of (24.1%, 79.8%). Among all shape parameters $0.5 \le \delta \le 2.0$, $\delta = 1.5$ has the most robust PWP-GT estimates throughout $0.4 \le P_c \le 1.0$, with relative MSE in the range of (6.2%, 9.5%), relative BIAS in the range of (5.3%, 11.7%), and relative MAD in the range of (19.0%, 24.1%). If the censoring probability is controlled below 0.8, the PWP-GT estimates perform well at the shape parameter range of $0.8 \le \delta \le 2.0$.

As for the case of $U = 120$, for the range of shape parameter $0.8 \le \delta \le 2.0$ with $P_c = 0.4$, the PWP-GT estimates have relative MSE in the range of (0.6%, 10.8%), relative BIAS in the range of (-4.5%, 16.0%), and relative MAD in the range of (5.9%, 16.5%). As the value of P_c is increased to 0.6, the PWP-GT estimates have relative MSE in the range of (0.5%, 10.9%), relative BIAS in the range of (-5.6%, 14.8%), and relative MAD in the range of (5.0%, 18.6%). Likewise, when P_c is increased to 0.8, the PWP-GT estimates have relative MSE in the range of (0.9%, 21.9%), relative BIAS in the range of (2.7%, 14.3%), and relative MAD in the range of (7.5%, 22.6%). However, when P_c is increased to 1.0, the PWP-GT estimates deteriorate substantially, with relative MSE in the range of (8.3%, 168.4%), relative BIAS in the range of (16.4%, 45.8%), and relative MAD in the range of (18.8%, 48.1%). Among all shape parameters $0.5 \le \delta \le 2.0$, the more favorable applications range of the PWP-GT estimates is $0.8 \le \delta \le 2.0$ throughout $0.4 \le P_c \le 0.8$, with relative MSE in the range of (0.5%, 21.9%), relative BIAS in the range of (-5.6%, 18.7%), and relative MAD in the range of (5.0%, 22.6%).

In the case of $U = 180$, the heaviest censoring probability ($P_c = 1.0$) is less damaging, compared to $U = 60$ and $U = 120$, except for the rapidly decreasing ROCOF, $\delta = 0.5$, where the PWP-GT estimates in $\delta = 0.5$ are on the rise in the range of $0.8 \le P_c \le 1.0$. Table 1 indicates that for shape parameter in the range of

$0.8 \le \delta \le 2.0$ and censoring probability in the range of $0.4 \le P_c \le 1.0$, the PWP-GT estimates have relative MSE in the range of (0.2%, 12.8%), relative BIAS in the range of (-7.0%, 17.8%), and relative MAD in the range of (3.3%, 21.8%).

Censoring probability and experimental units were chosen as the two factors to present in 3-D charts. Based on each shape parameter, 3-D charts were generated to present the PWP-GT model results. Figs.4.2 (a)-(c) provide the robustness evaluation of the PWP-GT model for three power-law intensity functions, and indicate that the error is on the rise as the censoring probability increases and the error is on the decrease as the sample size is increased. The sample size effect is exacerbated by heavy censoring. The BIAS values at $(U,\delta,P_c)=(60,1.8,1.0)$ and $(U,\delta,P_c)=(60,2.0,1.0)$ from Table 4.1 indicate a negative value due to high variability from the heavy censoring factor. As the sample size is increased to 120 and 180, the number of sample units compensates the heavy censoring effect.

Table 4.1 Summary of PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit)

N =10 failure events/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$													
δ	P_c	U	BIAS	MAD	MSE	U	BIAS	MAD	MSE	U	BIAS	MAD	MSE
0.5	0.4	60	0.40407	0.62746	0.53551	120	0.00465	0.33390	0.36198	180	-0.18178	0.18178	0.06075
0.5	0.6	60	0.48680	0.71696	0.77886	120	-0.00555	0.33308	0.34196	180	-0.18696	0.18696	0.06217
0.5	0.8	60	0.62978	0.86481	1.36832	120	-0.01706	0.31954	0.30120	180	-0.09188	0.24669	0.12328
0.5	1.0	60	1.00221	1.20024	2.27255	120	0.40857	0.69729	1.09745	180	0.06185	0.40939	0.27329
0.8	0.4	60	-0.00612	0.09325	0.01291	120	-0.04467	0.09571	0.01563	180	-0.05950	0.08415	0.01140
0.8	0.6	60	-0.01597	0.11158	0.01670	120	-0.05590	0.09200	0.01477	180	-0.07045	0.09304	0.01306
0.8	0.8	60	-0.05920	0.10437	0.01871	120	0.08223	0.22646	0.21895	180	-0.05754	0.09754	0.01753
0.8	1.0	60	0.65585	0.79796	2.47916	120	0.21174	0.34510	0.62236	180	-0.03674	0.12327	0.02774
1.0	0.4	60	0.04178	0.08594	0.01143	120	0.00232	0.05852	0.00590	180	-0.00516	0.03339	0.00195
1.0	0.6	60	0.00454	0.13398	0.02750	120	-0.00506	0.04979	0.00455	180	-0.01788	0.04288	0.00361
1.0	0.8	60	-0.00877	0.08845	0.01590	120	0.02708	0.07464	0.00923	180	-0.01083	0.06113	0.00719
1.0	1.0	60	0.32151	0.41818	0.63632	120	0.34912	0.38596	1.09905	180	0.00855	0.06957	0.00984
1.2	0.4	60	0.06894	0.11685	0.01989	120	0.04091	0.07307	0.00864	180	0.03653	0.05888	0.00615
1.2	0.6	60	0.07725	0.16187	0.03993	120	0.02710	0.06802	0.00811	180	0.01795	0.06244	0.00806
1.2	0.8	60	0.03604	0.13645	0.03159	120	0.04122	0.07453	0.00894	180	0.01806	0.07786	0.01176
1.2	1.0	60	0.21679	0.29794	0.32410	120	0.45836	0.48052	1.68430	180	0.02747	0.10220	0.01642
1.5	0.4	60	0.11742	0.19005	0.06218	120	0.09289	0.11596	0.02970	180	0.08892	0.11935	0.03006
1.5	0.6	60	0.11172	0.22042	0.07480	120	0.08247	0.11322	0.02931	180	0.07461	0.11632	0.03304
1.5	0.8	60	0.07272	0.21307	0.08060	120	0.08892	0.11935	0.03006	180	0.06877	0.13615	0.03924
1.5	1.0	60	0.05251	0.24110	0.09508	120	0.36274	0.40576	0.75671	180	0.07652	0.15275	0.04380
1.8	0.4	60	0.15846	0.24172	0.12377	120	0.13317	0.16240	0.07150	180	0.13741	0.15393	0.07494
1.8	0.6	60	0.14878	0.26557	0.13549	120	0.12388	0.15561	0.07106	180	0.11684	0.16411	0.07669
1.8	0.8	60	0.12972	0.27344	0.14486	120	0.12368	0.15971	0.07195	180	0.10765	0.16802	0.08241
1.8	1.0	60	-0.22053	0.63335	1.79361	120	0.16373	0.18841	0.08326	180	0.12651	0.19201	0.08858
2.0	0.4	60	0.17962	0.26974	0.17465	120	0.15956	0.16452	0.10789	180	0.17793	0.19784	0.10950
2.0	0.6	60	0.16955	0.29295	0.18666	120	0.14829	0.18587	0.10932	180	0.14400	0.19210	0.11578
2.0	0.8	60	0.15996	0.29812	0.19294	120	0.14281	0.18920	0.11142	180	0.12980	0.19181	0.12272
2.0	1.0	60	-0.21764	0.71010	2.14182	120	0.18721	0.22697	0.12494	180	0.14455	0.21804	0.12772

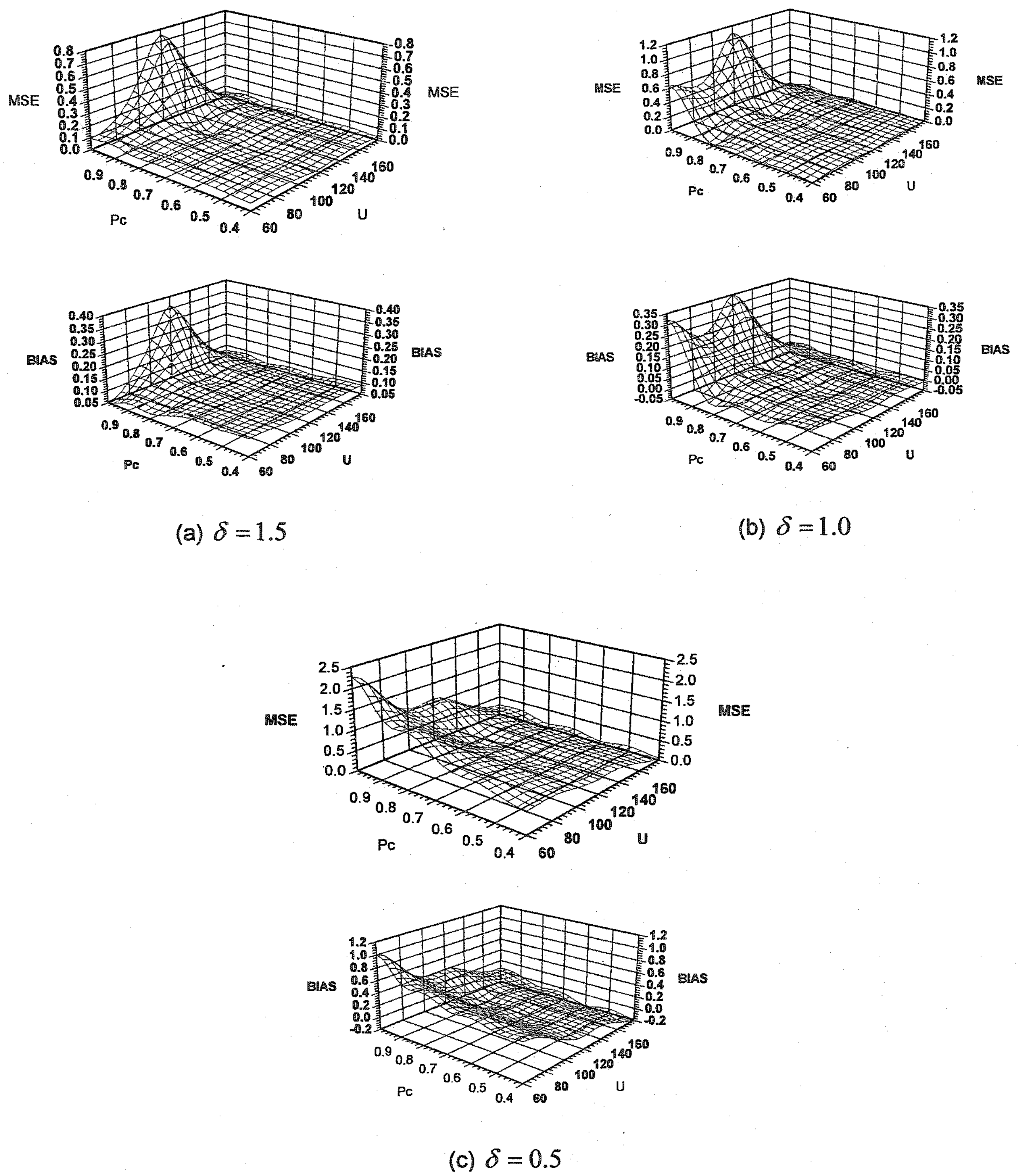


Fig. 4.2(a)-(c) PWP-GT model results for estimating $\hat{\boldsymbol{\beta}}_i$ (10 failures/unit), $\delta = 0.5$, $\delta = 1.0$, and $\delta = 1.5$

4.4.2 Effect of heavy censoring

To demonstrate that heavy censoring is producing a sample size effect, the number of replications was increased for the case of $\delta = 1.5$ and $P_c = 1.0$, where in Figs.4.2 (a)-(b) the errors are higher for $U = 120$ than for $U = 60$. Note that $U = 180$ performs the best among the three sample sizes. This section examines the 3-D error plots of the PWP-GT model for estimating $\hat{\boldsymbol{\beta}}_i$ by doubling the number of replications. Fig. 4.3 contains the comparisons (MSE and BIAS) of 3 replicates and 6 replicates in performing the PWP-GT model for the case of $\delta = 1.5$, and indicates that the error of U = 60 is not better than U = 120 as the number of replicates is increased.

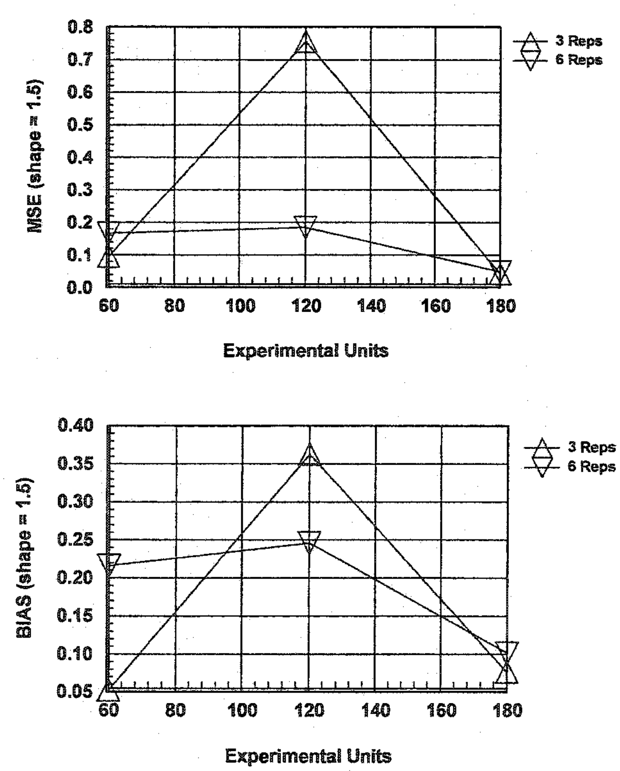


Fig. 4.3 Performance comparisons of 3 and 6 replicates (PWP-GT), $\delta = 1.5$

4.4.3 Complete data

This section compares the effects of right-censoring versus the base case of complete data ($P_c = 0$) for three values of shape parameter and three sample sizes. Tables 4.2 and 4.3 list the performance metrics. Table 4.2 examines three power-law intensity functions at sample size U =120, while Table 4.3 examines two other sample sizes for $\delta = 1.5$.

At the sample size U =120, three charts (Fig. 4.4, Table 4.2) of $\delta = 1.5$, $\delta = 1.0$, and $\delta = 0.5$ in the vertical order all indicate that the error is on the rise as the censoring probability increases. Among the five censoring probability levels, P_c =1.0 presents much higher error values compared to P_c =0, 0.4, 0.6, and 0.8. For shape parameter of $\delta = 1.5$, the performance metrics of the three sample sizes (U =60, 120, and 180) are documented in Table 4.3 and Fig. 4.4 (in the horizontal order). As the censoring probability increases, MAD and MSE are on the rise. Sample size effect is significant at heavy censoring, P_c =1.0.

Table 4.2 Performance metrics (PWP-GT) in three power-law intensity functions

N = 10 failures/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$					
U	δ	P_c	BIAS	MAD	MSE
120	0.5	0	-0.09376	0.24805	0.12655
		0.4	0.00465	0.33390	0.36198
		0.6	-0.00555	0.33308	0.34196
		0.8	-0.01706	0.31954	0.30120
		1	0.40857	0.69729	1.09745
120	1.0	0	-0.02346	0.05080	0.00380
		0.4	0.00232	0.05852	0.00590
		0.6	-0.00506	0.04979	0.00455
		0.8	0.02708	0.07464	0.00923
		1	0.34912	0.38596	1.09905
120	1.5	0	0.06580	0.09571	0.02556
		0.4	0.09289	0.11596	0.02970
		0.6	0.08247	0.11322	0.02931
		0.8	0.08892	0.11935	0.03006
		1	0.36274	0.40576	0.75671

Table 4.3 Performance metrics (PWP-GT) in three sample sizes, $\delta = 1.5$

N = 10 failures/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$					
U	δ	P_c	BIAS	MAD	MSE
60	1.5	0	0.09024	0.16699	0.05685
		0.4	0.11742	0.19005	0.06218
		0.6	0.11172	0.22042	0.07480
		0.8	0.07272	0.21307	0.08060
		1	0.05251	0.24110	0.09508
120	1.5	0	0.06580	0.09571	0.02556
		0.4	0.09289	0.11596	0.02970
		0.6	0.08247	0.11322	0.02931
		0.8	0.08892	0.11935	0.03006
		1	0.36274	0.40576	0.75671
180	1.5	0	0.08218	0.09191	0.02785
		0.4	0.08892	0.11935	0.03006
		0.6	0.07461	0.11632	0.03304
		0.8	0.06877	0.13615	0.03924
		1	0.07652	0.15275	0.04380

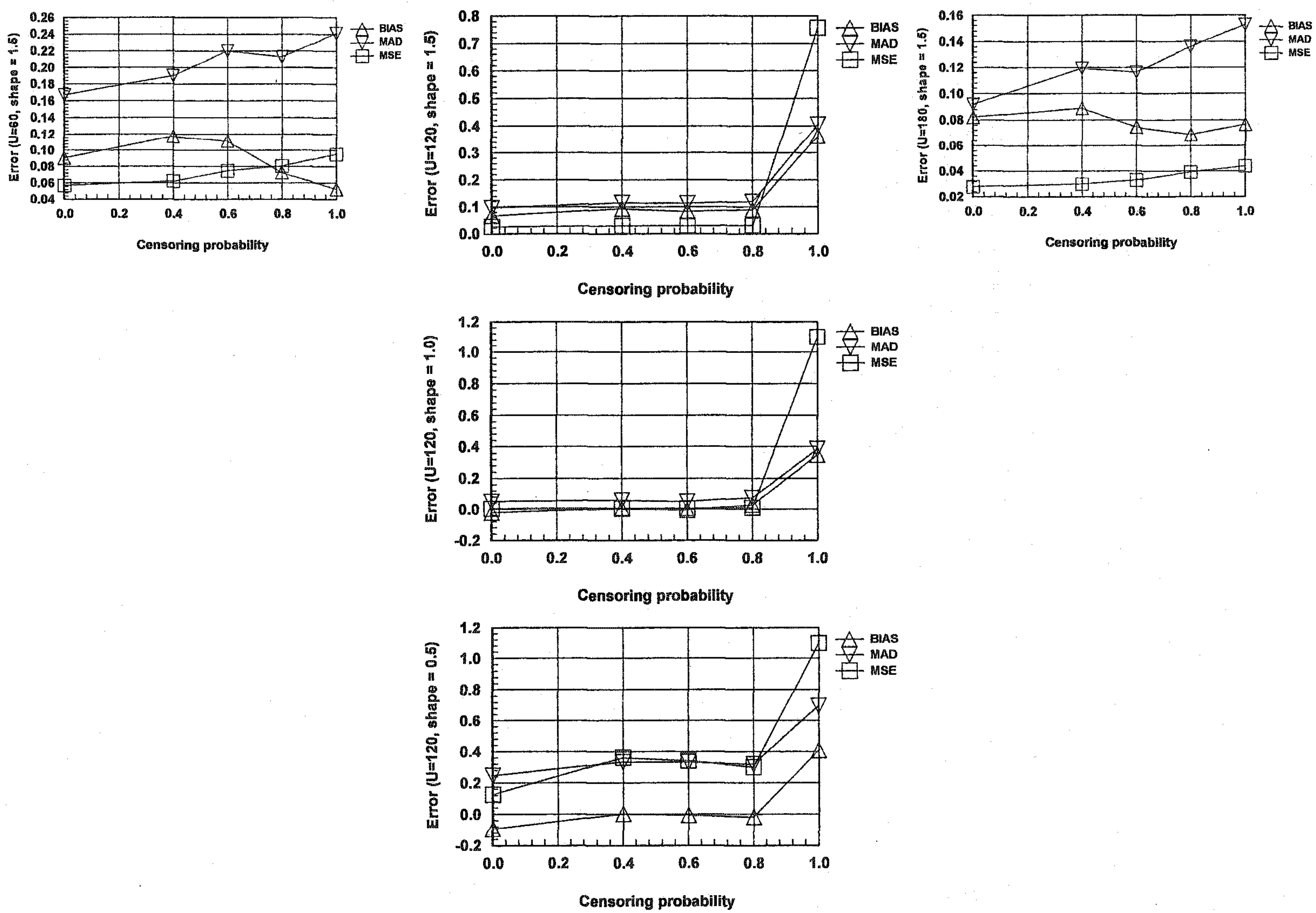


Fig. 4.4 Error plots (PWP-GT) in three shape parameters ($\delta = 1.5$, $\delta = 1.0$, and $\delta = 0.5$) and three sample sizes (U =60,120, and 180) at $\delta = 1.5$

4.4.4 95% confidence interval on $\hat{\beta}_i$

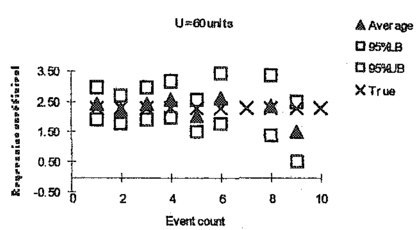
To visualize right-censoring effects upon the PWP-GT model, 95% confidence bounds were constructed on $\hat{\beta}_i$ for the HPP, where P_c is set to 1.0 (heavily censored). Three sample sizes, U = 60 (Fig. 4.5(a)), U =120 (Fig. 4.5(b)), and U =180 (Fig. 4.5(c)) at $P_c = 1.0$, are examined (Table 4.4 contains the resource data). To compare the three sample sizes, an equal range of Y-axis levels is set on [0.00, 3.50]. Due to the restriction of the range, the 7^{th} failure in $U = 120$ contains the 95% confidence interval [2.05119,3.62985], where the upper limit exceeds the maximum value 3.5.

In the case of $U = 60$ (Table 4.4, Fig. 4.5(a)), the 7^{th} and 10^{th} failures illustrate the heavy right-censoring effect, and in the case of $U = 120$ (Table 4.4, Fig. 4.5(b)), only the 95% C.I. for 10^{th} failure shows heavy censoring effect. The high variability of the PWP-GT estimate at 7^{th} or 10^{th} failure indicates a random pattern. When the sample size is increased to $U = 180$ (Table 4.4, Fig. 4.5(c)), the PWP-GT model $\hat{\beta}_i$ estimates at each failure event count are sufficient to provide tight bounds on $\hat{\beta}_i$ estimates. PWP-GT estimates tend to fluctuate more and the 95% C.I. limits tend to become wider as the event count progresses.

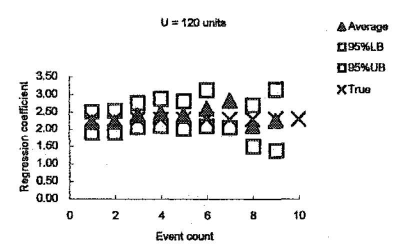
Table 4.4 95% C.I. on $\hat{\boldsymbol{\beta}}_i$, $(\delta, P_c) = (1.0,1.0)$, for three sample sizes, 60, 120, and 180

n	U	Average[a]	95%LB	95%UB	U	Average	95%LB	95%UB	U	Average	95%LB	95%UB
1	60	2.43927(0.25991)	1.92986	2.94868	120	2.20402(0.15571)	1.89883	2.50921	180	2.25030(0.12981)	1.99588	2.50472
2	60	2.25953(0.23653)	1.79593	2.72313	120	2.21734(0.16274)	1.89837	2.53631	180	2.30831(0.13910)	2.03567	2.58094
3	60	2.42571(0.27594)	1.88488	2.96655	120	2.41301(0.18051)	2.05921	2.76681	180	2.25617(0.14502)	1.97193	2.54041
4	60	2.56670(0.30891)	1.96126	3.17215	120	2.49076(0.20193)	2.09499	2.88653	180	2.48872(0.17226)	2.15110	2.82635
5	60	2.03069(0.26970)	1.50208	2.55930	120	2.41393(0.20798)	2.00630	2.82156	180	2.48429(0.17698)	2.13742	2.83116
6	60	2.60332(0.41418)	1.79155	3.41509	120	2.61077(0.26371)	2.09390	3.12763	180	2.49868(0.19974)	2.10721	2.89015
7	60	6.46199(126.22198)	-240.92818	253.85217	120	2.84052(0.40273)	2.05119	3.62985	180	2.62891(0.26423)	2.11103	3.14680
8	60	2.34825(0.50925)	1.35015	3.34636	120	2.10121(0.30771)	1.49810	2.70432	180	2.18766(0.26295)	1.67228	2.70304
9	60	1.50453(0.51508)	0.49498	2.51407	120	2.26362(0.45061)	1.38045	3.14679	180	2.30608(0.34864)	1.62275	2.98941
10	60	5.78885(459.66699)	-895.14055	906.71826	120	9.50955(154.25540)	-292.82503	311.84413	180	1.81369(0.46808)	0.89626	2.73112

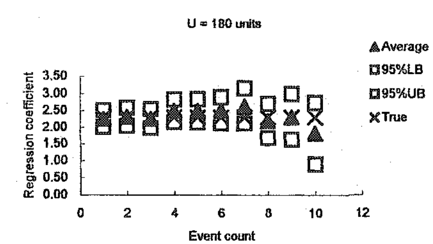
[a]: True β =2.30259



(a) U =60



(b) U =120



(c) U =180

Fig. 4.5(a)-(c) 95%CI on PWP-GT estimates at each failure count for three sample sizes 60, 120, and 180, $(\delta, P_c) = (1.0,1.0)$

4.4.5 PWP-TT, AG, and WLW models

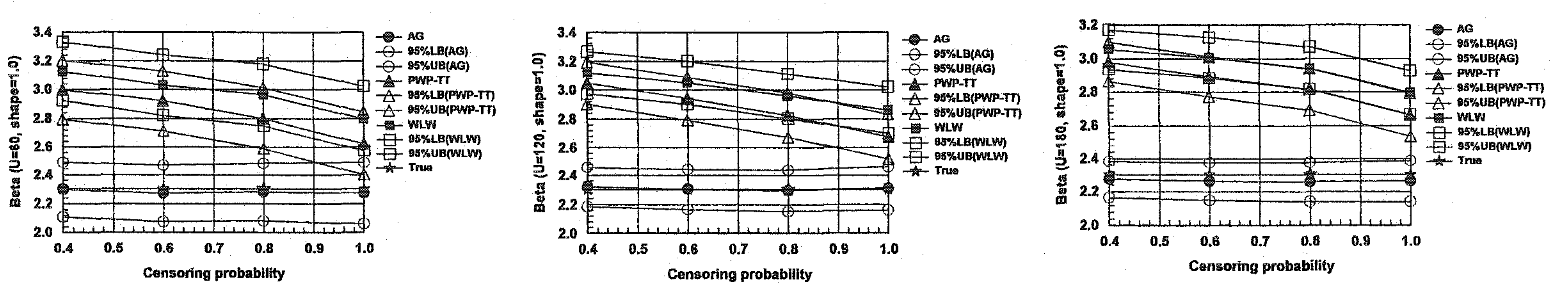
Fig. 4.6(a)-(c) visualize the estimating performance of the PWP-TT, AG, and WLW models for three sample sizes (U =60, 120, and 180) in an HPP case ($\delta = 1$). As the censoring probability increases, the AG estimate does not fluctuate, while the PWP-TT and WLW estimates slightly decrease. The sample size effect does improve the variability of the estimate resulting in narrower 95% confidence intervals. The total-time models (PWP-TT, AG, and WLW) are not affected by the shape parameter δ compared to the gap-time model (PWP-GT). The estimate and its variability using the PWP-TT, AG, or WLW model remain the same as the shape parameter setting varies. The reason the PWP-TT, AG and WLW estimates do not vary with shape parameter is that shape parameter does not influence the likelihood function in the total-time model. Consequently, the HPP case is chosen for the purpose of illustrating the PWP-TT, AG, and WLW models. The AG estimate provides the most reliable estimate in a right-censoring HPP case, among the PWP-TT, AG, and WLW models. The true β is 2.30259, and Table 4.5 summarizes the results that the true β lies within the 95% C.I. of the AG estimate in each combination of experimental units and censoring probability.

Table 4.5 Summary of semi-parametric AG model results for $\hat{\beta}$ ($\delta = 1$, an HPP)

Conditions	AG estimates [a]	95%LB	95%UB
(U, P_c)=(60,0.4)	2.29963(0.09692) [b]	2.10967	2.48959
(U, P_c)=(60,0.6)	2.27179(0.09984)	2.07611	2.46747
(U, P_c)=(60,0.8)	2.28192(0.10185)	2.08230	2.48155
(U, P_c)=(60,1.0)	2.27442(0.10851)	2.06175	2.48709
(U, P_c)=(120,0.4)	2.32087(0.06927)	2.18511	2.45664
(U, P_c)=(120,0.6)	2.30481(0.07077)	2.16611	2.44351
(U, P_c)=(120,0.8)	2.29507(0.07297)	2.15206	2.43808
(U, P_c)=(120,1.0)	2.31017(0.07653)	2.16018	2.46015
(U, P_c)=(180,0.4)	2.27485(0.05541)	2.16624	2.38346
(U, P_c)=(180,0.6)	2.26272(0.05726)	2.15049	2.37495
(U, P_c)=(180,0.8)	2.25832(0.05936)	2.14197	2.37467
(U, P_c)=(180,1.0)	2.26412(0.06257)	2.14149	2.38675

[a] Theoretical values of $\hat{\beta} = -\frac{1}{\delta}\ln\left(\frac{\upsilon_0}{\upsilon_1}\right)$ = 2.30259

[b] Estimated standard errors in parenthesis



(a) U = 60 (b) U = 120 (c) U = 180

Fig. 4.6(a)-(c) AG, PWP-TT, and WLW model results, U = 60, 120, and 180 ($\delta = 1.0$)

4.5 Conclusions

Previous studies (by Landers and Soroudi [5] and Qureshi et al. [6]) conducted on the PWP-GT model for the case of an underlying NHPP with power-law intensity function indicated good performance. This research has performed a right-censorship robustness study and examined other semi-parametric PI models with covariates for the case of right-censoring. Qureshi et al. [6] examined the PWP-GT model applied to recurrent data without censoring (complete data) and concluded that the PWP-GT estimator underestimates the covariate effect in a DROCOF case and overestimates the covariate effect in an IROCOF case. Qureshi et al. proved the PWP-GT model an accurate estimator in estimating the times to failures for NHPP power-law processes with shape parameter in the range $1.0 \le \delta \le 3.0$ and for larger sample sizes ($U \ge 60$). In comparing with other researchers, Section 4.4.3 examined both cases: complete data and right-censoring data. Section 4.4.3 has included the case of Qureshi's work (complete data) and produced results consistent with those of Qureshi.

The PWP-GT and AG prove to be models of choice, evaluated in terms of the BIAS, MAD, and MSE of covariate regression coefficients over ranges of sample sizes, shape parameters, and censoring severity encountered in engineering applications. The research domains of the three factors of interests are: (1) $60 \le U \le 180$, (2) $0.5 \le \delta \le 2.0$, and (3) $0.0 \le P_c \le 1.0$. The more favorable engineering applications ranges may be inferred from the results, as follows. At the smaller sample size (U = 60), the PWP-GT proves to perform well for moderate right-censoring ($0.0 \le P_c \le 0.8$) and moderately decreasing, constant, and moderately increasing ROCOF (power-law NHPP shape parameter in the range of $0.8 \leq \delta \leq 1.8$). In the case of U = 120, the PWP-gap time proves to perform well for moderate right-censoring ($0.0 \leq P_c \leq 0.8$) and moderately decreasing, constant, and moderately increasing ROCOF (power-law NHPP shape parameter in the range of $0.8 \leq \delta \leq 2.0$). For the large sample size (U = 180), the PWP-GT performs well for heavy right-censoring ($0.0 \leq P_c \leq 1.0$) and moderately decreasing, constant, and moderately increasing ROCOF (power-law NHPP shape parameter in the range of $0.8 \leq \delta \leq 2.0$). The AG model proves to outperform the WLW for a stationary process (HPP) across a wide range of right-censorship ($0.0 \leq P_c \leq 1.0$) and for sample sizes of 60 (30 per class) or more.

The sample sizes chosen for this engineering research were $60 \leq U \leq 180$ (30-90 units per class of a two-level covariate). Many of the medical studies reported in the literature contain sample sizes smaller than this range. Small sample sizes are common in medical studies due to the high cost of a clinical trial including requisite examinations required in the medical practices, such as X-Ray scans and blood tests. The numbers of qualified subjects are sometimes small because of cost and/or medical conditions. The AG model is designed to estimate the general covariate effect and can be useful if small sample size is unavoidable. The AG model adopts the stationary counting process model and assumes each occurrence as independent and identically distributed according to an exponential distribution. Thus, the number of observation for a subject can be utilized as the expansion of the sample size. For instance, there are ten subjects available and four observations for each subject are collected. The sample size effect for those ten subjects in using the AG model is equivalent to having forty subjects.

This research has addressed only the case of data from an NHPP with power-law intensity function. The log-linear intensity function is also encountered in the literature and may be important for industry. Future research to examine the right-censoring effect upon recurring events from an NHPP with log-linear intensity function could be beneficial to practitioners. Left-censoring also arises in some applications for recurrent failure data from repairable systems. An example case is field data where early life events were not recorded and records were lost. Future research could apply the methodology to examine PWP-GT robustness under left-censoring.

4.6 References

[1] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika, 1981; 68: 373-9.

[2] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 1982; 10(4): 1100-20.

[3] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, Dec.1989; 84(408): 1065-73.

[4] Cox DR. Regression models and lifetables (with discussion). Journal of the Royal Statistical Society B, 1972; 34:187-202.

[5] Landers TL, Soroudi HE. Robustness of a semi-parametric proportional intensity model. IEEE Transactions on Reliability, 1991; 40(2): 161-4.

[6] Qureshi WM, Landers TL, Edward EG. Robustness evaluation of a semi-parametric proportional intensity model. Reliability Engineering and System Safety, 1994; 44: 103-9.

[7] Vithala S. Robustness of a semi-parametric proportional intensity model for the case of a log-linear Non-homogeneous Poisson Process. A thesis of the Industrial Engineering Department at the University of Arkansas, 1994.

[8] Landers TL, Jiang ST, Peek, JR. Semi-parametric PWP model robustness for log-linear increasing rates of occurrence of failures. Reliability Engineering and System Safety, 2001; 73: 145-53.

[9] Ansell JI, Phillips MJ. Practical problems in the statistical analysis of reliability data (with Discussion). Applied Statistics, 1989;38(2):205-47.

[10] Ansell JI, Phillips MJ. Practical reliability data analysis. Reliability Engineering and System Safety, 1990;28(3):337-56.

[11] Ansell JI, Phillips MJ. Practical aspects of modeling of repairable systems data using proportional hazards models. Reliability Engineering and System Safety, 1997;58(2):165-71.

[12] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Assessing the maintenance in a process using a semi-parametric approach. Quality and Reliability Engineering International, 2001;17:163-7.

[13] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Analysing maintenance data to gain insight into systems performance. Journal of Operational Research Society, 2003;54:343-9.

[14] Ascher H. Regression analysis of repairable systems reliability. In: Skwirzynski JK, editor. Electronic systems effectiveness and life cycle costing. Berlin: Springer, 1983, p.119-33.

[15] Ascher H, Feingold H. Repairable systems reliability—modeling inference, misconceptions and their causes, Lecture notes in

statistics, New York: Marcel Dekker Inc, 1984;7.
[16] Dale C. Application of the proportional hazards model in the reliability field. Reliability Engineering, 1985;10:1-14.
[17] Wightman DW, Bendell A. The practical application of proportional hazards modeling. Reliability Engineering, 1986;15:29-53.
[18] Bendell A, Wightman DW, Walker EV. Applying proportional hazards modeling in reliability. Reliability Engineering and System Safety, 1991;34:35-53.
[19] Bowman ME. An evaluation of statistical models for the analysis of recurrent events data: with application to needlestick injuries among a cohort of female veterinarians. A dissertation of the Preventive Medicine at the Ohio State University, 1996.
[20] Lin DY. Cox's regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine, 1994; 13: 2233-47.
[21] Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Statistics in Medicine, 1997; 16: 833-9.
[22] Lipschutz KH, Snapinn SM. Discussion of paper by Wei and Glidden. Statistics in Medicine, 1997; 16: 841-51.
[23] Kelly PJ, Lim LL-Y. Survival analysis for recurrent data: An application to childhood infectious diseases. Statistics in Medicine, 2000; 19: 13-33.
[24] Blanks HS, Tordon MJ. Laplace and Mann aging trend test effectiveness. Proceeding of the Annual Reliability and Maintainability Symposium, 1987;165-70.
[25] Proschan F. Theoretical explanation of observed decreasing failure rate. Technometrics, Aug. 1963;5(3):375-83.
[26] SUGI Supplementary Library User's Guide, Version 5. SAS Institute, Cary, North Carolina, USA, 1985, p. 437-66.

5. Robustness of semi-parametric proportional intensity models for right-censored recurrent failure data from a stationary counting process

5.0 Abstract

The class of semi-parametric proportional intensity (PI) models applies to recurrent failure event modeling for a repairable system with covariates for a right-censored Homogeneous Poisson Process (HPP). Abundant federal funding received in biostatistics/ medical research has advanced the PI models to become well developed and widely referenced. Engineering applications of these four methods have been few because the models are not well known and the favorable ranges of applications have not been examined. This paper not only reports the robustness evaluation of the four PI models (Prentice-Williams-Peterson-gap time (PWP-GT), PWP-total time (PWP-TT), Andersen-Gill (AG), and Wei-Lin-Weissfeld (WLW)) under right-censorship, but also presents the comparison of the three event-specific baseline intensity function models (PWP-GT, PWP-TT, and WLW).

Landers and Soroudi (1991), Qureshi et al. (1994), and Landers et al. (2001) have examined robustness of the PWP-GT model for the case of an underlying NHPP with power-law and log-linear intensity functions and complete (uncensored) data. However, the phenomenon of censoring data is generally present in field data. This research has extended their work to the important case of right-censorship and has examined other semi-parametric PI models. This experimental design has incorporated three levels of censorship severity (light, moderate, and severe) to evaluate these four proposed PI models.

The more favorable engineering applications ranges are recommended, which are beneficial to practitioners in anticipating the more favorable applications domains and selecting appropriate PI models in applying to right-censored recurrent failure data. The PWP-gap time model has proven the most robust and accurate estimator (at the lowest error) among the three event-specific models. Compared to WLW, the PWP-TT estimator yields similar but slightly better results. The PWP-gap time presents a low-error region at the range of $120 \le U \le 180$ and $0 \le P_c \le 1$. For the small sample size $U = 60$, the more favorable applications range is $0 \le P_c \le 0.8$. For the other two estimators, when the sample size is increased from $U = 60$ to $U = 120$, PWP-TT and WLW have a slightly improved applications range $0 \le P_c \le 0.4$. As the sample size is increased to 180, the performance is poor but stable over applications range $0 \le P_c \le 0.8$ on both models. The results show that AG performs well for the case of smaller sample size (U=60) and heavy censoring ($P_c = 1.0$). The favorable applications region of the common baseline AG model lies at $60 \le U \le 180$ and $0 \le P_c \le 1$.

Keywords: repairable systems reliability, right-censored recurrent events, proportional intensity models

Nomenclature

Acronyms

AG	Andersen and Gill model
C.I.	Confidence interval
DROCOF	Decreasing rate of occurrence of failures
HPP	Homogeneous Poisson Process

IROCOF	Increasing rate of occurrence of failures
i.i.d	Independent and identically distributed
MTTF	Mean time to failure
MAD	Mean absolute deviation
MSE	Mean squared error
NHPP	Non-homogeneous Poisson Process
PI	Proportional intensity
PWP	Prentice, Williams, and Peterson model
PWP-GT	Prentice, Williams, and Peterson-gap time model
PWP-TT	Prentice, Williams, and Peterson-total time model
WLW	Wei, Lin, and Weissfeld model

Notation

C_{ki}	Censoring time for the i^{th} subject of the k^{th} type of failures
$h(t;z)$	Proportional hazard function
$h_0(t)$	Baseline hazard function
I_0	Number of sample units in class ϕ
I_1	Number of sample units in class 1
i.i.d.	Independent and identically distributed
N	Successive failure count
$N(t)$	Random variable for the number of failures in $(0,t]$; a counting process
n	An integer counting successive failure times; a stratification indicator subscript
P_c	Censoring probability
s.d.	Standard deviation
T_1, T_2	The beginning and end of an event; bivariate exponential variables
T_n	Random variable for cumulative time of occurrence of the n^{th} failure
t_n	Cumulative time of occurrence of the n^{th} failure; a realization of T_n
U	Sample size (number of units)

$\widetilde{X}$	Observation time
$Y_i^{(n)}$	an at-risk indicator in the AG model
$\mathbf{Z}(t)$	Covariate process up to time t
$\mathbf{z}$	$(k \times 1)$ vector of covariates, $\mathbf{z} = (z_1, z_2, \ldots z_k)'$
$\boldsymbol{\beta}_n$	$(k \times 1)$ vector of stratum-specific regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$
δ	Shape parameter of a power-law NHPP
Δ	Indicator of a failure or censored time; limit to time zero
λ_0	Baseline value of λ for power-law NHPP
$\lambda_0(t)$	Baseline intensity function
$\lambda_{0n}(t)$	Stratum-specific baseline intensity function
$\lambda(t;\mathbf{z})$	Proportional intensity function
υ	Scale parameter of a power-law NHPP
υ_0	Baseline value of υ, the scale parameter of a power-law NHPP
υ_1	Alternate value of υ, the scale parameter of a power-law NHPP
σ	Standard deviation
^	Denotes an estimator
'	Denotes the transpose of a vector

5.1 Introduction

Failure time data on a repairable system are realizations of a stochastic point process, in which the instantaneous rate of occurrence of failures (ROCOF) is $\lambda(t)$. Prentice, Williams, and Peterson (PWP) [1] proposed a semi-parametric approach to model recurrent failure event data from a repairable system using two methods: PWP-GT (gap time) and PWP-TT (total time). Several researchers have subsequently proposed alternate modeling methods by modifying the risk set (common or event-specific baseline intensity function) and the risk interval (gap time, total time, or counting process). These include the AG (Andersen-Gill) [2] and WLW (Wei-Lin-Weissfeld) [3] models.

Cox proposed the distribution-free (semi-parametric) proportional hazards model in 1972 [4]. The Cox-based regression models (PWP-GT, PWP-TT, AG, and WLW) have been applied to recurring events in medical studies (biostatistics field), such as recurrent infections of a patient. For engineering applications, Landers and Soroudi [5], Qureshi et al. [6], Vithala [7], and Landers et al. [8] have investigated robustness of the PWP-GT model, where the underlying recurrent failure time data are from a Non-homogeneous Poisson Process (NHPP) with a power-law or a log-linear intensity function. These references also report the engineering applications of the PWP-GT model cited in the literature. Qureshi et al. [6] found that the PWP-GT model performs best for constant and moderately increasing rate of occurrence of failures (IROCOF) and decreasing rate of occurrence of failures (DROCOF) and for larger sample sizes from power-law NHPPs. Vithala [7] considered the case of log-linear increasing rates of occurrence of failures, and concluded the PWP-GT model performs best for moderately increasing rates of occurrence of failures and for larger sample sizes. Both Qureshi et al [6] and Vithala [7] restricted their studies to the case of complete (uncensored) data. However, the phenomenon of censoring is generally present in field data. This research has extended their work to the important case of right-censorship and has examined other semi-parametric PI models (PWP-TT, AG, and WLW).

Compared to the extensive literature on applications of the Cox-based regression models in the biostatistics field, there have been few reported engineering applications. Abundant federal funding received in biostatistics / medical research has advanced the PI models to become well developed and widely referenced. PI models for medical applications could also apply to recurring failure/repair data in engineering problems. The PWP-GT, PWP-TT, AG, and WLW models are potentially powerful analytical tools for engineering practitioners as they become better recognized and understood. This paper reports the robustness of the PWP-GT, PWP-TT, AG, and WLW models for right-censored recurrent failure events and stationary data. The results are beneficial to practitioners in anticipating the more favorable applications domains and selecting appropriate PI models.

5.2 Semi-parametric Proportional Intensity models

Cox [4] proposed a proportional hazards formulation to include explanatory variables (covariates) in survival models. PWP proposed an extension of the Cox model to stochastic processes and applied the approach to model recurring infections in aplastic anemia and leukemia patients having received bone-marrow transplants. This application involved several subjects and a small number of events (up to five) for each subject. The paper by PWP did not address the baseline intensity function but rather reported the relative risks for the test and control groups. In reliability and maintainability engineering applications, a number of authors have applied the semi-parametric PI (PH) model, for example, Ansell and Phillips [9], Ansell and Phillips [10], Landers and Soroudi [5], Qureshi et al. [6], Ansell and Phillips [11], Landers et al. [8], Ansell et al. [12], and Ansell et al. [13]. A collection of the PI model applied to different industries includes: marine gas turbine engines (Asher [14]), semiconductor, electrical, and pipeline industries (Ansell and Phillips [11]), U.S. Army main battle tank (Landers et al. [8]), water supply industry (Ansell et al. [12], [13]), etc. Ascher [14] illustrated the use of the PWP model for analysis of reliability for marine gas turbine engines. Ascher and Feingold [15] suggested application of the PWP model in the field of reliability engineering. Dale [16] applied the PWP approach to simulated data for a reliability growth program with design improvements implemented after each of the five stages, resulting in a DROCOF. Wightman and Bendell [17] and Bendell et al. [18] cited the PWP model and advised caution in application for engineering studies.

Qureshi et al. [6] performed a robustness study to determine how well the PWP-GT method performed when applied to data from a failure process that was actually parametric (specifically the NHPP with power-law intensity function). The 2 σ bounds of the PWP-GT estimates can cover the true values with few exceptions. The PWP-GT method performed well, except at small values of shape parameter ($\delta < 0.6$). The PWP-GT method was best for larger sample size and for moderately decreasing, constant, and moderately increasing ROCOFs. The validation process for the case of an HPP in Section 2.2.3 (also refer to Table 2.10) indicated that the estimated $MTTF$ (mean time to failure) differences between the PWP-GT model and theoretical values were not statistically significant. As for the PWP-GT estimates of the covariate regression coefficient, the true value of coefficient β lies within the 2 σ confidence bounds on the estimate $\hat{\beta}$ for $1.0 \leq \delta \leq 1.4$. The PWP-GT method tends to underestimate β for a DROCOF (e.g., BIAS= -26% at $\delta = 0.5$) and overestimate β for an IROCOF (e.g., BIAS= 19% at $\delta = 3.0$).

The AG model (Andersen and Gill [2]) and the WLW model (Wei et al. [3]) are widely cited in the literature. Bowman [19] and Lin [20] surveyed and evaluated the AG, PWP-GT, PWP-TT, and WLW methods. Bowman conducted a simulation based on a bivariate exponential distribution to generate bivariate recurrent events, in order to control the correlation (θ) among recurring events. Bowman utilized the bivariate exponential distribution (T_1, T_2) to generate the consecutive recurring event times $T_n = T_1 + T_2$, where n is the event count and T_1 and T_2 represent the beginning and end of an event. The univariate event time T_n is composed of T_1 and T_2 with given correlation (θ). This type of simulation approach makes it possible to manage the correlation of recurring events. Bowman identified the PWP-GT model as superior and then used it to analyze needle-stick injury data.

Wei and Glidden [21] have reviewed the Cox-based methods designed to model recurrent data, and summarized the strengths and weaknesses for each method. In a commentary on the Wei and Glidden paper, Lipschutz and Snapinn [22] stressed two concepts of “event times” and “risk sets” as crucial to choosing the appropriate model. Event elapsed times are related to the total time, gap time, and counting process. The PWP-TT and WLW are modeled by total time, while only PWP-GT is modeled by gap time. The risk interval of the AG model belongs to the counting process class. Intuitively, total (global) times within a subject are highly correlated, with similar indication on the first recurrence and subsequent events. The total time model may indicate large treatment effect throughout the entire study, even though the gap time model has indicated little treatment effect beyond a certain recurrence. The counting process concept of the AG method implies each recurrence is not affected by previous events, and does not contribute to future events.

The risk set consists of the subjects at risk for a specified event (e.g., failure). There are three types of risk sets: conditional (e.g., PWP), counting process (e.g., AG), or marginal (e.g., WLW). As a marginal method, the WLW method assumes a subject is at risk regardless of event count until the observation for the subject terminates by censoring. The AG method also provides an index of a general covariate effect, which is expressed by the common baseline hazard (unrestricted risk set). However, a subject in the PWP method has event-specific baseline hazards (restricted risk set), in that the proportional intensity of event k only considers the subjects that have experienced $(k-1)$ events. Lipschutz and Snapinn [22] suggested guidelines as follows in choosing the appropriate models:

- Use total time, common baseline hazard (unrestricted risk set) when the general effect is of interest.
- Use gap time, event-specific baseline hazards (restricted risk set) when the primary concern is how the treatment will affect the recurring events beyond the first occurrence.

Kelly and Lim [23] noted that risk interval can be defined by three formulations {(1) gap time, (2) total time, and (3) counting process} demonstrated in Fig. 5.1(a)-(c). Risk interval determines whether a model is marginal in the total time or conditional in the gap time. The risk interval of any event in total time is not influenced by any previous events, but measures time from entry into the experiment (beginning of observation). However, the risk interval of the gap time begins from the end of last event (Kelly and Lim [23]). Counting processes use the total time scale and share the same elapsed time as does the gap time model. However, the risk interval starts from the previous event instead of the entry time. Based on the common or event-specific baseline intensities, the risk set is labeled as either unrestricted or restricted. Kelly and Lim [23] defined three possible risk sets {(1) unrestricted, (2) restricted, and (3) semi-restricted} in deciding which sample units are at risk of contributing to event k. Kelly and Lim [23] employed the concepts of the risk interval and risk set and categorized the PWP-GT, PWP-TT, AG, WLW, LWA (Lee-Wei-Amato), and other methods.

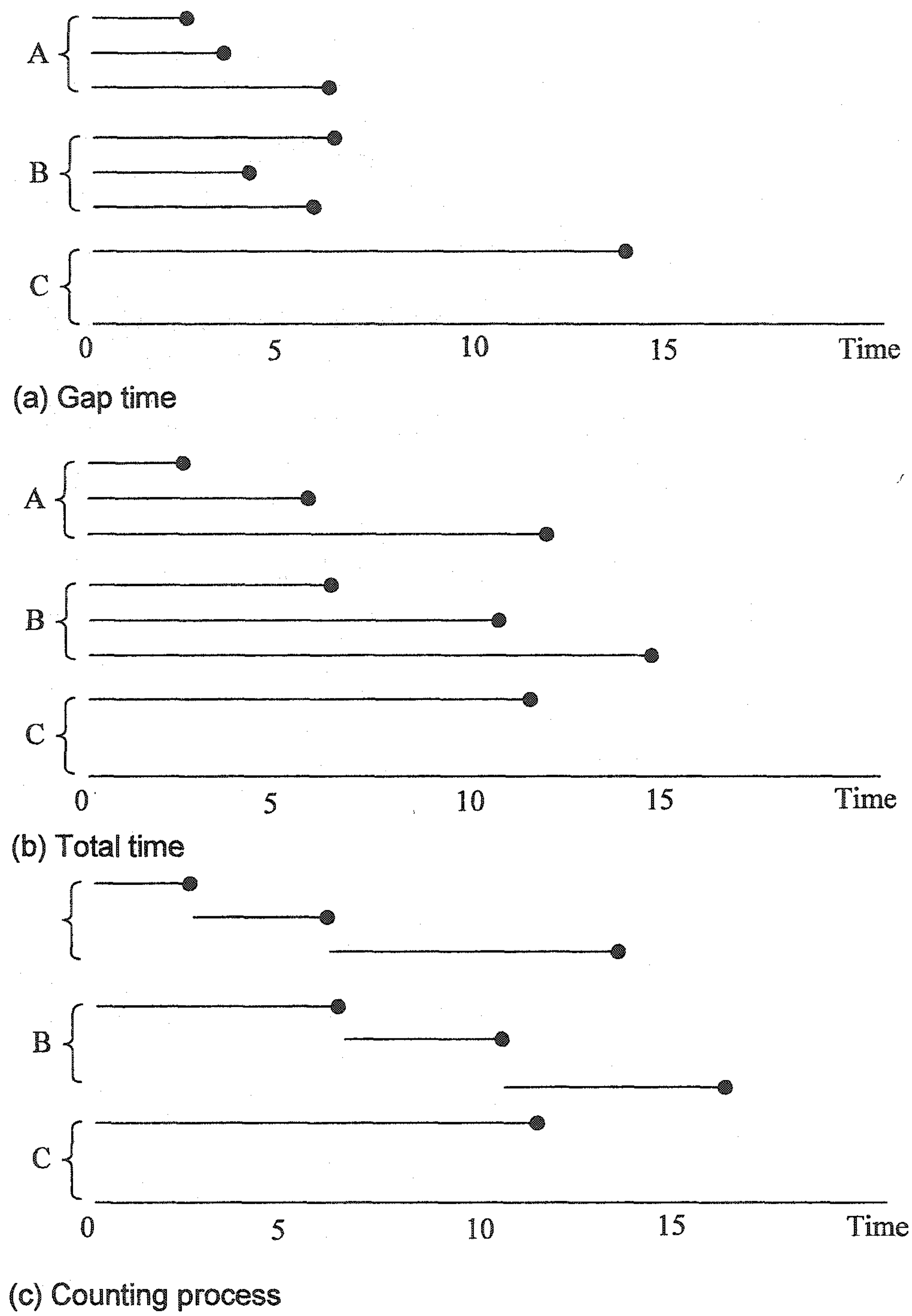


Fig. 5.1(a)-(c) Risk interval formulations (Kelly and Lim [23])

5.3 Models and methods

Sections 5.3.1-5.3.2 review the semi-parametric Cox regression model for single event and the related regression models for recurrent events. Sections 5.3.3-5.3.4 review two other alternate regression models (AG and WLW), and Section 5.3.5 describes the method used to assess the robustness of the semi-parametric PI models for the case of censored data from a true but unknown stationary counting process.

5.3.1 Cox regression model

For the case of a time-to-failure random variable, Cox [4] proposed a proportional hazards regression model of the form:

$$h(t;\mathbf{z}) = h_0(\mathrm{t})\exp(\boldsymbol{\beta}'\mathbf{z}), \tag{1}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The PH model is composed of two parts: baseline hazards function $h_0(t)$ and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect.

The Cox model can be used to describe the semi-parametric distribution of time-to-failure for non-repairable items with covariates. Under proportional hazards, the ratio of the hazard functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline hazard function. When the baseline hazard function is fully specified (e.g., Weibull) the analytical procedure is termed a parametric method. Alternatively, $h_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

5.3.2 Semi-parametric PWP model

The PWP model [1] is a generalization of the semi-parametric Cox proportional hazard function to a proportional intensity function $\lambda(t;\mathbf{z})$ for the case of repeated failure events. Under proportional intensities, the ratio of the intensity functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline intensity function. When the baseline intensity function is fully specified (e.g., power-law or log-linear) the analytical procedure is termed a parametric method. Alternatively, the $\lambda_0(t)$ baseline intensity function can be left arbitrary in which case the procedure is termed semi-parametric.

Given the counting and covariate processes at time t, the general semi-parametric intensity function was defined by Prentice, Williams and Peterson as follows:

$$\lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lim \ \Pr\{t \le T_{n(t)+1} < t + \Delta \mid N(t), \mathbf{Z}(t)\} / \Delta , \tag{2}$$

where $N(t)$ represents a random variable for the number of failures in $(0,t]$, $\mathbf{Z}(t)$ denotes the covariate process up to time t, and Δ limits the time span to zero. Among the semi-parametric regression models specified by Prentice, Williams and Peterson were the following:

$$PWP-GT : \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t - t_{n-1})\exp[\boldsymbol{\beta}_n' \mathbf{z}(t)] \tag{3}$$

$$PWP-TT : \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t)\exp[\boldsymbol{\beta}_n' \mathbf{z}(t)]. \tag{4}$$

In the PWP-GT model of Eq. (3) the time metric is the interval between times of failure t_{n-1} and t_n, defined as gap time. The PWP model stratifies a failure data set based on the failure event count. When a unit is placed into operation it has experienced no failures and so resides in stratum 1 ($n=1$), and when the first failure occurs the unit moves to the second stratum ($n=2$). In general, the unit moves to stratum n immediately following the $(n-1)^{st}$ failure and remains there until the n^{th} failure.

5.3.3 Semi-parametric AG model

Andersen and Gill [2] developed the AG method as an extension of the Cox proportional hazards model, to accommodate recurring events in a counting process. The AG method explains general covariate effects (common baseline intensity function in the concept of risk set), since each event count re-starts the failure process, and thus does not feature event-stratifying effects. The risk interval of an AG model follows a counting process associated with recurring events, where recurrences $(N_i^{(n)}, Y_i^{(n)}, Z_i^{(n)})$ are independent identically distributed (i.i.d.) replicates of (N, Y, Z), and the probability of the occurrence of two events at a given time is zero. Thus, the risk set of the $(n-1)^{st}$ event is identical to the risk set of the n^{th} event. The AG model is defined as

$$\lambda_i^{(n)}(t) = Y_i^{(n)}(\mathrm{t})\lambda_0(\mathrm{t})\exp\{\boldsymbol{\beta}\times\mathbf{z}_i^{(n)}(t)\}, \qquad (5)$$

where $Y_i^{(n)}$ is an at-risk indicator and $Y_i^{(n)}=1$ unless the subject is withdrawn from the study.

5.3.4 Semi-parametric WLW model

Wei et al. [3] proposed a marginal method, expanded from the conditional PWP method, in dealing with recurrent failure data. Compared to the PWP method, the WLW method has greater or equal risk set, depending on the sample size of the failure count. The PWP method estimates the intensity function by considering the subjects having a complete history of previous recurring events, while the WLW method additionally considers the subjects that have been withdrawn from the observation. The censored subjects are still in the risk set; thus, contributing influence on events that are followed after the censoring time. The risk set of each subject using the WLW method remains the same regardless of complete data or censoring events since a subject is still at risk when the subject has been withdrawn from the experiment.

Wei et al. [3] in a bladder cancer study examined treatment effects by using the PWP and WLW models about placebo and thiotepa therapies for tumor patients. This bladder cancer example collects four recurrence times of tumors $T_1 \sim T_4$ corresponding to four marginal proportional hazards models. Rather than fitting each T_i one model at a time, WLW fits four marginal models in one analysis, simultaneously. This example has two response variables {failure time and censoring status}, three covariates {treatment, tumour number, tumour size}, and four recurrent events over time.

For the k^{th} failure type and the i^{th} failure event count, the hazard function $\lambda_{ki}(t)$ in WLW is assumed to take the form:

$$\lambda_{ki}(t) = \lambda_{k0}(\mathrm{t})\exp\{\beta_k' \times \mathbf{z}_{ki}(t)\}, \mathrm{t} \geq 0, \tag{6}$$

where $\lambda_{k0}(t)$ is an unspecified baseline hazard function and $\boldsymbol{\beta}_k' = (\beta_{1k}, \ldots, \beta_{pk})'$ is a vector of failure-specific regression parameters. $\mathbf{z}_{ki}(t)$ denotes a $p \times 1$ vector of covariates for the i^{th} subject at time t with respect to the k^{th} type of failure, expressed as $\mathbf{z}_{ki}(t) = (z_{1ki}(t), z_{2ki}(t), \ldots z_{pki}(t))'$.

Let X_{ki} represent the failure time of the i^{th} subject for the k^{th} type of failure and let C_{ki} represent the censoring time. $\widetilde{X}_{ki}$ are observation values of X_{ki}, where $X_{ki} = \min\{\widetilde{X}_{ki}, C_{ki}\}$. The indicator variable Δ_i is utilized for determining the event as a failure or censoring. Let $\Delta_i = 1$, when $X_{ki} = \widetilde{X}_{ki}$; otherwise $\Delta_i = 0$. Key assumptions for the WLW method are: (1) $X_{ki} \perp C_{ki}$, i.e., the failure and censoring times are independent of each other; (2) (X_i, Δ_i, Z_i) are i.i.d. random vectors, where Z_i represent covariates and i represents event count; and (3) The regression coefficients $\hat{\boldsymbol{\beta}}_i$ follow a normal distribution with mean $\overline{\boldsymbol{\beta}}_i$ denoted

$$(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \ldots, \hat{\beta}_k) \xrightarrow{iid} Normal\ (\overline{\beta}_1, \overline{\beta}_2, \overline{\beta}_3, \ldots, \overline{\beta}_k).$$

5.3.5 Method

Unlike the gap time scale (PWP-GT), the total time scale (PWP-TT and WLW) is invariant to the shape parameter (δ) of the power-law form NHPP, because δ does not influence the likelihood function in the total-time model. The counting process (AG) adopts the total time scale, and thus becomes an estimator invariant to shape parameter. Kelly and Kim [23] and Lipschutz and Snapinn [22] suggested how to use the total time and counting process models. The gap time scale has been considered a better model to capture the dependence structure existing among failure times than has the total time scale. Thus, in any rate of occurrence of failures, utilizing the gap time scale can capture the trend and give a sound estimate of covariate effects, while the total time scale appears to overestimate covariate effects as the event count progresses. This overestimation is termed as a misspecification problem resulting from applying the total time scale and counting process scale. Consequently, the HPP case is chosen for the purpose of illustrating the PWP-TT, AG, and WLW models. The PWP-GT model is implemented for comparison purposes.

There is another limitation of using the total time scale: the maximum number of simulated failure events that is able to afford reliable estimates. As a result, four failure events were generated for each sample unit. In the concept of risk set, there are two types: common (AG) or event-specific baselines (PWP-GT, PWP-TT, and WLW). The common baseline provides a general covariate effect based on the generic analysis from all input data without ranks of failure events, while the event-specific baseline offers the estimated covariate effect in each stratum defined by failure count.

Simulation data with right-censored patterns, where the underlying distribution follows a power-law NHPP is generated by a modified Blanks & Tordon [24] simulation algorithm. Since stationary data are specified, $\delta = 1$ is set to convert a power-law NHPP into an HPP. In order to simulate right-censored recurrent data, two groups of sample units were generated, in which one group contains the sample units with complete data and the other group contains the sample units with right-censored data. In the group of censored units, the right-censored pattern is set random. The ratio (probability) of the sample units that have censored times to total sample units is defined as censored probability (P_c).

A discrete indicator covariate $\mathbf{z}_1$ was used to separate the data into two strata for an arbitrary treatment effect. For consistency with the work of Qureshi et al. [6], simulated data was generated from a power-law NHPP with like parameters, except that four recurring failure events were generated for each sample unit (compared to ten recurring events on Qureshi's work). A proportional intensity function dataset was created using two different values for the scale parameter ($\upsilon_0 = 0.001, \upsilon_1 = 0.01$) corresponding to the two values of the indicator covariate $\mathbf{z}_1$ ($\mathbf{z}_1 = 0, \mathbf{z}_1 = 1$).

There are two experimental factors: experimental units (U) and censoring probability (P_c). The levels for each factor are selected as follows: (1) U = 60, 120, and 180 and (2) P_c= 0, 0.4, 0.8, and 1.0. Note that $P_c = 0$ represents complete data, which provides the comparison of censored and complete data. The selection of the U and P_c levels has taken the following considerations: (1) the parameter settings in the previous relevant works (Proschan [25], Landers and Soroudi [5], Qureshi et al.[6], and Landers et al. [8]) (2) Severe right-censorship may cause the small sample size (U=20) to have insufficient data. The selection of P_c levels takes into account the light, moderate, and severe censoring. The selection of U levels is taken from the parameter settings in the previous research works, and it has also considered the small, median, and large sample sizes.

To implement the four regression methods (AG, PWP-GT, PWP-TT, and WLW), three types of datasets were generated consistent with the theory and methodology of each. For the AG method, the data set is formed from the time interval (T_1, T_2) with respect to the counting process formulation. Thus, the logic rule to form the dataset is: $T_2 > T_1$. As a result, all the censored failure times are removed from the dataset since $T_2 = T_1$ when it is a censored event as stipulated for the AG method. The concept of forming the dataset for the PWP method originates from the probability theory of conditionality. The later failure times after the n^{th} failure count cannot be included into the dataset when the intensity function at the n^{th} failure count is estimated. That is, for each censored unit, the censored times are removed from the dataset except for the first censored event count. Due to the marginal probability theory of the WLW method, the dataset contains full records including all censored events, such that censored units remain in the risk set.

The four semi-parametric methods were implemented using the SAS™ Users Group (SUGI) software code PHREG [26], which performs the semi-parametric Cox regression method with a blocking option to stratify for a covariate, such as failure event count, not satisfying the proportional hazards conditions. PHREG applies the product-limit method to estimate the reliability function within all strata defined by the failure count and for all values of the covariate. PHREG also applies the Cox method to estimate the vector of regression coefficients $\boldsymbol{\beta}$ and the covariance matrix. Appendix IV provides the programming code to perform the four semi-parametric methods. To measure and compare model performance, three robustness metrics were compiled:

- relative signed error (BIAS);

- relative mean absolute deviation (MAD) and
- relative mean squared error (MSE).

5.4 Results

5.4.1. Event-specific baseline models (PWP-GT, PWP-TT, and WLW)

Figure 5.2 presents the 95% C.I. on the PWP-GT, PWP-TT, and WLW estimators in the case of $(U,P_c)=(120,0.4)$, where U denotes sample units and P_c denotes censoring probability. This chart serves as an example that the PWP-GT model has proven the most robust and accurate estimator (at the lowest error) among the three models. The true value of β is derived from the formula $\hat{\beta}=-\frac{1}{\delta}\ln\left(\frac{\upsilon_0}{\upsilon_1}\right)$ = 2.30259. As the failure count proceeds, the PWP-GT model remains within its 95% C.I., while the other two estimators do not lie in their 95%C.I.

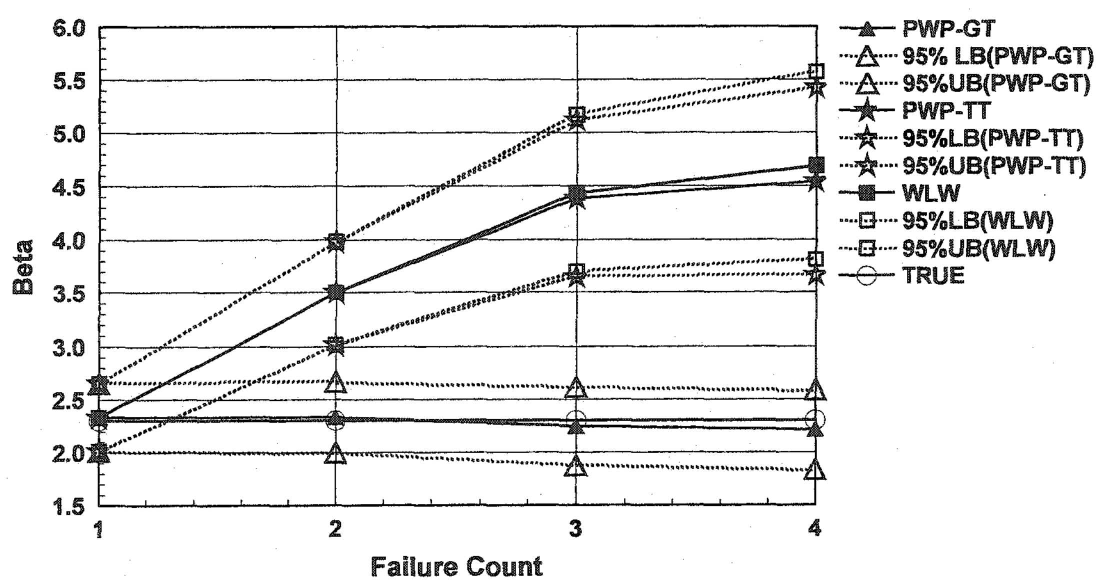


Fig. 5.2 Comparison of the three event-specific estimators of the covariate effect, $(U,P_c)=(120,0.4)$

Table 5.1 summarizes the robustness across the strata defined by failure count. Note that the parameter setting $P_c = 0$ (complete data) is included for comparison with censored data.

Table 5.1 Performance metrics of $\hat{\beta}_1$ in an HPP case (PWP-GT, PWP-TT, and WLW)

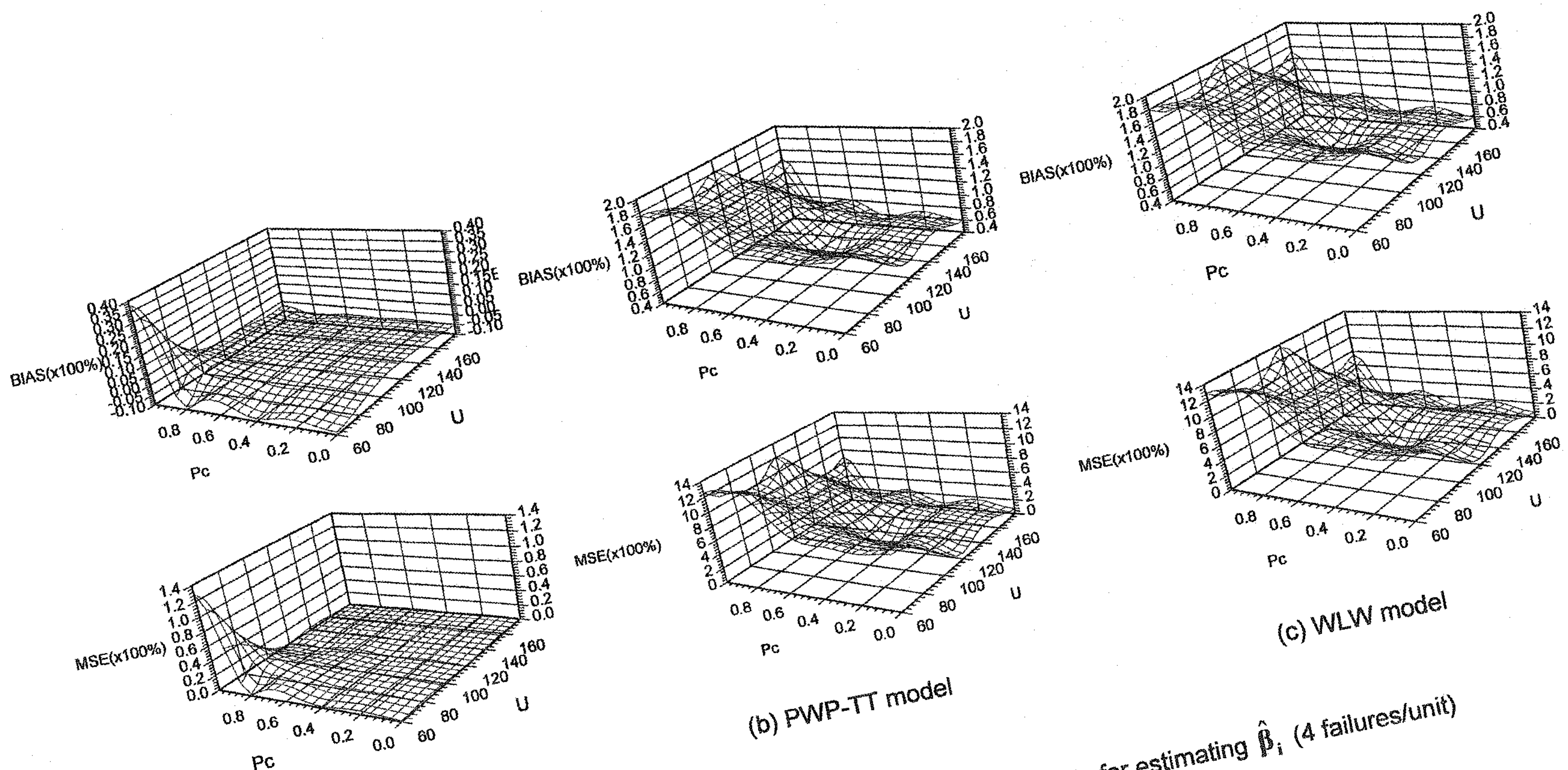
U	P_c	PWP-GT			PWP-TT			WLW		
		BIAS	MAD	MSE	BIAS	MAD	MSE	BIAS	MAD	MSE
60	0	-0.08739	0.08993	0.01591	1.37088	1.40005	7.14217	1.37088	1.40005	7.14217
	0.4	-0.08758	0.09292	0.01806	1.35651	1.38568	7.02349	1.38943	1.41860	7.36003
	0.8	-0.08634	0.10208	0.02009	1.81323	1.84240	13.11053	1.86904	1.89821	13.83614
	1	0.38466	0.59073	1.29475	1.75212	1.78129	12.55797	1.79718	1.82635	12.93005
120	0	-0.02181	0.03149	0.00206	0.63394	0.63394	0.76947	0.63394	0.63394	0.76947
	0.4	-0.00911	0.02261	0.00083	0.60218	0.60218	0.67732	0.62340	0.62340	0.73285
	0.8	-0.01586	0.02253	0.00103	1.49651	1.49651	7.04960	1.57647	1.57647	7.84370
	1	0.00138	0.01607	0.00039	1.83500	1.83500	12.05176	1.95337	1.95337	13.63281
180	0	-0.05047	0.05047	0.00540	0.59103	0.60241	0.72164	0.59103	0.60241	0.72164
	0.4	-0.04359	0.04366	0.00406	0.57569	0.58707	0.67309	0.59926	0.61063	0.73740
	0.8	-0.05770	0.05770	0.00633	0.62249	0.63387	0.81017	0.67436	0.68574	0.97323
	1	-0.02335	0.04742	0.00410	1.44098	1.45236	6.82499	1.50519	1.51657	7.40608

* Refer also to Fig. 5.3 (a)-(c)

Table 5.1 indicates the PWP-GT model has proven the most robust and accurate estimator (at the lowest error) among the three event-specific models. Compared to WLW, the PWP-TT estimator yields similar but slightly better results. Fig. 5.3(a)-(c) contains six charts of 3-D error graphs of performance metrics (BIAS and MSE) illustrated by each method (PWP-GT (Fig. 5.3(a)), PWP-TT (Fig. 5.3(b)), and WLW (Fig. 5.3(c))).

The PWP-GT error chart (Fig. 5.3(a)) presents a low-error region at the range of $120 \le U \le 180$ and $0 \le P_c \le 1$. For the small sample size $U = 60$, the more favorable applications range is $0 \le P_c \le 0.8$, having relative MSE in the range of (1.6%, 2.0%), relative BIAS in the range of (-8.8%, -8.6%), and relative MAD in the range of (9.0%, 10.2%). As the sample size is increased to $U = 120$ the more favorable applications range is widened to $0 \le P_c \le 1.0$, having relative MSE in the range of (0.0%, 0.2%), relative BIAS in the range of (-2.2%, 0.1%), and relative MAD in the range of (1.6%, 3.1%). In the case of $U = 180$, PWP-GT estimates have relative MSE is in the range of (0.4%, 0.6%), relative BIAS is in the range of (-5.8%, -2.3%), and relative MAD is in the range of (4.4%, 5.8%).

The other two estimators (PWP-TT and WLW) present a similar pattern in model performance (Figs. 5.3(b) and (c)). As the sample size increases, the error is on the decrease. As the censoring increases, the error is on the rise. Sample size $U = 60$ does not provide sufficient data, and thus both PWP-TT and WLW yield a poor result. When the sample size is increased to $U = 120$, PWP-TT and WLW have a slightly improved applications range $0 \le P_c \le 0.4$. PWP-TT estimates have relative MSE in the range of (67.7%, 76.9%), relative BIAS in the range of (60.2%, 63.4%), and relative MAD in the range of (60.2%, 63.4%). Likewise, WLW estimates have relative MSE in the range of (73.3%, 76.9%), relative BIAS in the range of (62.3%, 63.4%), and relative MAD in the range of (62.3%, 63.4%). As the sample size is increased to 180, the performance is poor but stable over applications range $0 \le P_c \le 0.8$ on both methods. PWP-TT estimates have relative MSE in the range of (67.3%, 81.0%), relative BIAS in the range of (57.6%, 62.2%), and relative MAD in the range of (58.7%, 63.4%). Likewise, WLW estimates have relative MSE in the range of (72.2%, 97.3%), relative BIAS in the range of (59.1%, 67.4%), and relative MAD in the range of (60.2%, 68.6%).



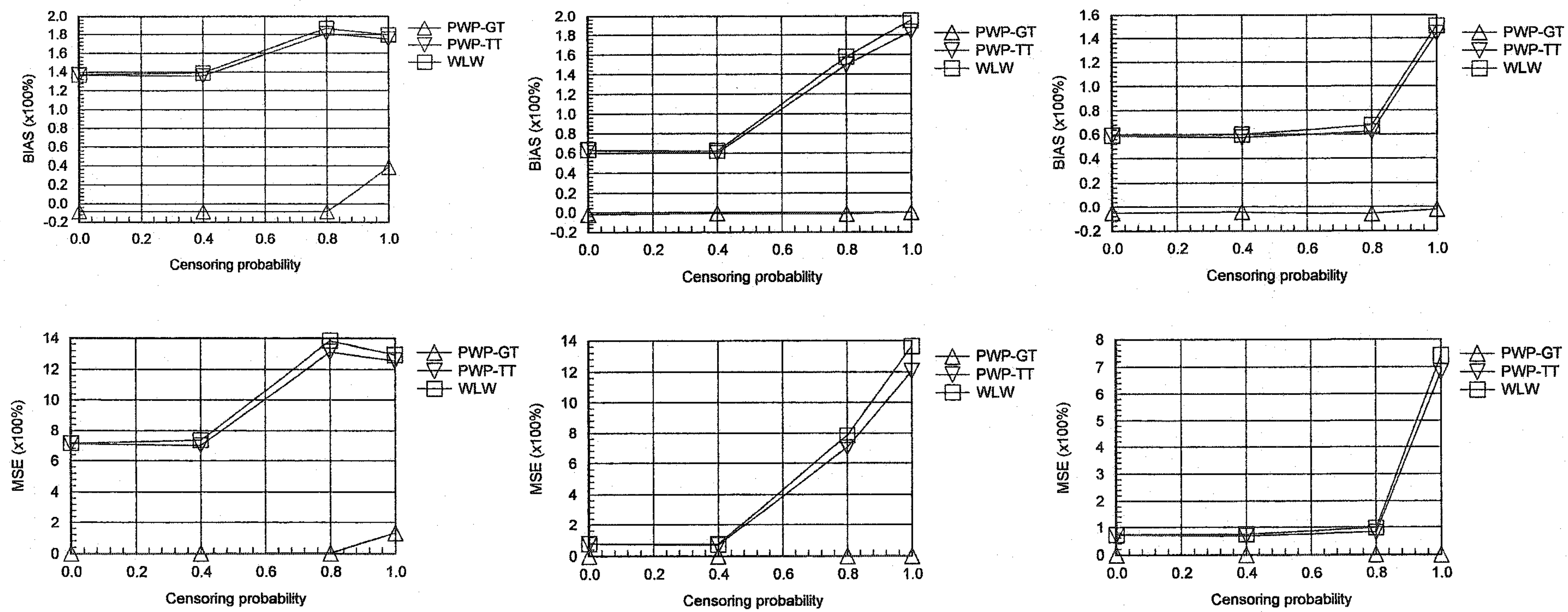
(a) PWP-GT model

(b) PWP-TT model

(c) WLW model

Fig. 5.3(a)-(c) Comparison of PWP-GT, PWP-TT, and WLW model results for estimating $\hat{\beta}_i$ (4 failures/unit)

Fig. 5.4 (a)-(c) examines BIAS and MSE versus P_c for the three event-specific models at each of three sample sizes ($U = 60$, $U = 120$, and $U = 180$). PWP-GT is shown to perform best, since PWP-GT BIAS and MSE are beneath PWP-TT and WLW at any sample size. As the sample size increases, the disparity between the gap time and total time groups is reduced/compensated.



(a) U=60

(b) U=120

(c) U=180

Fig. 5.4(a)-(c) Comparison of three event-specific models at three sample sizes (refer also to Table 5.1)

5.4.2. Common baseline model (AG)

The AG model is used to estimate general covariate effects from all strata of recurrent failure events. The results show that AG performs well for smaller sample size (U = 60) and heavy censoring case ($P_c = 1.0$). In an NHPP case, dependence structure among recurring events hinders the AG model to estimate covariate effects. Thus, an HPP case is chosen to examine the AG model in a right-censoring setting.

Table 5.2 and Fig. 5.5 portray the 95% C.I. of the AG estimates $\hat{\beta}$ in the cases of $60 \le U \le 180$ and $0 \le P_c \le 1$. The true β is 2.30259, and Fig. 5.5 (Table 5.2) indicates that the true β lies within the 95% C.I. of the AG estimates. The favorable applications region lies at $60 \le U \le 180$ and $0 \le P_c \le 1$. As the sample size increases, the variability of the AG estimate becomes smaller, producing narrower confidence intervals.

Table 5.2 95% C.I. of AG estimates for $\hat{\beta}$ ($\delta = 1$, a HPP)

Conditions	AG estimates [a]	95%LB	95%UB
(U, P_c)=(60,0.0)	2.14240(0.17321) [b]	1.80291	2.48188
(U, P_c)=(60,04)	2.09753(0.17830)	1.74808	2.44699
(U, P_c)=(60,0.8)	2.09901(0.19110)	1.72446	2.47356
(U, P_c)=(60,1.0)	2.06188(0.19345)	1.68272	2.44103
(U, P_c)=(120,0.0)	2.22307(0.12272)	1.98254	2.46360
(U, P_c)=(120,0.4)	2.18404(0.12587)	1.93734	2.43073
(U, P_c)=(120,0.8)	2.26325(0.13494)	1.99878	2.52772
(U, P_c)=(120,1.0)	2.22105(0.13600)	1.95450	2.48760
(U, P_c)=(180,0.0)	2.17635(0.09455)	1.99104	2.36166
(U, P_c)=(180,0.4)	2.13810(0.09750)	1.94701	2.32920
(U, P_c)=(180,0.8)	2.15384(0.10222)	1.95350	2.35419
(U, P_c)=(180,1.0)	2.16989(0.10456)	1.96496	2.37482

[a] theoretical values of $\hat{\beta} = -\frac{1}{\delta}\ln\left(\frac{\upsilon_0}{\upsilon_1}\right)$ = 2.30259

[b] Estimated standard errors in parenthesis

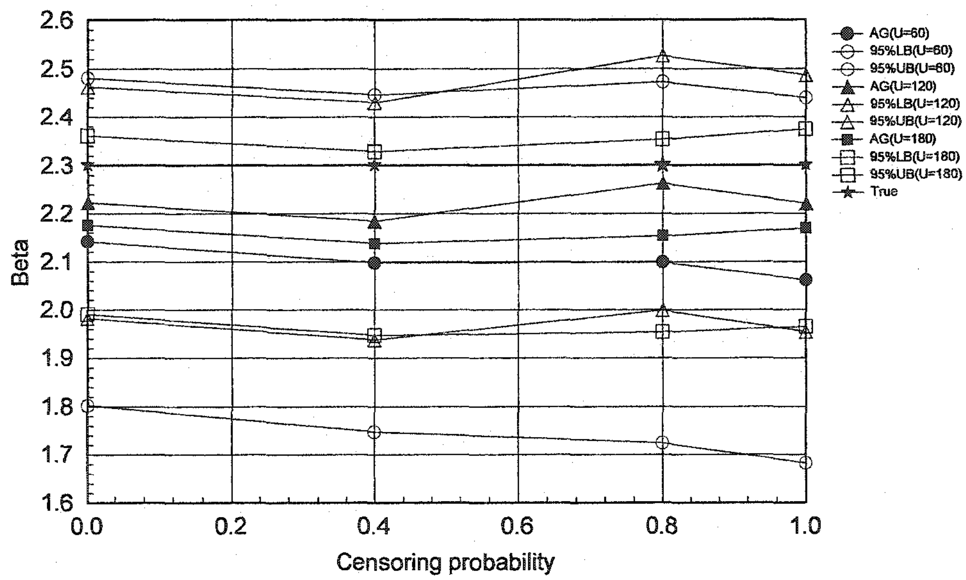


Fig. 5.5 95% C.I. of AG estimates for $\hat{\beta}$

5.5 Conclusions

The research studied the robustness of three event-specific baseline models (PWP-GT, PWP-TT, and WLW) and a common baseline model (AG) to recurring failure events with right-censoring effect from an HPP. The PWP-GT and AG prove to be models of choice, evaluated in terms of the BIAS, MAD, and MSE of covariate regression coefficients over ranges of sample sizes and censoring severity encountered in engineering applications. The favorable engineering applications ranges are recommended.

The research domains of the two factors of interests are: (1) $60 \le U \le 180$ and (2) $0.0 \le P_c \le 1.0$. The parameter setting $P_c = 0$ (complete data) is included for comparison with censored data. The PWP-GT model has proven the most robust and accurate estimator (at the lowest error) among the three event-specific models. Compared to WLW, the PWP-TT estimator yields similar but slightly better results. The PWP-GT presents a low-error region at the range of $120 \le U \le 180$ and $0 \le P_c \le 1$. For the small sample size $U = 60$, the more favorable applications range is $0 \le P_c \le 0.8$. For the other two estimators, when the sample size is increased from $U = 60$ to $U = 120$, PWP-TT and WLW have a slightly improved applications range $0 \le P_c \le 0.4$. As the sample size is increased to 180, the performance is poor but stable over applications range $0 \le P_c \le 0.8$ on both models. The results show that AG performs well for the case of smaller sample size (U = 60) and heavy censoring ($P_c = 1.0$). The favorable applications region of the common baseline AG model lies at $60 \le U \le 180$ and $0 \le P_c \le 1$.

The sample sizes chosen for this engineering research were $60 \le U \le 180$ (30-90 units per class of a two-level covariate). Many of the medical studies reported in the literature contain sample sizes smaller than this range. Small sample sizes are common in medical studies due to the high cost of a clinical trial including requisite examinations required in the medical practices, such as X-Ray scans and blood tests. The numbers of qualified subjects are sometimes small because of cost and/or medical conditions. The AG model is designed to estimate the general covariate effect and can be useful if small sample size is unavoidable. The AG model adopts the stationary counting process model and assumes each occurrence as independent and identically distributed according to an exponential distribution. Thus, the number of observation for a subject can be utilized as the expansion of the sample size. For instance, there are ten subjects available and four observations for each subject are collected. The sample size effect for those ten subjects in using the AG model is equivalent to having forty subjects.

This research has addressed only the case of data from an NHPP with power-law intensity function. The log-linear intensity function is also encountered in the literature and may be important for industry. Future research to examine the right-censoring effect upon recurring events from an NHPP with log-linear intensity function could be beneficial to practitioners.

5.6 References

[1] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika, 1981; 68: 373-9.

[2] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 1982; 10(4): 1100-20.

[3] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, Dec.1989; 84(408): 1065-73.

[4] Cox DR. Regression models and lifetables (with discussion). Journal of the Royal Statistical Society B, 1972; 34:187-202.

[5] Landers TL, Soroudi HE. Robustness of a semi-parametric proportional intensity model. IEEE Transactions on Reliability, 1991; 40(2): 161-4.

[6] Qureshi WM, Landers TL, Edward EG. Robustness evaluation of a semi-parametric proportional intensity model. Reliability Engineering and System Safety, 1994; 44: 103-9.

[7] Vithala S. Robustness of a semi-parametric proportional intensity model for the case of a log-linear Non-homogeneous Poisson Process. A thesis of the Industrial Engineering Department at the University of Arkansas, 1994.

[8] Landers TL, Jiang ST, Peek, JR. Semi-parametric PWP model robustness for log-linear increasing rates of occurrence of failures. Reliability Engineering and System Safety, 2001; 73: 145-53.

[9] Ansell JI, Phillips MJ. Practical problems in the statistical analysis of reliability data (with Discussion). Applied Statistics, 1989;38(2):205-47.

[10] Ansell JI, Phillips MJ. Practical reliability data analysis. Reliability Engineering and System Safety, 1990;28(3):337-56.

[11] Ansell JI, Phillips MJ. Practical aspects of modeling of repairable systems data using proportional hazards models. Reliability Engineering and System Safety, 1997;58(2):165-71.

[12] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Assessing the maintenance in a process using a semi-parametric approach. Quality and Reliability Engineering International, 2001;17:163-7.

[13] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Analysing maintenance data to gain insight into systems performance. Journal of Operational Research Society, 2003;54:343-9.

[14] Ascher H. Regression analysis of repairable systems reliability. In: Skwirzynski JK, editor. Electronic systems effectiveness and life cycle costing. Berlin: Springer, 1983, p.119-33.

[15] Ascher H, Feingold H. Repairable systems reliability--modeling inference, misconceptions and their causes, Lecture notes in

statistics, New York: Marcel Dekker Inc, 1984;7.
[16] Dale C. Application of the proportional hazards model in the reliability field. Reliability Engineering, 1985;10:1-14.
[17] Wightman DW, Bendell A. The practical application of proportional hazards modeling. Reliability Engineering, 1986;15:29-53.
[18] Bendell A, Wightman DW, Walker EV. Applying proportional hazards modeling in reliability. Reliability Engineering and System Safety, 1991;34:35-53.
[19] Bowman ME. An evaluation of statistical models for the analysis of recurrent events data: with application to needlestick injuries among a cohort of female veterinarians. A dissertation of the Preventive Medicine at the Ohio State University, 1996.
[20] Lin DY. Cox's regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine, 1994; 13: 2233-47.
[21] Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Statistics in Medicine, 1997; 16: 833-9.
[22] Lipschutz KH, Snapinn SM. Discussion of paper by Wei and Glidden. Statistics in Medicine, 1997; 16: 841-51.
[23] Kelly PJ, Lim LL-Y. Survival analysis for recurrent data: An application to childhood infectious diseases. Statistics in Medicine, 2000; 19: 13-33.
[24] Blanks HS, Tordon MJ. Laplace and Mann aging trend test effectiveness. Proceeding of the Annual Reliability and Maintainability Symposium, 1987;165-70.
[25] Proschan F. Theoretical explanation of observed decreasing failure rate. Technometrics, Aug. 1963;5(3):375-83.
[26] SUGI Supplementary Library User's Guide, Version 5. SAS Institute, Cary, North Carolina, USA, 1985, p. 437-66.

6. Semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals

6.0 Abstract

The class of semi-parametric proportional intensity (PI) models applies to recurrent failure event modeling for a repairable system with covariates. Certain systems (e.g., aircraft and power plants) experience a substantial period of downtime due to performing maintenance (i.e. major overhaul) following a major failure. This discontinuity in observation time has been a concern in the accuracy of estimating the covariate effect. Hansen and Ascher examined an automobile for intermittent failures, which often lead to a series of unsuccessful repair attempts, and reported that repair times for intermittent failures cannot be assumed negligible and the model must be designed to account for machine downtimes. Therneau and Hamilton proposed a discontinuous risk-free-intervals method for biomedical applications that could also apply to this engineering problem. This paper has examined three semi-parametric PI models (Prentice-Williams-Peterson-gap time (PWP-GT), Andersen-Gill (AG), and Wei-Lin-Weissfeld (WLW)), and has recommended selecting appropriate PI models as a function of the overhaul duration.

The experimental design in this research has incorporated two levels of overhaul duration (short: $R \leq 0.5$ and long: $3.0 \leq R \leq 5.0$, where R is defined as a gap-time-ratio indicating a proportion of the previous $MTTF$ (mean time to failure)) to evaluate these three proposed PI models. The more favorable engineering applications ranges for the overhaul duration based on the sample size (U) and shape parameter (δ) are recommended. The PWP-GT model

proves to perform well for sample sizes 60 (30 per class) or more, moderately decreasing, constant, and moderately increasing rate of occurrence of failures (power-law NHPP shape parameter in the range $0.8 \le \delta \le 1.8$) if the overhaul duration is short ($R \le 0.5$). If it is a long overhaul duration ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2$. In the large sample size 120 (60 per class), the PWP-GT model performs well in the range of $0.5 \le \delta \le 2.0$, if the overhaul duration is short ($R \le 0.5$). If the overhaul duration is long ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2$. As for the other two common baseline intensity model (i.e. AG and WLW), the AG model performs consistently well in the small sample size (20) regardless of the overhaul duration ($R \le 5.0$) in an HPP case. The WLW model performance improves as the overhaul duration increases ($R \ge 5.0$).

Keywords: repairable systems, semi-parametric proportional intensity models, major repairs, overhauls, preventive maintenance, risk-free-intervals

Nomenclature

Acronyms

AG	Andersen and Gill model
C.I.	Confidence interval
DROCOF	Decreasing rate of occurrence of failures
HPP	Homogeneous Poisson Process
IROCOF	Increasing rate of occurrence of failures
i.i.d	Independent and identically distributed
LWA	Lee, Wei, and Amato model
MTTF	Mean time to failure

MAD	Mean absolute deviation
MSE	Mean squared error
NHPP	Non-homogeneous Poisson Process
PH	Proportional hazards
PI	Proportional intensity
PM	Preventive maintenance
PWP	Prentice, Williams, and Peterson model
PWP-GT	Prentice, Williams, and Peterson-gap time model
PWP-TT	Prentice, Williams, and Peterson-total time model
WLW	Wei, Lin, and Weissfeld model

Notation

C_{ki}	Censoring time for the i^{th} subject of the k^{th} type of failures
D	Overhaul duration
F	The event number that represents a major failure
$h(t;z)$	Proportional hazard function
$h_0(t)$	Baseline hazard function
I_0	Number of sample units in class ϕ
I_1	Number of sample units in class 1
i.i.d.	Independent and identically distributed
N	Successive failure count
$N(t)$	Random variable for the number of failures in $(0,t]$; a counting process
n	An integer counting successive failure times; a stratification indicator subscript
R	Gap time ratio
s.d.	Standard deviation
T_1, T_2	The beginning and end of an event; bivariate exponential variables
t_F	Major failure times (clock hour)
T_n	Random variable for cumulative time of occurrence of the n^{th} failure

t_n	Cumulative time of occurrence of the n^{th} failure; a realization of T_n
U	Sample size (number of units)
$\widetilde{X}$	Observation time
Y_{F-1}	The gap time associated with the minor event prior to a major event
$Y_i^{(n)}$	an at-risk indicator in the AG model
$\mathbf{Z}(t)$	Covariate process up to time t
$\mathbf{z}$	$(k \times 1)$ vector of covariates, $\mathbf{z} = (z_1, z_2, ... z_k)'$
$\boldsymbol{\beta}_n$	$(k \times 1)$ vector of stratum-specific regression coefficients
	$\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)$
δ	Shape parameter of a power-law NHPP
Δ	Indicator of a failure or censored time; limit to time zero
λ_0	Baseline value of λ for power-law NHPP
$\lambda_0(t)$	Baseline intensity function
$\lambda_{0n}(t)$	Stratum-specific baseline intensity function
$\lambda(t;\mathbf{z})$	Proportional intensity function
υ	Scale parameter of a power-law NHPP
υ_0	Baseline value of υ, the scale parameter of a power-law NHPP
υ_1	Alternate value of υ, the scale parameter of a power-law NHPP
σ	Standard deviation
^	Denotes an estimator
'	Denotes the transpose of a vector

6.1 Introduction

Failure time data on a repairable system are realizations of a stochastic point process, in which the instantaneous rate of occurrence of failures (ROCOF) is $\lambda(t)$. Prentice, Williams, and Peterson (PWP) [1] proposed a semi-parametric approach to model recurrent failure event data from a repairable system using two methods: PWP-GT (gap time) and PWP-TT (total time). Several researchers have subsequently proposed alternate modeling methods by modifying the risk set (common or event-specific baseline intensity function) and the risk interval (gap time, total time, or counting process). These include the AG (Andersen-Gill) [2] and WLW (Wei-Lin-Weissfeld) [3] models.

Cox proposed the distribution-free (semi-parametric) proportional hazards (PH) model in 1972 [4]. The Cox-based regression models (PWP-GT, PWP-TT, AG, and WLW) have been applied to recurring events in medical studies (biostatistics field), such as recurrent infections of a patient. For engineering applications, Landers and Soroudi [5], Qureshi et al. [6], Vithala [7], and Landers et al. [8] have investigated robustness of the PWP-GT model, where the underlying recurrent failure time data are from a Non-homogeneous Poisson Process (NHPP) with a power-law or a log-linear intensity function. These references also report the engineering applications of the PWP-GT model cited in the literature. Qureshi et al. [6] found that the PWP-GT model performs best for constant and moderately increasing rate of occurrence of failures (IROCOF) and decreasing rate of occurrence of failures (DROCOF) and for larger sample sizes from power-law NHPPs. Vithala [7] considered the case of log-linear increasing rates of occurrence of failures, and concluded the PWP-GT model performs best for moderately increasing rates of occurrence of failures and for larger sample sizes.

Compared to the extensive literature on applications of the proportional intensity (PI) models in the biostatistics field, there have been few reported engineering applications. Abundant federal funding received in biostatistics/ medical research has advanced the PI models to become well developed and widely referenced. PI models for medical applications could also apply to recurring failure/repair data in engineering problems. The AG, PWP-GT, PWP-TT, and WLW models are potentially powerful analytical tools for engineering practitioners as they become better recognized and understood.

Hansen and Ascher [9] examined an automobile for intermittent failures, which often lead to a series of unsuccessful repair attempts, and reported that repair times for intermittent failures cannot be assumed negligible and the model must be designed to account for machine downtimes. Kobbacy and Jeon [10] considered both failure times and machine downtimes in the PI model for preventive maintenance (PM) scheduling in a deteriorating repairable system. Therneau and Hamilton [11] introduced the concept of discontinuous risk-free-intervals that may be applied in reliability engineering as the duration of performing major overhauls. This paper reports progress on continuing work after Landers and Soroudi [4], Qureshi et al. [5], Vithala [6], and Landers et al. [7], and investigates the robustness of the semi-parametric PI models for repairable systems subject to prolonged risk-free-intervals for major repairs (overhauls).

6.2 Semi-parametric Proportional Intensity models

Cox [4] proposed a proportional hazards (PH) formulation to include explanatory variables (covariates) in survival models. PWP proposed an extension of the Cox model to stochastic processes and applied the approach to model recurring infections in aplastic anemia and leukemia patients having received bone-marrow transplants. This application involved several subjects and a small number of events (up to five) for each subject. The paper by PWP did not address the baseline intensity function but rather reported the relative risks for the test and control groups. In reliability and maintainability engineering applications, a number of authors have applied the semi-parametric PI (PH) model, for example, Ansell and Phillips [12], Ansell and Phillips [13], Landers and Soroudi [5], Qureshi et al. [6], Ansell and Phillips [14], Landers et al. [8], Ansell et al. [15], and Ansell et al. [16]. A collection of the PI model applied to different industries includes: marine gas turbine engines (Asher [17]), semiconductor, electrical, and pipeline industries (Ansell and Phillips [14]), U.S. Army main battle tank (Landers et al. [8]), water supply industry (Ansell et al. [12], [16]), etc. Ascher [17] illustrated the use of the PWP model for analysis of reliability for marine gas turbine engines. Ascher and Feingold [18] suggested application of the PWP model in the field of reliability engineering. Dale [19] applied the PWP approach to simulated data for a reliability growth program with design improvements implemented after each of the five stages, resulting in a decreasing rate of occurrence of failures (DROCOF). Wightman and Bendell [20] and Bendell et al. [21] cited the PWP model and advised caution in application for engineering studies.

Qureshi et al. [6] performed a robustness study to determine how well the PWP-GT method performed when applied to data from a failure process that was actually parametric (specifically the NHPP with power-law intensity function). The 2 σ bounds of the PWP-GT estimates can cover the true values with few exceptions. The PWP-GT method performed well, except at small values of shape parameter ($\delta < 0.6$). The PWP-GT method was best for larger sample size and for moderately decreasing, constant, and moderately increasing ROCOF. The validation process for the case of an HPP in Section 2.2.3 (also refer to Table 2.10) indicated that the estimated $MTTF$ (mean time to failure) differences between the PWP-GT model and theoretical values were not statistically significant. As for the PWP-GT estimates of the covariate regression coefficient, the true value of coefficient β lies within the 2 σ confidence bounds on the estimate $\hat{\beta}$ for $1.0 \leq \delta \leq 1.4$. The PWP-GT method tends to underestimate β for a DROCOF (e.g., BIAS= -26% at $\delta = 0.5$) and overestimate β for an IROCOF (e.g., BIAS= 19% at $\delta = 3.0$).

The AG model (Andersen and Gill [2]) and the WLW model (Wei et al. [3]) are widely cited in the literature. Bowman [22] and Lin [23] surveyed and evaluated the AG, PWP-GT, PWP-TT, and WLW methods. Bowman identified the local time model (PWP-GT) as superior and then used it to analyze needle-stick injury data. Wei and Glidden [24] have reviewed the Cox-based methods designed to model recurrent data, and summarized the strengths and weaknesses for each method. In a commentary on the Wei and Glidden paper, Lipschutz and Snapinn [25] stressed two concepts of “event times” and “risk sets” as crucial to choosing the appropriate model. Event elapsed times are related to the total time, gap time, and counting process. The PWP-TT and WLW are modeled by total time, while only PWP-GT is modeled by gap time. The risk interval of the AG model belongs to the counting process class. Intuitively, total (global) times within a subject are highly correlated. The total time model may indicate large treatment effect throughout the entire study, even though the gap time model has indicated little treatment effect beyond a certain recurrence. The counting process concept of the AG method implies each recurrence is not affected by previous events, and does not contribute to future events.

The risk set consists of the subjects at risk for a specified event (e.g., failure). There are three types of risk sets: conditional (e.g., PWP), counting process (e.g., AG), or marginal (e.g., WLW). As a marginal method, the WLW method assumes a subject is at risk regardless of event count until the observation for the subject terminates by censoring. The AG method also provides an index of a general covariate effect, which is expressed by the common baseline hazard (unrestricted risk set). However, a subject in the PWP method has event-specific baseline hazards (restricted risk set), in that the proportional intensity of event k only considers the subjects that have experienced $(k-1)$ events. Lipschutz and Snapinn [25] suggested guidelines as follows in choosing the appropriate models:

- Use total time, common baseline hazard (unrestricted risk set) when the general effect is of interest.
- Use gap time, event-specific baseline hazards (restricted risk set) when the primary concern is how the treatment will affect the recurring events beyond the first occurrence.

Kelly and Lim [26] noted that risk interval can be defined by three formulations {(1) gap time, (2) total time, and (3) counting process} demonstrated in Fig. 6.1(a)-(c). Risk interval determines whether a model is marginal in the total time or conditional in the gap time. The risk interval of any event in total time is not influenced by any previous events, but measures time from entry into the experiment (beginning of observation). However, the risk interval of the gap time begins from the end of last event (Kelly and Lim [26]). Counting processes use the total time scale and share the same elapsed time as does the gap time model. However, the risk interval starts from the previous event instead of the entry time. Based on the common or event-specific baseline intensities, the risk set is labeled as either unrestricted or restricted. Kelly and Lim [26] defined three possible risk sets {(1) unrestricted, (2) restricted, and (3) semi-restricted} in deciding which sample units are at risk of contributing to event k. Kelly and Lim [26] employed the concepts of the risk interval and risk set and categorized the AG, PWP-gap time (PWP-GP), PWP-total time (PWP-TT), WLW, LWA(Lee-Wei-Amato), and other methods.

Hansen and Ascher [9] examined an automobile for intermittent failures, which often lead to a series of unsuccessful repair attempts, and reported that repair times for intermittent failures cannot be assumed negligible and the model must be designed to account for machine downtimes. Kobbacy and Jeon [10] considered both failure times and machine downtimes in the PI model for preventive maintenance (PM) scheduling in a deteriorating repairable system.

Therneau and Hamilton [11] introduced the concept of discontinuous risk-free-intervals. A study of rhDNase in patients with cystic fibrosis involved a seven-day discontinuous risk-free-interval, initiated by intravenous (IV) administration of antibiotics. The concept of risk-free-intervals may be applied in reliability engineering as the duration of performing major overhauls. The discontinuous-risk-free-intervals modeling relaxes the assumption of zero time to perform a major overhaul, and thus better describes the typical field life cycle. For instance, suppose three failures have taken place at days 25, 60, and 90, where two days of performing a major overhaul are required after the second failure. The data records, expressed as $(n, t_1, t_2, status)$ for the three failure times in the Cox-based models, can be written as $(1,0,25,1)$, $(2,25,60,1)$, and $(3,62,90,1)$, where $(n, t_1, t_2, status)$ denotes (failure count, start time, stop time, (0,1) indicator variable for censor (0) event or failure (1) event). The value $t_2 = 90$ of the third failure with a major overhaul records global time to failure with the third failure coinciding with the beginning of a risk-free-interval. However, consideration of major overhaul of duration (D) requires a change from interval (t_1, t_2) to interval $(t_1 + D, t_2)$. In the aircraft industry, D could be as long as one year after flying for 3000 hours for a major overhaul or as short as a few hours for a minor repair. This robustness study examines how the magnitude of D affects the PI methods, as measured by the regression estimates ($\hat{\beta}_i$).

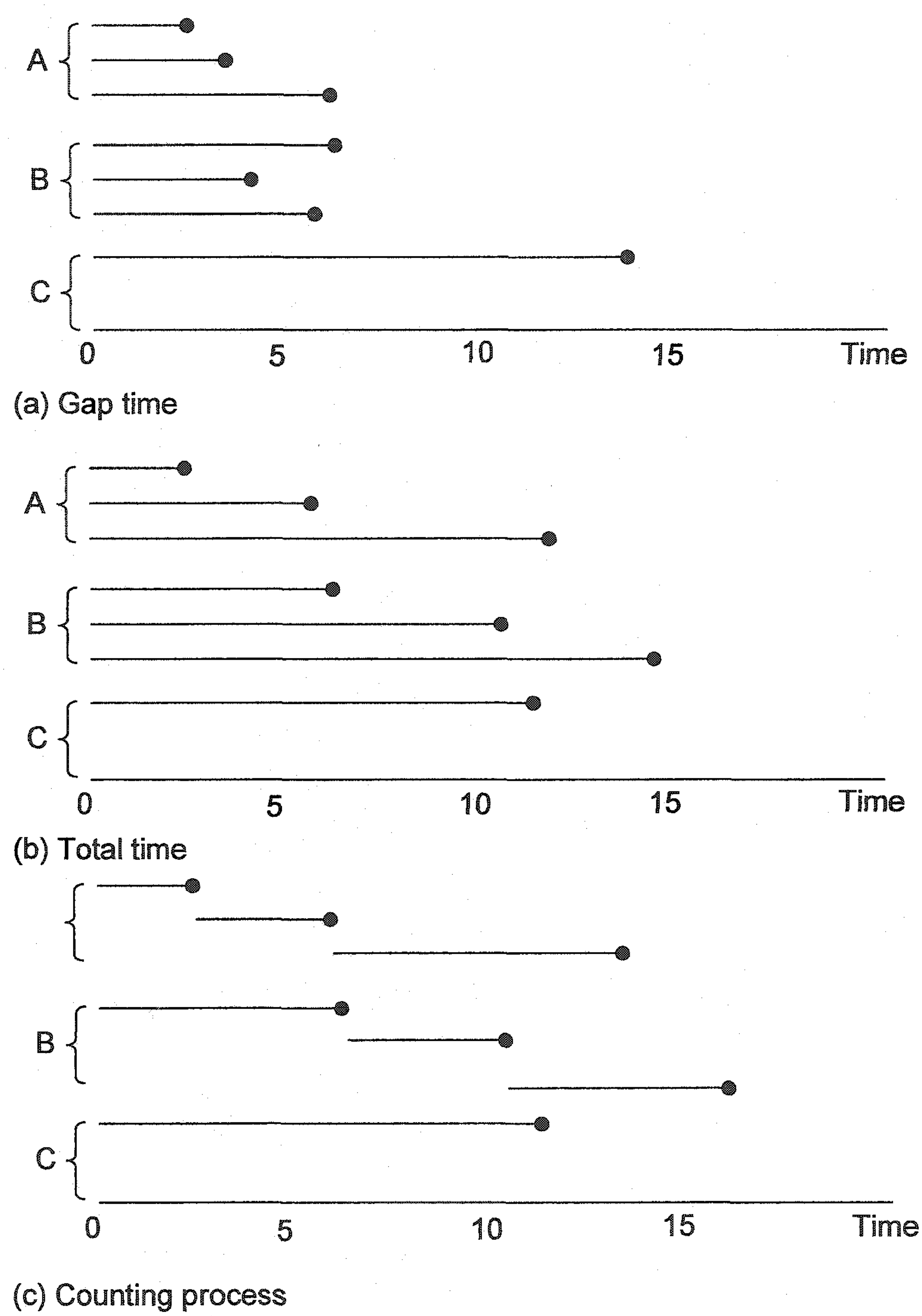


Fig. 6.1(a)-(c) Risk interval formulations (Kelly and Lim [26])

6.3 Models and methods

Sections 6.3.1-6.3.4 review the semi-parametric Cox regression model for single event and the related regression models for recurrent events. Section 6.3.5 reviews an NHPP with power-law intensity function. Section 6.3.6 describes the method used to assess the robustness of the semi-parametric PI methods for the case of complete data from a true but unknown power-law NHPP.

6.3.1 Cox regression model

For the case of a time-to-failure random variable, Cox [4] proposed a PH regression model of the form:

$$h(t;\mathbf{z}) = h_0(\mathrm{t})\exp(\boldsymbol{\beta}'\mathbf{z}), \tag{1}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The PH model is composed of two parts: baseline hazards function $h_0(t)$ and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect. The Cox model can be used to describe the semi-parametric distribution of time-to-failure for non-repairable items with covariates. Under proportional hazards, the ratio of the hazard functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline hazard function. When the baseline hazard function is fully specified (e.g.,Weibull) the analytical procedure is termed a parametric method. Alternatively, $h_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

6.3.2 Semi-parametric PWP model

The PWP model [1] is a generalization of the semi-parametric Cox proportional hazard function to a proportional intensity function $\lambda(t;\mathbf{z})$ for the case of repeated failure events. Under proportional intensities, the ratio of the intensity functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline intensity function. When the baseline intensity function is fully specified (e.g., power-law or log-linear) the analytical procedure is termed a parametric method. Alternatively, the baseline intensity function can be left arbitrary in which case the procedure is termed semi-parametric.

Given the counting and covariate processes at time t, the general semi-parametric intensity function was defined by PWP as follows:

$$\lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lim \ \Pr\{t \leq T_{n(t)+1} < t + \Delta t \mid N(t), \mathbf{Z}(t)\} / \Delta , \tag{2}$$

where $N(t)$ represents a random variable for the number of failures in $(0,t]$, $\mathbf{Z}(t)$ denotes the covariate process up to time t, and Δ limits the time span to zero. Among the semi-parametric regression models specified by PWP were the following:

$$PWP-GT : \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t - t_{n-1})\exp[\boldsymbol{\beta}_n' \mathbf{z}(t)] \tag{3}$$

$$PWP-TT : \lambda\{t \mid N(t), \mathbf{Z}(t)\} = \lambda_{0n}(t)\exp[\boldsymbol{\beta}_n' \mathbf{z}(t)]. \tag{4}$$

In the PWP-GT model of Eq. (3), the time metric is the interval between times of failure t_{n-1} and t_n, defined as gap time. The PWP model stratifies a failure data set based on the failure event count. When a unit is placed into operation it has experienced no failures and so resides in stratum 1 ($n=1$), and when the first failure occurs the unit moves to the second stratum ($n=2$). In general, the unit moves to stratum n immediately following the $(n-1)^{st}$ failure and remains there until the n^{th} failure.

Unlike the gap time model, the limitation of the event-specific total time model restricts the number of recurring events. Ten recurring failure events generated from a power-law NHPP in this study have shown a highly correlated relationship. Thus, the PWP-TT model is modified to a special case of Eq. (4), where the baseline intensity function is set to a common baseline intensity function denoted as $\lambda_{on}(t)=\lambda_0(t)$.

6.3.3 Semi-parametric AG model

Andersen and Gill [2] developed the AG method as an extension of the Cox proportional hazards model, to accommodate recurring events in a counting process. The AG method explains general covariate effects (common baseline intensity function in the concept of risk set), since each event count re-starts the failure process, and thus does not feature event-stratifying effects. The risk interval of an AG model follows a counting process associated with recurring events, where recurrences $(N_i^{(n)},Y_i^{(n)},Z_i^{(n)})$ are independent and identically distributed (i.i.d.) replicates of (N,Y,Z), and the probability of the occurrence of two events at a given time is zero. Symbols: (N,Y,Z) represent the successive failure count, an at-risk indicator, and covariates. Thus, the risk set of the $(n-1)^{st}$ event is identical to the risk set of the $(n)^{th}$ event. The AG model is defined as

$$\lambda_i^{(n)}(t) = Y_i^{(n)}(\mathrm{t})\lambda_0(\mathrm{t})\exp\{\boldsymbol{\beta} \times \mathbf{z}_i^{(n)}(t)\}, \tag{5}$$

where $Y_i^{(n)}$ is an at-risk indicator and $Y_i^{(n)} = 1$ unless the subject is withdrawn from the study.

6.3.4 Semi-parametric WLW model

WLW proposed a marginal method, expanded from the conditional PWP method, in dealing with recurrent failure data. Compared to the PWP method, the WLW method has greater or equal risk set, depending on the sample size associated with the failure count. The PWP method estimates the intensity function by considering the subjects having a complete history of previous recurring events, while the WLW method additionally considers the subjects that have been withdrawn from observation. The subjects that have been censored are still in the risk set; thus, contributing influence on events that are followed after the censoring time. The risk set of each subject using the WLW method remains the same regardless of complete data or censoring events since a subject is still at risk when the subject has been withdrawn from the experiment.

Wei et al. [3] in a bladder cancer study examined treatment effects by using the PWP and WLW models about placebo and thiotepa therapies for tumor patients. This bladder cancer example collects four recurrence times of tumors $T_1 \sim T_4$ corresponding to four marginal proportional hazards models. Rather than fitting each T_i one model at a time, WLW fits four marginal models in one analysis, simultaneously. This example has two response variables {failure time and censoring status}, three covariates {treatment, tumour number, tumour size}, and four recurrent events over time.

For the k^{th} failure type and the i^{th} failure event count, the hazard function $\lambda_{ki}(t)$ in WLW is assumed to take the form:

$$\lambda_{ki}(t) = \lambda_{k0}(\mathrm{t})\exp\{\boldsymbol{\beta}'_k \times \mathbf{z}_{ki}(t)\}, \mathrm{t} \geq 0, \tag{6}$$

where $\lambda_{k0}(t)$ is an unspecified baseline hazard function and $\boldsymbol{\beta}'_k = (\beta_{1k},....,\beta_{pk})'$ is a vector of failure-specific regression parameters. $\mathbf{z}_{ki}(t)$ denotes a $p \times 1$ vector of covariates for the i^{th} subject at time t with respect to the k^{th} type of failure, expressed as $\mathbf{z}_{ki}(t) = (z_{1ki}(t), z_{2ki}(t),...z_{pki}(t))'$. Let X_{ki} represent the failure time of the i^{th} subject for the k^{th} type of failure and let C_{ki} represent the censoring time. $\widetilde{X}_{ki}$ are observation values of X_{ki}, where $X_{ki} = \min\{\widetilde{X}_{ki}, C_{ki}\}$. The indicator variable Δ_i is utilized for determining the event as a failure or censoring. Let $\Delta_i = 1$, when $X_{ki} = \widetilde{X}_{ki}$; otherwise $\Delta_i = 0$. Key assumptions for the WLW method are: (1) $X_{ki} \perp C_{ki}$, i.e., the failure and censoring times are independent of each other; (2) (X_i, Δ_i, Z_i) are i.i.d. random vectors, where Z_i represent covariates and i represents event count; and (3) The regression coefficients $\hat{\boldsymbol{\beta}}_i$ follow a normal distribution with mean $\overline{\boldsymbol{\beta}}_i$ denoted $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3,..., \hat{\beta}_k) \xrightarrow{iid} Normal$ $(\overline{\beta}_1, \overline{\beta}_2, \overline{\beta}_3,..., \overline{\beta}_k)$.

Unlike the gap time model, the limitation of the event-specific total time model restricts the number of recurring events. Ten recurring failure events generated from a power-law NHPP in this study have shown a highly correlated relationship.

Thus, the baseline intensity function of Eq. (6) is set to a common baseline intensity function denoted as $\lambda_{k0}(t) = \lambda_0(t)$. This simplified model is then termed as Lee-Wei-Amato (LWA) model designated to measure general covariate effects. In addition, the WLW model in this study is equivalent to a PWP-TT model when the failure count N for each sample unit is equal.

6.3.5 Power-law intensity function

For a power-law NHPP, the baseline intensity function is

$$\lambda_0(t) = v_0 \delta \times t^{\delta-1}, \tag{7}$$

where δ is the shape parameter and υ is the scale parameter of the power-law form. If we define $\upsilon_0 = \exp(\beta_0 \times z_0)$ and $z_0 = 1$, then the power-law PI model becomes

$$\lambda(t;\mathbf{z}) = \delta \times t^{\delta-1} \exp(\boldsymbol{\beta}'\mathbf{z}), \tag{8}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The power-law intensity function is composed of two parts: baseline intensity function that follows a power-law form and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect.

This process could model the reliability of a repairable system with rapid deterioration, since the failure intensity is increasing at an exponential rate with time. The analogous case for maintainability is a rapid learning process. The intensity function $\lambda(t)$ is strictly decreasing for δ <1, constant for δ = 1, and strictly increasing for δ >1. Thus, we have a DROCOF for δ < 1, an HPP for δ = 1, and an IROCOF for δ >1.

6.3.6 Method

In this study, simulated recurring data are generated from a modified Blanks & Tordon [27] algorithm. To determine the time to perform major overhauls in the counting process, a uniformly distributed $U(0,1)$ random variate is introduced to select the event number F, where the major overhaul is performed. The major overhaul is arranged after the F^{th} event, and we assume that a period D is required to perform a major overhaul. Thus, the next event time, which belongs to the $(F+1)^{st}$ event, occurs depending on the F^{th} event time plus the major overhaul duration.

The duration to perform a major overhaul is inserted into the interval (t_F, t_{F+1}), which makes the interval of risk become $(t_F + D, t'_{F+1})$, where new event time t'_{F+1} is determined by $t_F + D$ in the Blanks & Tordon formula. Consequently, the gap time and (t_F, t_{F+1}) have been altered compared to the recurrent data without the interruption of a major overhaul interval. However, the discontinuous risk interval concept in Therneau and Hamilton [11] is different in terms of (t_F, t_{F+1}), while the gap time remains unchanged. "For instance, in a study of patients with hip fracture, a subject who fractured at day 100, followed by a 15 day hospital stay and then 300 more days of uneventful follow-up would be represented as two at-risk intervals: $(0,100], (115,415]$" (Therneau and Hamilton [11]). The gap times of 100 days and 300 day remain the same, while the risk interval has been shifted forward from $\{(0,100], (100,400]\}$ to $\{(0,100], (115,415]\}$.

The magnitude of D is determined based on the previous gap time Y_{F-1}, where F is a random variate indicating the F^{th} event is a major failure event; otherwise, the event is a minor failure. The relationship between D and Y_{F-1} is:

$D = R \times Y_{F-1} \longrightarrow R = D / Y_{F-1}$,

where

R is the gap time ratio that controls the magnitude of D,
Y_{F-1} represents the gap time associated with the minor event prior to a major event,
F is the event number that represents a major failure.

The concept of utilizing the gap time ratio R in determining the major overhaul duration strengthens the model, since there are three types of power-law intensity functions (IROCOF, constant ROCOF, and DROCOF). The recurrent failure interval can vary from one time unit to a large value depending on the shape parameter.

The parameter settings are as follows when a discontinuous-risk-intervals model is associated with the repair time: (1) Scale parameters in CLASS0 (υ_0) and CLASS1 (υ_1) are set to 0.001 and 0.01; (2) Number of failures $N = 10$; (3) F^{th} event represents a major failure, followed by a major overhaul; and (4) Seed numbers for three replicates are 539, 255, and 59. The magnitude of D is examined as the primary factor that affects the performance of the semi-parametric PI models.

In the experimental design regarding the discontinuous risk-free-intervals model, there are three experimental factors: (1) number of the experimental units (U), (2) shape parameter (δ), and (3) gap time ratio (R) that controls the major overhaul duration (D). I_0 and I_1 represent the number of units in each class.

The levels of experimental factors are as follows. (1) U: 20, 60, and 120 (2) δ: 0.5, 0.8, 1.0, 1.2, 1.5, 1.8, and 2.0 (3) R: 0.001, 0.1, 0.3, 0.5, 3.0, and 5.0. The selection of the U, δ, and R levels has taken the following considerations: (1) the parameter settings in the previous relevant works (Proschan [28], Landers and Soroudi [5], Qureshi et al. [6], and Landers et al. [8]) (2) Gap time ratio (R) controls the magnitude of the duration performing an overhaul that is designated to cover two levels of overhaul duration (short: $R \le 0.5$ and long: $3.0 \le R \le 5.0$; R defined as a gap-time-ratio indicating a proportion of the previous $MTTF$ (mean time to failure)). The selection of U and δ levels is taken from the parameter settings in the previous research works, and it has also considered the small, median, and large sample sizes for U.

To implement the three Cox-Based regression methods (AG, PWP-GT, and WLW), requires formulation of three types of datasets (i.e. three formats for the same set of failure events, according to the theory underlying each methodology). First, for the AG method, the data set is formed from the time interval (T_1, T_2) with respect to the following counting process formulation:

$$\lim_{h \to 0} \frac{1}{h} p[N(t+h) - N(t) = 1 \mid T > t] = \lambda(t), \tag{9}$$

where

$\lambda(t)$: proportional intensity function of failure process,
$N(t)$: random variable for number of failures in $(0, t]$.

Eq. (9) defines the instantaneous failure rate between t and $t + \Delta$ under the condition that this individual has survived after time t. Thus, the logic rule to form the dataset is: $T_2 > T_1$. As a result, all the censored failure times are removed from the dataset since $T_2 = T_1$ when it is a censored event as stipulated for the AG method. The concept of forming the dataset for the PWP method originates from the probability theory of conditionality. The later failure times after the n^{th} failure count cannot be included into the dataset when the intensity function at the n^{th} failure count is estimated. That is, for each censored unit, the censored times are removed from the dataset except for the first censored event count. Due to the marginal probability theory of the WLW method, the dataset contains full records including all censored events, such that censored units remain in the risk set.

The three Cox-based semi-parametric methods were implemented using the SAS™ Users Group (SUGI) software code PHREG [29], which performs the semi-parametric Cox regression method with a blocking option to stratify for a covariate, such as failure event count, not satisfying the proportional hazards conditions. PHREG applies the product-limit method to estimate the reliability function within all strata defined by the failure count and for all values of the covariate. PHREG also applies the Cox method to estimate the vector of regression coefficients $\boldsymbol{\beta}$ and the covariance matrix. In the special case of an HPP, two models (AG and WLW) were compared, in terms of 95% confidence intervals. Appendix V provides the programming code to perform the three semi-parametric methods. To measure and compare the performances, three robustness metrics were compiled:

- relative signed error (BIAS);
- relative mean absolute deviation (MAD) and

- relative mean squared error (MSE).

6.4 Results

6.4.1 PWP-GT model results

This section examines the PWP-GT model robustness in estimating the covariate effect denoted as $\hat{\boldsymbol{\beta}}_1$. The experimental values of R considered in this study may be grouped into short maintenance interval ($0.001 \le R \le 0.5$) and long maintenance interval ($3.0 \le R \le 5.0$) categories, based on the PWP-GT performance. Table 6.1 summarizes the performance information for sample sizes 20, 60, and 120. Two factors (Gap Time Ratio (R) and the shape parameter (δ)) are portrayed in 3-D error charts (Fig. 6.2(a)-(b)), for sample size equal to 120. The larger errors occur at small sample sizes ($U = 20$) and DROCOF ($\delta = 0.5$). Fig. 6.2(b) indicates that the bias tends to be positive in the region of low gap time ratio and high shape parameter ($\delta \ge 1.2$). In addition, for $U = 60$ and 120, bias tends to be negative in the long maintenance interval ($3.0 \le R \le 5.0$) category. For $U = 60$ and 120, MSE (MAD) presents a convex function of shape parameter with a minimum point at $\delta = 1.0$. The unsigned performance metric BIAS tends to be negative when $R \ge 3.0$, $U > 20$, and $0.5 \le \delta \le 2.0$, which indicates that the PWP-GT estimator underestimates the covariate effect in the long maintenance interval category and $U > 20$.

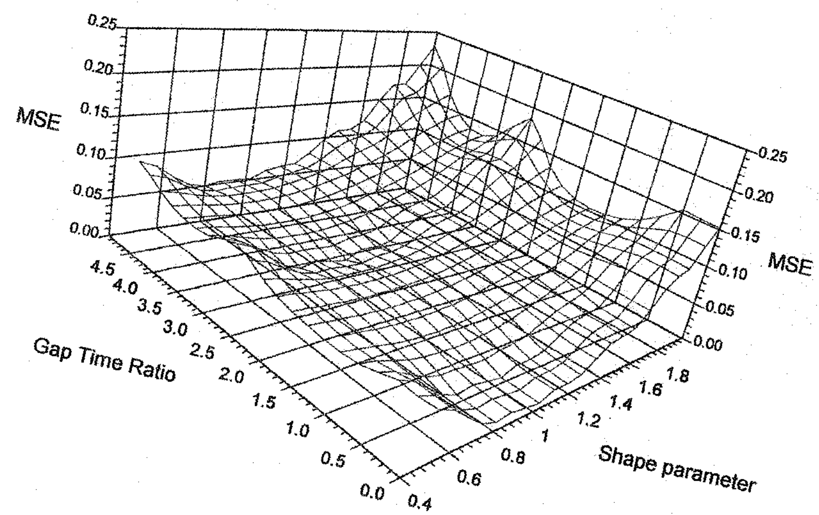
In the small sample size ($U = 20$), the maximum error occurs at $\delta = 0.5$. The more favorable applications range for shape parameter is $1.5 \le \delta \le 2.0$, where PWP-GT estimates have relative MSE in the range of (3.6%, 19.5%), relative BIAS in the range of (-9.2%, 24.3%), and relative MAD in the range of (18.0%, 34.1%) across all values of gap time ratio, $0.001 \le R \le 5.0$. In the case of $U = 60$

and short interval, the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.8$, producing relative MSE in the range of (1.0%, 12.2%), relative BIAS in the range of (-2.5%, 16.3%), and relative MAD in the range of (7.4%, 21.3%). In longer interval ($3.0 \le R \le 5.0$), the favorable applications range contracts to $0.8 \le \delta \le 1.2$, having relative MSE in the range of (2.5%, 5.4%), relative BIAS in the range of (-16.4%, -7.7%), and relative MAD in the range of (12.4%, 20.0%).

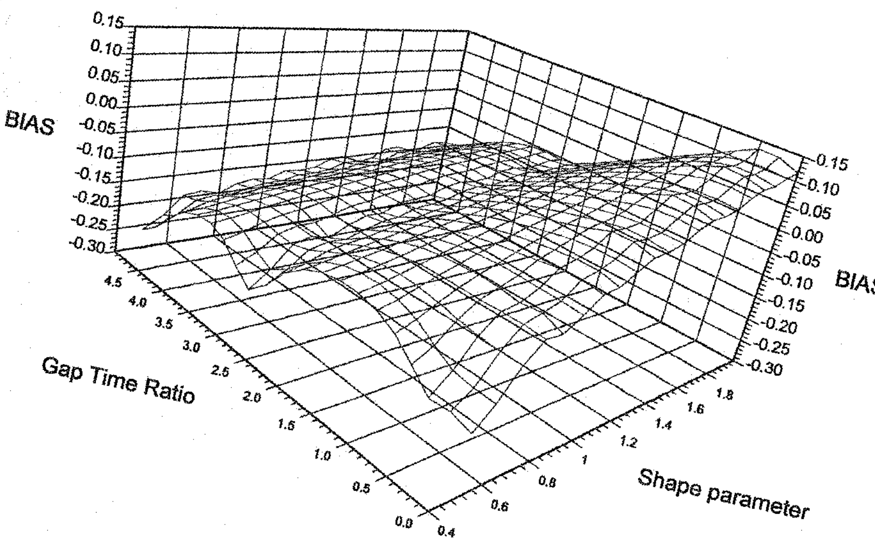
Large sample size ($U = 120$) yields significant improvement upon the PWP-GT model across all shape parameters in $0.5 \le \delta \le 2.0$ and large gap time ratio ($3.0 \le R \le 5.0$). For $U = 120$ and short interval ($0.001 \le R \le 0.5$), PWP-GT estimates have relative MSE in the range of (0.9%, 15.5%), relative BIAS in the range of (-18.9%, 13.5%), and relative MAD in the range of (7.6%, 19.9%) across $0.5 \le \delta \le 2.0$. At $U = 120$ and longer interval, the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2$, having relative MSE in the range of (2.4%, 6.4%), relative BIAS in the range of (-20.0%, -9.6%), and relative MAD in the range of (13.1%, 22.6%).

Table 6.1 Summary of PWP-GT model results for estimating $\hat{\boldsymbol{\beta}}_i$ (10 failures/unit)

N=10 failure events/unit, $\upsilon_0=0.001, \upsilon_1=0.01$																	
U	δ	R	BIAS	MAD	MSE	U	δ	R	BIAS	MAD	MSE	U	δ	R	BIAS	MAD	MSE
20	0.5	0.001	1.05529	1.16554	1.85949	60	0.5	0.001	0.20635	0.43678	0.32301	120	0.5	0.001	-0.18869	0.18869	0.05788
20	0.5	0.1	1.30546	1.41571	2.84997	60	0.5	0.1	0.21266	0.43848	0.33449	120	0.5	0.1	-0.18207	0.18207	0.05648
20	0.5	0.3	1.30546	1.41571	2.84997	60	0.5	0.3	0.21629	0.44508	0.34457	120	0.5	0.3	-0.18075	0.18075	0.05590
20	0.5	0.5	1.40641	1.51666	3.15124	60	0.5	0.5	0.12094	0.35131	0.23542	120	0.5	0.5	-0.18314	0.18314	0.05717
20	0.5	3.0	1.25770	1.36795	2.44395	60	0.5	3.0	-0.04106	0.32454	0.17335	120	0.5	3.0	-0.22196	0.22365	0.07620
20	0.5	5.0	1.13721	1.24746	2.09833	60	0.5	5.0	-0.06841	0.33925	0.17441	120	0.5	5.0	-0.25972	0.25972	0.09196
20	0.8	0.001	0.53642	0.70429	1.65169	60	0.8	0.001	-0.02523	0.09390	0.01273	120	0.8	0.001	-0.07059	0.08352	0.01028
20	0.8	0.1	0.73095	0.87817	2.11346	60	0.8	0.1	-0.00511	0.11086	0.01943	120	0.8	0.1	-0.05928	0.07736	0.00927
20	0.8	0.3	0.73024	0.87888	2.11358	60	0.8	0.3	-0.00035	0.11643	0.02004	120	0.8	0.3	-0.05589	0.07640	0.00951
20	0.8	0.5	0.72966	0.87830	2.11146	60	0.8	0.5	-0.00728	0.11616	0.01967	120	0.8	0.5	-0.06111	0.07968	0.01017
20	0.8	3.0	0.51381	0.71972	1.08973	60	0.8	3.0	-0.11318	0.12380	0.02537	120	0.8	3.0	-0.14593	0.14593	0.02817
20	0.8	5.0	0.30858	0.61291	0.82874	60	0.8	5.0	-0.16369	0.16369	0.04071	120	0.8	5.0	-0.19979	0.19979	0.05010
20	1.0	0.001	0.36482	0.50068	0.89836	60	1.0	0.001	0.02922	0.07409	0.01040	120	1.0	0.001	-0.02056	0.05770	0.00541
20	1.0	0.1	0.38641	0.49692	0.89813	60	1.0	0.1	0.03829	0.08355	0.01157	120	1.0	0.1	-0.01052	0.05933	0.00546
20	1.0	0.3	0.40193	0.51243	0.90876	60	1.0	0.3	0.04065	0.08796	0.01219	120	1.0	0.3	-0.00418	0.06040	0.00576
20	1.0	0.5	0.58564	0.70513	1.50187	60	1.0	0.5	0.04110	0.08851	0.01222	120	1.0	0.5	-0.00879	0.05870	0.00567
20	1.0	3.0	0.30274	0.54852	1.01084	60	1.0	3.0	-0.08752	0.12458	0.02573	120	1.0	3.0	-0.11605	0.13064	0.02415
20	1.0	5.0	0.22762	0.53811	0.85849	60	1.0	5.0	-0.16122	0.18018	0.04865	120	1.0	5.0	-0.18350	0.19461	0.05112
20	1.2	0.001	0.52718	0.62346	1.48737	60	1.2	0.001	0.08294	0.10237	0.02007	120	1.2	0.001	0.01729	0.08407	0.01378
20	1.2	0.1	0.54539	0.63709	1.49004	60	1.2	0.1	0.08721	0.10823	0.02189	120	1.2	0.1	0.02854	0.08712	0.01449
20	1.2	0.3	0.56270	0.64774	1.49094	60	1.2	0.3	0.08639	0.11477	0.02275	120	1.2	0.3	0.03602	0.08946	0.01523
20	1.2	0.5	0.56659	0.65885	1.50033	60	1.2	0.5	0.07704	0.10878	0.02105	120	1.2	0.5	0.03184	0.08891	0.01492
20	1.2	3.0	0.42775	0.60969	1.43261	60	1.2	3.0	-0.07749	0.15887	0.03877	120	1.2	3.0	-0.09647	0.15264	0.03401
20	1.2	5.0	0.11501	0.39631	0.66547	60	1.2	5.0	-0.14027	0.19979	0.05414	120	1.2	5.0	-0.17221	0.22554	0.06364
20	1.5	0.001	0.14351	0.20067	0.07007	60	1.5	0.001	0.12099	0.13857	0.05603	120	1.5	0.001	0.05896	0.11890	0.04552
20	1.5	0.1	0.15381	0.21500	0.07804	60	1.5	0.1	0.13055	0.15026	0.05831	120	1.5	0.1	0.06858	0.12353	0.04674
20	1.5	0.3	0.17885	0.23507	0.08925	60	1.5	0.3	0.13352	0.16294	0.06181	120	1.5	0.3	0.07718	0.12973	0.04794
20	1.5	0.5	0.16995	0.23333	0.08713	60	1.5	0.5	0.13113	0.17285	0.06438	120	1.5	0.5	0.07687	0.13007	0.04788
20	1.5	3.0	0.00965	0.17953	0.04545	60	1.5	3.0	-0.06821	0.19745	0.07920	120	1.5	3.0	-0.07678	0.19344	0.06959
20	1.5	5.0	-0.05651	0.20130	0.06404	60	1.5	5.0	-0.14634	0.27074	0.11214	120	1.5	5.0	-0.16706	0.28371	0.10922
20	1.8	0.001	0.15371	0.21054	0.08734	60	1.8	0.001	0.14288	0.18523	0.11359	120	1.8	0.001	0.09526	0.16026	0.10167
20	1.8	0.1	0.16737	0.22714	0.09556	60	1.8	0.1	0.15321	0.20032	0.11720	120	1.8	0.1	0.10241	0.16431	0.10291
20	1.8	0.3	0.20559	0.26150	0.11308	60	1.8	0.3	0.16319	0.21332	0.12219	120	1.8	0.3	0.11266	0.16858	0.10386
20	1.8	0.5	0.20608	0.26418	0.11203	60	1.8	0.5	0.15190	0.21324	0.12158	120	1.8	0.5	0.11207	0.16920	0.10416
20	1.8	3.0	0.07352	0.19724	0.03610	60	1.8	3.0	-0.04887	0.24272	0.14477	120	1.8	3.0	-0.06343	0.24342	0.13061
20	1.8	5.0	-0.09228	0.27500	0.13071	60	1.8	5.0	-0.13560	0.32489	0.18600	120	1.8	5.0	-0.15313	0.33312	0.17400
20	2.0	0.001	0.17147	0.23629	0.11558	60	2.0	0.001	0.17371	0.21793	0.16879	120	2.0	0.001	0.11650	0.19043	0.15191
20	2.0	0.1	0.19565	0.26536	0.13024	60	2.0	0.1	0.18346	0.23015	0.17159	120	2.0	0.1	0.12551	0.19436	0.15323
20	2.0	0.3	0.24120	0.30757	0.15913	60	2.0	0.3	0.19435	0.25063	0.17792	120	2.0	0.3	0.13451	0.19786	0.15448
20	2.0	0.5	0.24284	0.31447	0.16141	60	2.0	0.5	0.18042	0.25024	0.17828	120	2.0	0.5	0.13304	0.19856	0.15453
20	2.0	3.0	0.06127	0.23904	0.11183	60	2.0	3.0	-0.04489	0.27914	0.20505	120	2.0	3.0	-0.05519	0.27740	0.18535
20	2.0	5.0	-0.07470	0.34099	0.19487	60	2.0	5.0	-0.13881	0.37135	0.25824	120	2.0	5.0	-0.14248	0.36469	0.22878



(a) MSE



(b) BIAS

Fig. 6.2 3-D plots, PWP-GT model results for sample size of 120

6.4.2 AG and WLW models results

Tables 6.2-6.4 provide details of AG and WLW performance for the case of a stationary process (HPP) with risk-free interval, as sample size increases (from $U = 20$ to $U = 60$ and $U = 120$). Table 6.2-6.4 and Fig. 6.3 summarize performance of the AG and WLW methods (HPP case) for gap time ratio in the range $0.001 \le R \le 5.0$. The AG performance is consistently good across all values of the gap time ratio $0.001 \le R \le 5.0$. However, the WLW estimates improve as R and sample size increase.

In a small sample size ($U = 20$, Table 6.2), the variability of the WLW or the AG estimate is high, and the WLW estimates fluctuate more rapidly for $R > 0.5$. The more favorable applications range for AG lies within $R \le 5.0$, while the more favorable applications rage for WLW estimate is restricted in $R > 0.5$. The AG estimate lies between (2.31413, 2.43232), while the WLW estimate lies between (2.78490, 3.22381). The AG estimates have relative BIAS 0.03046, relative MAD 0.03046, and relative MSE 0.00145, while the WLW estimates have relative BIAS 0.33841, relative MAD 0.33841, and relative MSE 0.14360. The AG model appears capable of handling the recurrent data better than the WLW model.

At sample size $U = 60$ (Table 6.3), variability of both the AG and WLW estimates is reduced, leading to a narrower confidence interval. The more favorable applications range of the AG and WLW estimates lie within $R \le 5.0$ and $R > 0.5$, respectively. For the gap time ratio between $0.001 \le R \le 5.0$, the AG estimates lie between (2.33405, 2.41263), while the WLW estimates lie between (2.70554, 3.32081). The AG estimates have relative BIAS 0.03774, relative MAD

0.03774, and relative MSE 0.00189, while the WLW estimates have relative BIAS 0.36267, relative MAD 0.36267, and relative MSE 0.17043.

At the large sample case of $U = 120$ (Tables 6.4 and Fig. 6.3), the 95% C.I. show that the more favorable applications range of the AG and WLW estimates lie within $R \leq 5.0$ and $R > 0.5$, respectively. At the gap time ratio, $0.001 \leq R \leq 5.0$, the AG estimates lie between (2.29771, 2.34221), while the WLW estimates lie between (2.40654, 3.26163). The AG estimates have relative BIAS 0.01096, relative MAD 0.01167, and relative MSE 0.00020, while the WLW estimates have relative BIAS 0.30787, relative MAD 0.30787, and relative MSE 0.13860.

Table 6.2 Performance metrics of the AG and WLW models in an HPP, U =20

	AG model				WLW model			
R	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2
0.001	2.39562	0.04040	0.04040	0.00163	3.22381	0.40008	0.40008	0.16006
0.1	2.39388	0.03965	0.03965	0.00157	3.21717	0.39719	0.39719	0.15776
0.3	2.35567	0.02305	0.02305	0.00053	3.17877	0.38052	0.38052	0.14479
0.5	2.34470	0.01829	0.01829	0.00033	3.15384	0.36969	0.36969	0.13667
3.0	2.31413	0.00501	0.00501	0.00003	2.93244	0.27354	0.27354	0.07482
5.0	2.43232	0.05634	0.05634	0.00317	2.78490	0.20946	0.20946	0.04387
	BIAS(AG)= 0.03046				**BIAS(WLW)= 0.33841**			
		MAD(AG)= 0.03046				**MAD(WLW)= 0.33841**		
			MSE(AG)= 0.00145				**MSE(WLW)= 0.14360**	

[a] True β =2.30259

Table 6.3 Performance metrics of the AG and WLW models in an HPP, U =60

	AG model				WLW model			
R	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2
0.001	2.41245	0.04771	0.04771	0.00228	3.32081	0.44221	0.44221	0.19555
0.1	2.40282	0.04353	0.04353	0.00189	3.31100	0.43794	0.43794	0.19180
0.3	2.41263	0.04779	0.04779	0.00228	3.29226	0.42981	0.42981	0.18473
0.5	2.40234	0.04332	0.04332	0.00188	3.26629	0.41853	0.41853	0.17517
3.0	2.37265	0.03043	0.03043	0.00093	2.93013	0.27254	0.27254	0.07428
5.0	2.33405	0.01366	0.01366	0.00019	2.70554	0.17500	0.17500	0.03063
	BIAS(AG)= 0.03774				**BIAS(WLW)= 0.36267**			
		MAD(AG)= 0.03774				**MAD(WLW)= 0.36267**		
			MSE(AG)= 0.00189				**MSE(WLW)= 0.17043**	

[a] True β =2.30259

Table 6.4 Performance metrics of the AG and WLW models and the 95% C.I. in an HPP, U =120

	AG model						WLW model					
R	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2	95% LB	95% UB	Average $\hat{\beta}$ [a]	e_n	$\lvert e_n \rvert$	e_n^2	95% LB	95% UB
0.001	2.33565	0.01436	0.01436	0.00021	2.20459	2.46671	3.26163	0.41650	0.41650	0.17348	3.12090	3.40236
0.1	2.33665	0.01479	0.01479	0.00022	2.20584	2.46746	3.25388	0.41314	0.41314	0.17069	3.11342	3.39435
0.3	2.33680	0.01486	0.01486	0.00022	2.20683	2.46678	3.23144	0.40339	0.40339	0.16273	3.09203	3.37085
0.5	2.34221	0.01721	0.01721	0.00030	2.21281	2.47161	3.20009	0.38978	0.38978	0.15193	3.06246	3.33772
3.0	2.31796	0.00668	0.00668	0.00004	2.20375	2.43217	2.71537	0.17927	0.17927	0.03214	2.60338	2.82736
5.0	2.29771	-0.00212	0.00212	0.00000	2.19213	2.40329	2.40654	0.04514	0.04514	0.00204	2.30701	2.50606
	BIAS(AG)= 0.01096 **MAD(AG)= 0.01167** **MSE(AG)= 0.00020**						**BIAS(WLW)= 0.30787** **MAD(WLW)= 0.30787** **MSE(WLW)= 0.13860**					

[a] True β =2.30259

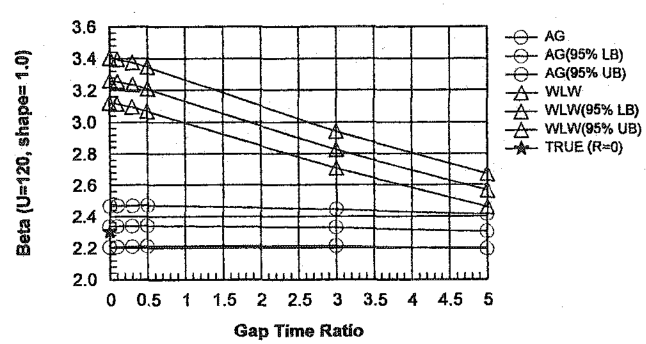


Fig. 6.3 HPP, AG and WLW estimates vs. Gap time ratio, U =120

6.5 Conclusions

The class of semi-parametric PI models applies to recurrent failure event modeling for a repairable system with covariates. A substantial period of down time, due to performing maintenance (i.e. major overhaul) after a major failure, has been a concern in the accuracy of estimating the covariate effect. This research examines the robustness of three semi-parametric PI models as a function of the overhaul duration. Qureshi et al. [6] assumed zero repair times in the PWP-GT model ($R = 0$). In comparing with other researchers, this study has defined the research domains for gap time ratio from a value close to zero repair times ($R = 0.001$) to long maintenance intervals ranging from 0.001 to 5.0. Qureshi et al. examined the PWP-GT model applied to recurrent data without considering the repair time process and concluded that the PWP-GT estimator underestimates the covariate effect in a DROCOF case (e.g., BIAS= -26% at $\delta = 0.5$) and overestimates the covariate effect in an IROCOF case (e.g., BIAS= 19% at $\delta = 3.0$). Qureshi et al. proved the PWP-GT model an accurate estimator in estimating the times to failures for NHPP power-law processes with shape parameter in the range $1.0 \leq \delta \leq 3.0$ and for larger sample sizes ($U \geq 60$). This study has considered both cases: zero repair times and long maintenance intervals and verified that the PWP-GT model results are consistent with those of Qureshi.

Recommendations to practitioners in selecting the more favorable applications ranges on (U, R, δ) are as follows. The PWP-GT model proves to perform well for sample sizes 60 (30 per class) or more, moderately decreasing, constant, and moderately increasing ROCOFs (power-law NHPP shape parameter in the range $0.8 \le \delta \le 1.8$) if the overhaul duration is within half time ($R \le 0.5$) of the previous instantaneous $MTTF$. If the overhaul duration is between three to five times of the previous instantaneous $MTTF$ ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2$. In the large sample size 120 (60 per class), the PWP-GT model performs well in the range of $0.5 \le \delta \le 2.0$, if the overhaul duration is within half time ($R \le 0.5$) of the previous instantaneous $MTTF$. If the overhaul duration is between three to five times of the previous instantaneous $MTTF$ ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2$. Within the short maintenance interval, increasing the sample size from 60 to 120 does not improve/widen the more favorable applications range for maintenance interval ($R \le 0.5$). As for the other two common baseline intensity model (i.e. AG and WLW), the AG model performs consistently well in the small sample size (20) regardless of the overhaul duration ($R \le 5.0$) in an HPP case. The WLW model performance improves as the overhaul duration increases ($R \ge 5.0$).

6.6 References

[1] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika, 1981; 68: 373-9.

[2] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 1982; 10(4): 1100-20.

[3] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, Dec.1989; 84(408): 1065-73.

[4] Cox DR. Regression models and lifetables (with discussion). Journal of the Royal Statistical Society B, 1972; 34:187-202.

[5] Landers TL, Soroudi HE. Robustness of a semi-parametric proportional intensity model. IEEE Transactions on Reliability, 1991; 40(2): 161-4.

[6] Qureshi WM, Landers TL, Edward EG. Robustness evaluation of a semi-parametric proportional intensity model. Reliability Engineering and System Safety, 1994; 44: 103-9.

[7] Vithala S. Robustness of a semi-parametric proportional intensity model for the case of a log-linear Non-homogeneous Poisson Process. A thesis of the Industrial Engineering Department at the University of Arkansas, 1994.

[8] Landers TL, Jiang ST, Peek, JR. Semi-parametric PWP model robustness for log-linear increasing rates of occurrence of failures. Reliability Engineering and System Safety, 2001; 73: 145-53.

[9] Hansen CK, Ascher HE. Modeling repair events under intermittent failures and failures subject to unsuccessful repair. Quality and Reliability Engineering International, 2002;18:453-65.

[10] Kobbacy KAH, Jeon J. Generalized non-stationary preventive maintenance model for deteriorating repairable systems. Quality and Reliability Engineering International, 2002;18:363-72.

[11] Therneau TM, Hamilton SA. rhDNase as an example of recurrent event analysis. Statistics in Medicine 1997; 16: 2029-47.

[12] Ansell JI, Phillips MJ. Practical problems in the statistical analysis of reliability data (with Discussion). Applied Statistics, 1989;38(2):205-47.

[13] Ansell JI, Phillips MJ. Practical reliability data analysis. Reliability Engineering and System Safety, 1990;28(3):337-56.

[14] Ansell JI, Phillips MJ. Practical aspects of modeling of repairable systems data using proportional hazards models. Reliability Engineering and System Safety, 1997;58(2):165-71.

[15] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Assessing the maintenance in a process using a semi-parametric approach. Quality and Reliability Engineering International, 2001;17:163-7.

[16] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Analysing maintenance data to gain insight into systems performance. Journal of Operational Research Society, 2003;54:343-9.
[17] Ascher H. Regression analysis of repairable systems reliability. In: Skwirzynski JK, editor. Electronic systems effectiveness and life cycle costing. Berlin: Springer, 1983, p.119-33.
[18] Ascher H, Feingold H. Repairable systems reliability--modeling inference, misconceptions and their causes, Lecture notes in statistics, New York: Marcel Dekker Inc, 1984;7.
[19] Dale C. Application of the proportional hazards model in the reliability field. Reliability Engineering, 1985;10:1-14.
[20] Wightman DW, Bendell A. The practical application of proportional hazards modeling. Reliability Engineering, 1986;15:29-53.
[21] Bendell A, Wightman DW, Walker EV. Applying proportional hazards modeling in reliability. Reliability Engineering and System Safety, 1991;34:35-53.
[22] Bowman ME. An evaluation of statistical models for the analysis of recurrent events data: with application to needlestick injuries among a cohort of female veterinarians. A dissertation of the Preventive Medicine at the Ohio State University, 1996.
[23] Lin DY. Cox's regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine, 1994; 13: 2233-47.
[24] Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Statistics in Medicine, 1997; 16: 833-9.
[25] Lipschutz KH, Snapinn SM. Discussion of paper by Wei and Glidden. Statistics in Medicine, 1997; 16: 841-51.
[26] Kelly PJ, Lim LL-Y. Survival analysis for recurrent data: An application to childhood infectious diseases. Statistics in Medicine, 2000; 19: 13-33.
[27] Blanks HS, Tordon MJ. Laplace and Mann aging trend test effectiveness. Proceeding of the Annual Reliability and Maintainability Symposium, 1987;165-70.
[28] Proschan F. Theoretical explanation of observed decreasing failure rate. Technometrics, Aug. 1963;5(3):375-83.
[29] SUGI Supplementary Library User's Guide, Version 5. SAS Institute, Cary, North Carolina, USA, 1985, p. 437-66.

7. Covariate proportional intensity modeling for recurrent data of two failure types (major and minor)

7.0 Abstract

This paper examines covariate proportional intensity (PI) modeling as an approach for explicit treatment of multiple (two) recurrent failure types (major and minor) with complete data following a power-law Non-homogeneous Poisson Process (NHPP). Although covariates are typically used to incorporate treatment effects, a covariate is shown to conceptually model multiple failure types in the special case where the proportional intensities rule holds. The Prentice-Williams-Peterson-gap time (PWP-GT) model has proven a robust and accurate estimator in handling recurrent data of two failure types. The more favorable engineering application ranges are recommended, which are beneficial to practitioners in anticipating the favorable application domains.

For the minor type, the PWP-GT model proves to perform well for sample size 120 (60 per class) or more, decreasing, constant, and increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 2.0$). For the major type, the PWP-GT model proves to perform well for sample size 180 (90 per class) or more, decreasing, constant, and increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 1.8$).

Keywords: repairable systems reliability, recurrent events, multiple failure types, covariate proportional intensity modeling

Nomenclature

Acronyms

AG	Andersen and Gill model
DROCOF	Decreasing rate of occurrence of failures
HPP	Homogeneous Poisson Process
IROCOF	Increasing rate of occurrence of failures
i.i.d	Independent and identically distributed
MTTF	Mean time to failure
MAD	Mean absolute deviation
MSE	Mean squared error
NHPP	Non-homogeneous Poisson Process
PI	Proportional intensity
PWP	Prentice, Williams, and Peterson model
PWP-GT	Prentice, Williams, and Peterson-gap time model
PWP-TT	Prentice, Williams, and Peterson-total time model
WLW	Wei, Lin, and Weissfeld model

Notation

$h(t;z)$	Proportional hazard function
$h_0(t)$	Baseline hazard function
$N(t)$	Random variable for the number of failures in $(0,t]$; a counting process
n	An integer counting successive failure times; a stratification indicator subscript
R_{1i}	Treatment factor for major type failures, i denotes the level number
R_{2i}	Treatment factor for minor type failures, i denotes the level number
T_n	Random variable for cumulative time of occurrence of the n^{th} failure
U	Sample size (number of units)
$\mathbf{Z}(t)$	Covariate process up to time t
$\mathbf{Z}_1(t)$	A two-dimensional covariate for major type failures

$\mathbf{Z}_2(t)$	A two-dimensional covariate for minor type failures
$\mathbf{z}$	$(k \times 1)$ vector of covariates, $\mathbf{z} = (z_1, z_2, ... z_k)'$
$\boldsymbol{\beta}_n$	$(k \times 1)$ vector of stratum-specific regression coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)$
$\boldsymbol{\beta}_{1n}$	regression coefficient for major failure events, n: event count
$\boldsymbol{\beta}_{2n}$	regression coefficient for minor failure events, n: event count
δ	Shape parameter of a power-law NHPP
δ_1	Shape parameter of the major type events
δ_2	Shape parameter of the minor type events
Δ	Indicator of a failure or censored time; limit to time zero
$\lambda_{0n}(t)$	Stratum-specific baseline intensity function
$\lambda(t;\mathbf{z})$	Proportional intensity function
υ	Scale parameter of a power-law NHPP
υ_0	Baseline value of υ, the scale parameter of a power-law NHPP
υ_1	Alternate value of υ, the scale parameter of a power-law NHPP
σ	Standard deviation
^	Denotes an estimator
′	Denotes the transpose of a vector

7.1 Introduction

Failure time data on a repairable system are realizations of a stochastic point process, in which the instantaneous rate of occurrence of failures (ROCOF) is $\lambda(t)$. Prentice, Williams, and Peterson (PWP) [1] proposed a semi-parametric approach to model recurrent failure event data from a repairable system using two methods: PWP-GT (gap time) and PWP-TT (total time). Cox proposed the distribution-free (semi-parametric) proportional hazards (PH) model in 1972 [2]. The proportional intensity (PI) models (PWP-GT, PWP-TT, Andersen-Gill (AG)

[3], and Wei-Lin-Weissfeld (WLW) [4]) have been applied to recurring events in medical studies (biostatistics field), such as recurrent infections of a patient. For engineering applications, Landers and Soroudi [5], Qureshi et al. [6], Vithala [7], and Landers et al. [8] have investigated robustness of the PWP-GT model, where the underlying recurrent failure time data are from a Non-homogeneous Poisson Process (NHPP) with a power-law or a log-linear intensity function. These references also report the engineering applications of the PWP-GT model cited in the literature. Qureshi et al. [6] found that the PWP-GT model performs best for constant and moderately increasing rate of occurrence of failures (IROCOF) and decreasing rate of occurrence of failures (DROCOF) and for larger sample sizes from power-law NHPPs. Vithala [7] considered the case of log-linear increasing rates of occurrence of failures, and concluded the PWP-GT model performs best for moderately increasing rates of occurrence of failures and for larger sample sizes. This research has extended their work to the important case of recurrent data of two failure types (major and minor).

Compared to the extensive literature on applications of the Cox-based regression models in the biostatistics field, there have been few reported engineering application. Abundant federal funding received in biostatistics / medical research has advanced the PI models to become well developed and widely referenced. PI models for medical applications could also apply to recurring failure/repair data in engineering problems. The PWP-GT, PWP-TT, AG, and WLW models offer powerful analytical tools for engineering practitioners as they become better recognized and understood.

Major and minor failure events are commonly seen in industry, where minor failure rate is typically proportionally higher than major failure rate. Most researchers have formulated the problem as univariate and pooled the major and minor failures as though they are identical. The Lin method of multi-dimensional covariates permits explicit modeling of major and minor failures as distinct types in a single, stratified model so long as the proportional intensity rule holds.

7.2 Semi-parametric Proportional Intensity models

Cox [4] proposed a PH formulation to include explanatory variables (covariates) in survival models. PWP proposed an extension of the Cox model to stochastic processes and applied the approach to model recurring infections in aplastic anemia and leukemia patients having received bone-marrow transplants. This application involved several subjects and a small number of events (up to five) for each subject. The paper by PWP did not address the baseline intensity function but rather reported the relative risks for the test and control groups. In reliability and maintainability engineering applications, a number of authors have applied the semi-parametric proportional intensity (hazards) model, for example, Ansell and Phillips [9], Ansell and Phillips [10], Landers and Soroudi [5], Qureshi et al. [6], Ansell and Phillips [11], Landers et al. [8], Ansell et al. [12], and Ansell et al. [13]. A collection of the PI model applied to different industries includes: marine gas turbine engines (Asher [14]), semiconductor, electrical, and pipeline industries (Ansell and Phillips [11]), U.S. Army main battle tank (Landers et al. [8]), water supply industry (Ansell et al. [12], [13]), etc. Ascher [14] illustrated the use of the PWP model for analysis of reliability for marine gas turbine engines.

Ascher and Feingold [15] suggested application of the PWP model in the field of reliability engineering. Dale [16] applied the PWP approach to simulated data for a reliability growth program with design improvements implemented after each of the five stages, resulting in a decreasing rate of occurrence of failures (DROCOF). Wightman and Bendell [17] and Bendell et al. [18] cited the PWP model and advised caution in application for engineering studies.

Qureshi et al. [6] performed a robustness study to determine how well the PWP-GT method performed when applied to data from a failure process that was actually parametric (specifically the NHPP with power-law intensity function). The 2σ bounds of the PWP-GT estimates can cover the true values with few exceptions. The PWP-GT method performed well, except at small values of shape parameter ($\delta < 0.6$). The PWP-GT method was best for larger sample size and for moderately decreasing, constant, and moderately increasing ROCOF. The validation process for the case of an HPP in Section 2.2.3 (also refer to Table 2.10) indicated that the estimated $MTTF$ (mean time to failure) differences between the PWP-GT model and theoretical values were not statistically significant. As for the PWP-GT estimates of the covariate regression coefficient, the true value of coefficient β lies within the 2σ confidence bounds on the estimate $\hat{\beta}$ for $1.0 \le \delta \le 1.4$. The PWP-GT method tends to underestimate β for a DROCOF (e.g., BIAS= -26% at $\delta = 0.5$) and overestimate β for an IROCOF (e.g., BIAS= 19% at $\delta = 3.0$).

Lin [19, 20] studied chronic granulomatous disease and employed a multiple dimensional covariate method to handle the recurrent data with multiple failure types. Lin considered three types of failure outcomes by defining a three-dimensioned covariate array. For the special case of two failure types considered in this research, let two covariates Z_1, Z_2 represent the major and minor failure types in two dimensions. That is, the major and minor failure types are coded as follows. (1) Major type: $Z_1 = [R_{1i}, 0]$ and (2) Minor type: $Z_2 = [0, R_{2i}]$, where $R_i = 1$ for class1 and $R_i = 0$ for class ϕ. The corresponding regression coefficient estimates are interpreted as the covariate effect applied to the major and minor failure types. To estimate the general covariate effect for major/minor event types altogether, a single covariate is introduced, and $Z_{ik} = R_i$, where $R_i = 0$, for class=ϕ and $R_i = 1$, for class=1 (i: sample unit; k: the failure type). Refer to Lin [19] for an illustration of the data set structure.

In industry, minor failure rate is typically higher than major failure rate. Most researchers have formulated this problem as univariate and pooled the multiple failure types as though they are identical. The Lin method of multi-dimensional covariates permits explicit modeling of major and minor failures as distinct failure types in a single, stratified model so long as the proportional intensity rule holds.

7.3 Models and methods

Sections 7.3.1-7.3.2 review the semi-parametric Cox regression model for single event and the PWP-GT model for recurrent events. Section 7.3.3 reviews the NHPP with power-law intensity function. Section 7.3.4 describes the method used to assess the robustness of the semi-parametric PI method for the case of completed data from a true but unknown power-law NHPP.

7.3.1 Cox regression model

For the case of a time-to-failure random variable, Cox [2] proposed a proportional hazards regression model of the form:

$$h(t;\mathbf{z}) = h_0(t)\exp(\boldsymbol{\beta}'\mathbf{z}), \tag{1}$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The PH model is composed of two parts: baseline hazards function $h_0(t)$ and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect. The Cox model can be used to describe the semi-parametric distribution of time-to-failure for non-repairable items with covariates. Under proportional hazards, the ratio of the hazard functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline hazard function. When the baseline hazard function is fully specified (e.g.,Weibull) the analytical procedure is termed a parametric method. Alternatively, $h_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

7.3.2 Semi-parametric PWP model

The PWP model is a generalization of the semi-parametric Cox proportional hazard function to a proportional intensity function $\lambda(t;\mathbf{z})$ for the case of repeated failure events. Under proportional intensities, the ratio of the intensity functions of two units (A and B) with covariate vectors $\mathbf{z}_A$ and $\mathbf{z}_B$ is constant over time. The covariates have a multiplicative effect on the baseline intensity function. When the baseline intensity function is fully specified (e.g., power-law or log-linear) the analytical procedure is termed a parametric method. Alternatively, $\lambda_0(t)$ can be left arbitrary, in which case the procedure is termed semi-parametric.

Given the counting and covariate processes at time t, the general semi-parametric intensity function was defined by PWP as follows:

$$\lambda\{t \mid N(t), \mathbb{Z}(t)\} = \lim \ \Pr\{ t \leq T_{n(t)+1} < t + \Delta \mid N(t), \mathbb{Z}(t)\} / \Delta , \tag{2}$$

where $N(t)$ represents a random variable for the number of failures in $(0,t]$, $\mathbb{Z}(t)$ denotes the covariate process up to time t, and Δ limits the time span to zero. Among the semi-parametric regression models specified by PWP were the following:

$$PWP-GT : \lambda\{t \mid N(t), \mathbb{Z}(t)\} = \lambda_{0n}(t - t_{n-1}) \exp[\boldsymbol{\beta}_n' \mathbf{z}(t)] \tag{3}$$

$$PWP-TT : \lambda\{t \mid N(t), \mathbb{Z}(t)\} = \lambda_{0n}(t) \exp[\boldsymbol{\beta}_n' \mathbf{z}(t)]. \tag{4}$$

In the PWP-GT model of Eq. (3), the time metric is the interval between times of failure t_{n-1} and t_n, defined as gap time. The PWP model stratifies a failure data set based on the failure event count. When a unit is placed into operation it has experienced no failures and so resides in stratum 1 ($n = 1$), and when the first failure occurs the unit moves to the second stratum ($n = 2$). In general, the unit moves to stratum n immediately following the $(n-1)^{st}$ failure and remains there until the n^{th} failure.

7.3.3 Power-law intensity function

For a power-law NHPP, the baseline intensity function is

$$\lambda_0(t) = v_0 \delta \times t^{\delta-1} , \tag{5}$$

where δ is the shape parameter and υ is the scale parameter of the power-law form. If we define $\upsilon_0 = \exp(\beta_0 \times z_0)$ and $z_0 = 1$, then the power-law PI model becomes

$$\lambda(t;\mathbf{z}) = \delta \times t^{\delta-1} \exp(\boldsymbol{\beta}'\mathbf{z}), \qquad (6)$$

where $\boldsymbol{\beta}$ is the regression coefficient vector and $\mathbf{z}$ represents a covariate vector. The power-law intensity function is composed of two parts: baseline intensity function that follows a power-law form and an exponential link function, where $\boldsymbol{\beta}$ is designated to measure the covariate effect.

This process could model the reliability of a repairable system with rapid deterioration, since the failure intensity is increasing at an exponential rate with time. The analogous case for maintainability is a rapid learning process. The intensity function $\lambda(t)$ is strictly decreasing for δ < 1, constant for δ = 1, and strictly increasing for δ > 1. Thus, we have a DROCOF for δ < 1, an HPP for δ = 1, and an IROCOF for δ > 1.

7.3.4 Method

Simulation data with recurring patterns (complete data), where the underlying distribution follows a power-law NHPP, is generated by a modified Blanks & Tordon [21] simulation algorithm. The concept of forming the dataset for the PWP method originates from the probability theory of conditionality. The later failure times cannot be included into the dataset when developing the intensity function at the n^{th} failure count. Consequently, for each censored unit, the censored failure times are removed from the dataset, except the first censored failure event. The pattern of the dataset shows that the last record in each censored unit has one and only one censored status. The logic in generating the dataset for the PWP method is to remove the record if both of the following conditions hold: (1) the current record is marked censored and (2) the previous record is marked censored.

The semi-parametric PWP-GT method was implemented using the SAS™ Users Group (SUGI) software code PHREG [22], which performs the semi-parametric Cox regression method with a blocking option to stratify for a covariate, such as failure event count, not satisfying the proportional hazards conditions. PHREG applies the product-limit method to estimate the reliability function within all strata defined by the failure count and for all values of the covariate. PHREG also applies the Cox method to estimate the vector of regression coefficients $\boldsymbol{\beta}$ and the covariance matrix. Appendix VI provides the programming code to perform the semi-parametric PWP-gap time method.

The simulation method of Blanks & Tordon [21] is modified to generate an NHPP with two failure types, where the underlying distribution is a power-law intensity function. Most of the parameters remain unchanged except that the sample unit size has been increased due to the insufficient sample size of major events. The parameter setting is as follows: $U = 120$, $F = 10$, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$.

The fixed time-invariant covariate vector $z_i, i = 1,2$ is defined as follows:

Major event, Class=ϕ: $z_1 = (0,0)$

Major event, Class=1: $z_1 = (1,0)$

Minor event, Class=ϕ: $z_2 = (0,0)$

Minor event, Class=1: $z_2 = (0,1)$

Ten failure events were generated for each sample unit. To determine the sequence of major and minor failure events in the counting process, a uniformly distributed random variate $U(0,1)$ was introduced to decide the event number F for occurrence of the major failure. Consequently, the F^{th} event time to have a major failure is generated as: $F = FLOOR(10 \times RANUNI(SEED)) + 1$. The remaining nine events are minor failure events. In this way, a counting process contains major and minor failure events, where the one major failure is inserted randomly among the $N-1$ minor failures. The event number for the major failure is randomly selected depending on the F value. Large enough sample size is generated in order to obtain sufficient data for each failure count in a PWP-GT model.

There are two experimental factors: experimental units (U) and the shape parameter (δ). The levels for each factor are selected as follows: (1) U = 120, 180, and 240 (2) δ=0.5, 0.8, 1.0, 1.2, 1.5, 1.8, and 2.0. The selection of the υ_0, υ_1 and δ levels has taken the following considerations: (1) the parameter settings in the previous relevant works (Proschan [23], Landers and Soroudi [5], Qureshi et al.[6], and Landers et al. [8]) (2) The selection of U levels is taken from the parameter settings in the previous research works, and it has also considered the small, median, and large sample sizes for both major and minor types of failure events. To measure and compare model performance, three robustness metrics were compiled:

- relative signed error (BIAS);
- relative mean absolute deviation (MAD) and

- relative mean squared error (MSE).

7.4 Results

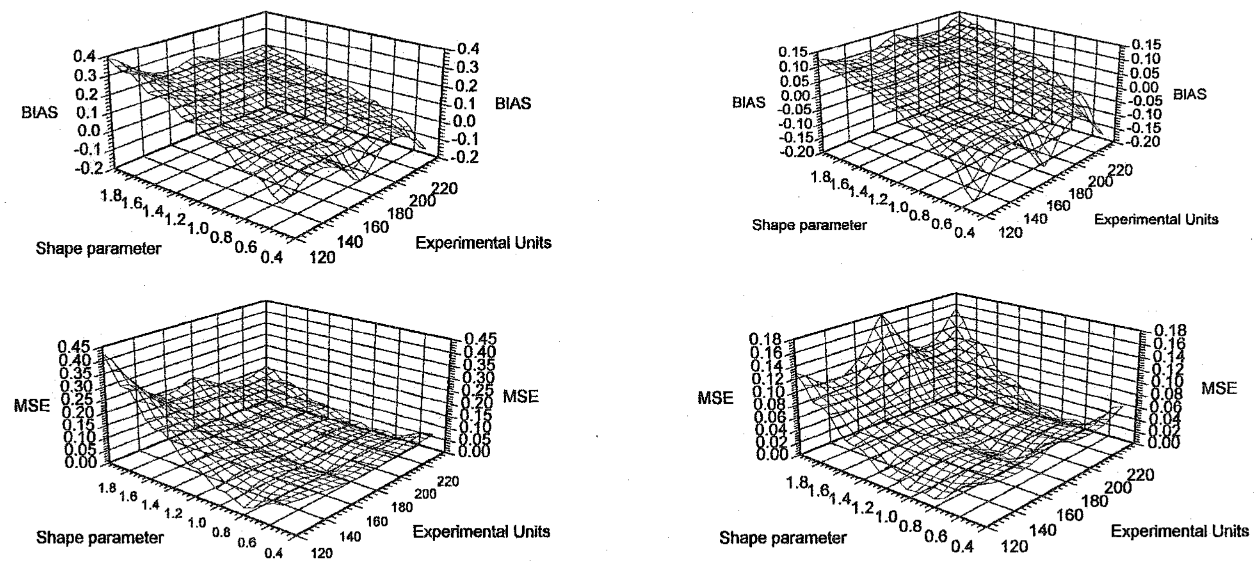
Table 7.1 and Fig. 7.1 summarize all the performance metrics in terms of U and δ, categorized by the major and the minor types. The treatment effect (CLASS= ϕ., CLASS=1) for the major type is summarized as follows. In the case of $U = 120$, the more favorable applications range for PWP-GT estimates is $0.5 \le \delta \le 1.2$, having relative MSE within the range of (2.7%, 10.8%), relative BIAS within the range of (-4.9%, 19.2%), and relative MAD within the range of (12.6%, 25.3%). In the case of $U = 180$, increasing the sample size (from 120 to 180) has significantly enhanced the accuracy of the PWP-GT model resulting in a wider applicable range of the shape parameter $0.5 \le \delta \le 1.8$ $(U = 180)$ than $0.5 \le \delta \le 1.2$ $(U = 120)$. For the more favorable applications range ($0.5 \le \delta \le 1.8$), the PWP-GT model estimates have relative MSE within the range of (1.0%, 15.8%), relative BIAS within the range of (-14.7%, 22.4%), and relative MAD within the range of (8.5%, 26.0%). As for $U = 240$, the more favorable applications range is $0.5 \le \delta \le 2.0$, with relative MSE within the range of (1.0%, 16.0%), relative BIAS within the range of (-17.3%, 19.1%), and relative MAD within the range of (7.9%, 24.9%).

The treatment effect (CLASS=ϕ., CLASS=1) for the minor type is summarized as follows. In the case of $U = 120$, the more favorable applications range for PWP-GT model estimates is $0.5 \le \delta \le 2.0$, having relative MSE within the range of (1.0%, 13.0%), relative BIAS within the range of (-16.5%, 12.5%), and relative MAD within the range of (7.1%, 22.4%). In the case of $U = 180$, the

PWP-GT model performs well within the range ($0.5 \le \delta \le 2.0$), having relative MSE within the range of (0.9%, 17.9%), relative BIAS within the range of (-18.1%, 13.4%), and relative MAD within the range of (6.4%, 21.9%). As for the case of $U = 240$, the more favorable applications range is $0.5 \le \delta \le 2.0$, with relative MSE within the range of (0.5%, 15.1%), relative BIAS within the range of (-18.5%, 13.5%), and relative MAD within the range of (5.4%, 19.8%).

Table 7.1 Summary of PWP-GT model results for estimating $\hat{\boldsymbol{\beta}}_i$ (10 failures/unit)

N = 10 failures/unit, $\upsilon_0 = 0.001, \upsilon_1 = 0.01$							
		Major events			Minor events		
U	δ	BIAS	MAD	MSE	BIAS	MAD	MSE
120	0.5	-0.04875	0.19879	0.08948	-0.16487	0.16487	0.05101
	0.8	0.04247	0.12553	0.02734	-0.05810	0.07934	0.01233
	1.0	0.12466	0.18186	0.05697	-0.00542	0.07079	0.00951
	1.2	0.19176	0.25272	0.10830	0.03509	0.10386	0.01791
	1.5	0.27228	0.33229	0.19928	0.07186	0.15342	0.04673
	1.8	0.34679	0.41784	0.32304	0.10206	0.19847	0.09108
	2.0	0.38869	0.46815	0.42425	0.12479	0.22399	0.12983
180	0.5	-0.14726	0.14726	0.04191	-0.18054	0.18054	0.05189
	0.8	-0.00717	0.08513	0.00958	-0.06023	0.07216	0.00962
	1.0	0.06193	0.10476	0.01664	-0.00458	0.06357	0.00853
	1.2	0.11448	0.15047	0.03710	0.03664	0.10441	0.02097
	1.5	0.17808	0.20435	0.08590	0.07986	0.15328	0.06061
	1.8	0.22432	0.26021	0.15800	0.11377	0.19468	0.12357
	2.0	0.25701	0.29262	0.22060	0.13394	0.21890	0.17854
240	0.5	-0.17263	0.18084	0.05756	-0.18549	0.18549	0.05539
	0.8	-0.03609	0.09453	0.01157	-0.05482	0.06759	0.00790
	1.0	0.02633	0.07927	0.00969	-0.00027	0.05360	0.00498
	1.2	0.07002	0.11655	0.02048	0.03953	0.08610	0.01430
	1.5	0.12523	0.16866	0.05612	0.07040	0.14857	0.05234
	1.8	0.16515	0.21480	0.11089	0.11493	0.17533	0.10220
	2.0	0.19105	0.24873	0.16020	0.13536	0.19800	0.15075



(a) Major failures

(b) Minor failures

Fig. 7.1 3-D plots, PWP-GT model results for estimating $\hat{\beta}_i$ (10 failures/unit)

7.5 Conclusions

This research also examines covariate proportional intensity modeling as an approach for explicit treatment of multiple (two) recurrent failure types (major and minor) with complete data. Although covariates are typically used to incorporate treatment effects, a covariate is shown to conceptually model multiple failure types in the special case where the proportional intensities rule holds. The PWP-GT model proves to be the model of choice to handle two failure types of recurring events, evaluated in terms of bias, mean absolute deviation, and mean squared error of covariate regression coefficients over ranges of sample sizes and shape parameters encountered in engineering applications. The more favorable engineering applications ranges are recommended.

The research domains of the two factors of interests are: (1) $120 \le U \le 240$ and (2) $0.5 \le \delta \le 2.0$. For the minor failure type, the PWP-GT proves to perform well for sample sizes 120 (60 per class) or more, decreasing, constant, and increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 2.0$). For the major failure type, the PWP-GT performs well for sample sizes 180 (90 per class) or more, decreasing, constant, and increasing rates of occurrence of failures (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 1.8$).

The recurring events of two failure types (major and minor failures) in this study were generated from a single NHPP stream with power-law intensity function, where the major and minor failure events share the same shape parameter (δ) of the power-law form. To meet the requirement of proportionality in the semi-parametric proportional regression method, the shape parameters of the major type (δ_1) and the minor type (δ_2) are set to equal, expressed as $\delta_1 = \delta_2 = \delta$. In practice, the case $\delta_2 \neq \delta_1$ is likely. Future research may propose a model that handles (1) $\delta_1 < \delta_2$ or (2) $\delta_1 > \delta_2$.

7.6 References

[1] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika, 1981; 68: 373-9.

[2] Cox DR. Regression models and lifetables (with discussion). Journal of the Royal Statistical Society B, 1972; 34:187-202.

[3] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 1982; 10(4): 1100-20.

[4] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, Dec.1989; 84(408): 1065-73.

[5] Landers TL, Soroudi HE. Robustness of a semi-parametric proportional intensity model. IEEE Transactions on Reliability, 1991; 40(2): 161-4.

[6] Qureshi WM, Landers TL, Edward EG. Robustness evaluation of a semi-parametric proportional intensity model. Reliability Engineering and System Safety, 1994; 44: 103-9.

[7] Vithala S. Robustness of a semi-parametric proportional intensity model for the case of a log-linear Non-homogeneous Poisson Process. A thesis of the Industrial Engineering Department at the University of Arkansas, 1994.

[8] Landers TL, Jiang ST, Peek, JR. Semi-parametric PWP model robustness for log-linear increasing rates of occurrence of failures. Reliability Engineering and System Safety, 2001; 73: 145-53.

[9] Ansell JI, Phillips MJ. Practical problems in the statistical analysis of reliability data (with Discussion). Applied Statistics, 1989;38(2):205-47.

[10] Ansell JI, Phillips MJ. Practical reliability data analysis. Reliability Engineering and System Safety, 1990;28(3):337-56.

[11] Ansell JI, Phillips MJ. Practical aspects of modeling of repairable systems data using proportional hazards models. Reliability Engineering and System Safety, 1997;58(2):165-71.

[12] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Assessing the maintenance in a process using a semi-parametric approach. Quality and Reliability Engineering International, 2001;17:163-7.

[13] Ansell J, Archibald T, Dagpunar J, Thomas L, Abell P, Duncalf D. Analysing maintenance data to gain insight into systems performance. Journal of Operational Research Society, 2003;54:343-9.

[14] Ascher H. Regression analysis of repairable systems reliability. In: Skwirzynski JK, editor. Electronic systems effectiveness and life cycle costing. Berlin: Springer, 1983, p.119-33.

[15] Ascher H, Feingold H. Repairable systems reliability--modeling inference, misconceptions and their causes, Lecture notes in

statistics, New York: Marcel Dekker Inc, 1984;7.
[16] Dale C. Application of the proportional hazards model in the reliability field. Reliability Engineering, 1985;10:1-14.
[17] Wightman DW, Bendell A. The practical application of proportional hazards modeling. Reliability Engineering, 1986;15:29-53.
[18] Bendell A, Wightman DW, Walker EV. Applying proportional hazards modeling in reliability. Reliability Engineering and System Safety, 1991;34:35-53.
[19] Lin DY. MULCOX2: A general computer program for the Cox regression analysis of multivariate failure time data. Computer Methods and Programs in Biomedicine. 1993; 40: 279-93.
[20] Lin DY. Cox's regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine 1994; 13: 2233-47.
[21] Blanks HS, Tordon MJ. Laplace and Mann aging trend test effectiveness. Proceeding of the Annual Reliability and Maintainability Symposium, 1987;165-70.
[22] SUGI Supplementary Library User's Guide, Version 5. SAS Institute, Cary, North Carolina, USA, 1985, p. 437-66.
[23] Proschan F. Theoretical explanation of observed decreasing failure rate. Technometrics, Aug. 1963;5(3):375-83.

8 Conclusions and recommendations

The class of semi-parametric PI models applies to recurrent failure event modeling for a repairable system with covariates. This research has provided a thorough robustness study of four semi-parametric PI models (PWP-GT, PWP-TT, AG, and WLW) subject to right-censoring severity and two distinct types of recurring events ((1) major overhaul duration and (2) major and minor failures). Two modeling extensions are examined for the case of multiple event types: multi-dimensional covariate (Lin (1993, 1994)) and discontinuous risk-free-intervals (Therneau and Hamilton (1997)). Recommendations for the more favorable applications range on the parameters in each individual study (from Chapters 4-7) are available for the prospective industrial applications, such as aircraft and power plants. The results are beneficial to practitioners in anticipating the more favorable applications domains and selecting appropriate PI models in repairable system reliability.

8.1 Conclusions

8.1.1 Right-censoring recurring events on four PI models

Previous studies (by Landers and Soroudi (1991) and Qureshi et al. (1994)) conducted on the PWP-GT model for the case of an underlying NHPP with power-law intensity function indicated good performance. This research has performed a right-censorship robustness study and examined other semi-parametric PI models with covariates. The PWP-GT and AG prove to be models of choice, evaluated in terms of the BIAS, MAD, and MSE of covariate regression coefficients over ranges of sample sizes, shape parameters, and censoring severity typically encountered in engineering applications. The more favorable engineering applications ranges are recommended (Section 8.2.1).

8.1.2 Right-censoring effect on event-specific PI models

The research studied the robustness of three event-specific baseline models (PWP-GT, PWP-TT, and WLW) and a common baseline model (AG) to recurring failure events with right-censoring effect from a Poison Process. The PWP-GT and AG prove to be models of choice, evaluated in terms of the BIAS, MAD, and MSE of covariate regression coefficients over ranges of sample sizes and censoring severity typically encountered in engineering applications. The favorable engineering applications ranges are recommended (Section 8.2.2).

8.1.3 Discontinuous risk-free-intervals

The class of semi-parametric PI models applies to recurrent failure event modeling for a repairable system with covariates. A substantial period of downtime, due to performing maintenance (i.e. major overhaul) following a major failure, has been a concern in the accuracy of estimating the covariate effect. The event-specific PWP-GT model proves to be the model of choice to estimate the covariate effect, stratum by stratum, if the overhaul duration is short. The AG model performs well in an HPP regardless of sample size and overhaul duration. This research examines the robustness of three semi-parametric PI models as a function of the overhaul duration (Section 8.2.3).

8.1.4 Covariate PI modeling

This research also examines covariate PI modeling as an approach for explicit treatment of multiple (two) recurrent failure types (major and minor) with complete data. Although covariates are typically used to incorporate treatment effects, a covariate is shown to conceptually model multiple failure types in the special case where the proportional intensities rule holds. The PWP-GT model proves to be the model of choice to handle two failure types of recurring events, evaluated in terms of BIAS, MAD, and MSE of covariate regression coefficients over ranges of sample sizes and shape parameters typically encountered in engineering applications. The more favorable engineering applications ranges are recommended (Section 8.2.4).

8.2 Recommendations

8.2.1 Right- censoring effect on four PI models

The research domains of the three factors of interests are: (1) $60 \le U \le 180$, (2) $0.5 \le \delta \le 2.0$, and (3) $0.0 \le P_c \le 1.0$. At the smaller sample size (U =60), the PWP-GT proves to perform well for moderate right-censoring ($0.0 \le P_c \le 0.8$) and moderately decreasing, constant, and moderately increasing ROCOFs (power-law NHPP shape parameter in the range of $0.8 \le \delta \le 1.8$). In the case of U =120, the PWP-GT proves to perform well for moderate right-censoring ($0.0 \le P_c \le 0.8$) and moderately decreasing, constant, and moderately increasing ROCOFs (power-law NHPP shape parameter in the range of $0.8 \le \delta \le 2.0$). For the large sample size (U =180), the PWP-GT performs well for severe right-censoring ($0.0 \le P_c \le 1.0$) and moderately decreasing, constant, and moderately increasing ROCOFs (power-law NHPP shape parameter in the range of $0.8 \le \delta \le 2.0$). The AG model proves to outperform the WLW for stationary process (HPP) across a wide range of right-censorship ($0.0 \le P_c \le 1.0$) and for sample sizes of 60 (30 per class) or more.

8.2.2 Right-censoring effect on event-specific PI models

The research domains of the two factors of interests are: (1) $60 \le U \le 180$ and (2) $0.0 \le P_c \le 1.0$. The parameter setting $P_c = 0$ (complete data) is included for comparison with censored data. The PWP-GT model has proven the most robust and accurate estimator (at the lowest error) among the three event-specific models. Compared to WLW, the PWP-TT estimator yields similar but slightly better results. The PWP-GT presents a low-error region at the range of $120 \le U \le 180$ and $0 \le P_c \le 1$. For the small sample size $U = 60$, the more favorable applications range is $0 \le P_c \le 0.8$. For the other two estimators, when the sample size is increased from $U = 60$ to $U = 120$, PWP-TT and WLW have a slightly improved applications range ($0 \le P_c \le 0.4$). As the sample size is increased to 180, the performance is poor but stable over applications range $0 \le P_c \le 0.8$ on both models. The results show that AG performs well for the case of smaller sample size (U=60) and severe censoring ($P_c = 1.0$). The favorable applications region of the common baseline AG model is $60 \le U \le 180$ and $0 \le P_c \le 1$.

8.2.3 Discontinuous risk-free-intervals

The research domains of the three factors of interests are: (1) $20 \le U \le 120$, (2) $0.5 \le \delta \le 2.0$, and (3) $0.001 \le R \le 5.0$. The PWP-GT model proves to perform well for sample sizes 60 (30 per class) or more, moderately decreasing, constant, and moderately increasing rate of occurrence of failures (power-law NHPP shape parameter in the range $0.8 \le \delta \le 1.8$) if the overhaul duration is within half time ($R \le 0.5$) of the previous instantaneous $MTTF$. If the overhaul duration is between three to five times of the previous instantaneous $MTTF$ ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2.$ In the large sample size 120 (60 per class), the PWP-GT model performs well in the range of $0.5 \le \delta \le 2.0,$ if the overhaul duration is within half time ($R \le 0.5$) of the previous instantaneous $MTTF$. If the overhaul duration is between three to five times of the previous instantaneous $MTTF$ ($3.0 \le R \le 5.0$), the more favorable applications range of PWP-GT for shape parameter is $0.8 \le \delta \le 1.2.$ As for the other two common baseline intensity model (i.e. AG and WLW), the AG model performs consistently well in the small sample size (20) regardless of the overhaul duration ($R \le 5.0$) in an HPP case. The WLW model performance improves as the overhaul duration increases ($R \ge 5.0$).

8.2.4 Covariate PI modeling

The research domains of the two factors of interests are: (1) $120 \le U \le 240$ and (2) $0.5 \le \delta \le 2.0.$ For the minor failure type, the PWP-GT proves to perform well for sample sizes 120 (60 per class) or more and decreasing, constant, and increasing ROCOFs (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 2.0$). For the major failure type, the PWP-GT performs well for sample sizes 180 (90 per class) or more and decreasing, constant, and increasing ROCOFs (power-law NHPP shape parameter in the range of $0.5 \le \delta \le 1.8$).

8.3 Contributions

The major contributions of this study are summarized as follows:

1. Robustness study of the right-censoring effect upon the four PI models in an NHPP for $0.5 \leq \delta \leq 2.0$ and ten failures per unit on 60, 120, and 180 units with censoring severity of $0.0 \leq P_c \leq 1.0$, including the development of the program to perform the robustness analysis, where the right-censoring probability is a variable. The more favorable engineering applications ranges of the right-censoring severity were recommended based on the sample size and shape parameter, including recommendations for selecting appropriate PI models in repairable system reliability.
2. Robustness study of the right-censoring effect upon the four semi-parametric PI models in an HPP for four failures per unit on 60, 120, and 180 units with censoring severity of $0.0 \leq P_c \leq 1.0$. Comparison of three event-specific PI models (PWP-GT, PWP-TT, and WLW) was presented as an indicator of selecting appropriate PI models. The more favorable engineering applications ranges of the right-censoring severity were recommended based on the sample size.
3. Development of a methodology/plan and a program for applying the discontinuous risk-free-intervals modeling to incorporate the overhaul duration following a major failure (discontinuity of observation time in system/machine downtime). The more favorable engineering applications ranges of the major duration were recommended based on the sample size and shape parameter.

4. Development of a methodology and a program on multi-dimensional covariate PI modeling for handling two distinct failure types (major and minor types), while the proportional intensities rule holds. The more favorable engineering applications ranges for the major and minor types were recommended based on the sample size and shape parameter.

8.4 Future research

8.4.1 Right-censored recurrent events

Like the power-law intensity function, the log-linear form is frequently encountered in industry. To examine the right-censoring effect upon recurring events, which follow an NHPP with log-linear intensity function, is beneficial to practitioners. Future research may evaluate how the event-specific PWP-GT model handles the recurrent data with the log-linear intensity function in terms of sample sizes, shape parameters, and censoring severity.

8.4.2 Multiple failure types

The recurring events of two failure types (major and minor failures) in this study were generated from a single NHPP stream with power-law intensity function, where the major and minor failure events share the same shape parameter (δ) of the power-law form. To meet the requirement of proportionality in the semi-parametric proportional regression method, the shape parameters of the major type (δ_1) and the minor type (δ_2) are set equal, expressed as $\delta_1 = \delta_2 = \delta$. In practice, the case $\delta_2 \neq \delta_1$ is likely. Future research may propose a model that handles (1) $\delta_1 < \delta_2$ or (2) $\delta_1 > \delta_2$.

8.4.3 Left-censored recurrent events

Left-censoring also arises in some applications for recurrent failure data from repairable systems. An example case is filed data where early life events were not recorded and records were lost. Future research could apply the methodology of Chapter 4 to examine PWP-GT robustness under left-censoring.

Appendix I (Semi-parametric proportional intensity models robustness for right-censored recurrent failure data)

I.1 Experimental units (singular type)

Table A. I. 1 Experimental units effect of the PWP-GT model ($\delta = 1.8$)

	U = 60			U = 120			U = 180		
	MSE	BIAS	MAD	MSE	BIAS	MAD	MSE	BIAS	MAD
$P_c = 0.4$	0.12377	0.15846	0.24172	0.07150	0.13317	0.16240	0.07150	0.13741	0.15393
$P_c = 0.6$	0.13549	0.14878	0.26557	0.07106	0.12388	0.15561	0.07106	0.11684	0.16411
$P_c = 0.8$	0.14486	0.12972	0.27344	0.07195	0.12368	0.15971	0.07195	0.10765	0.16802
$P_c = 1.0$	1.79361	-0.22053	0.63335	0.08326	0.16373	0.18841	0.08326	0.12651	0.19201

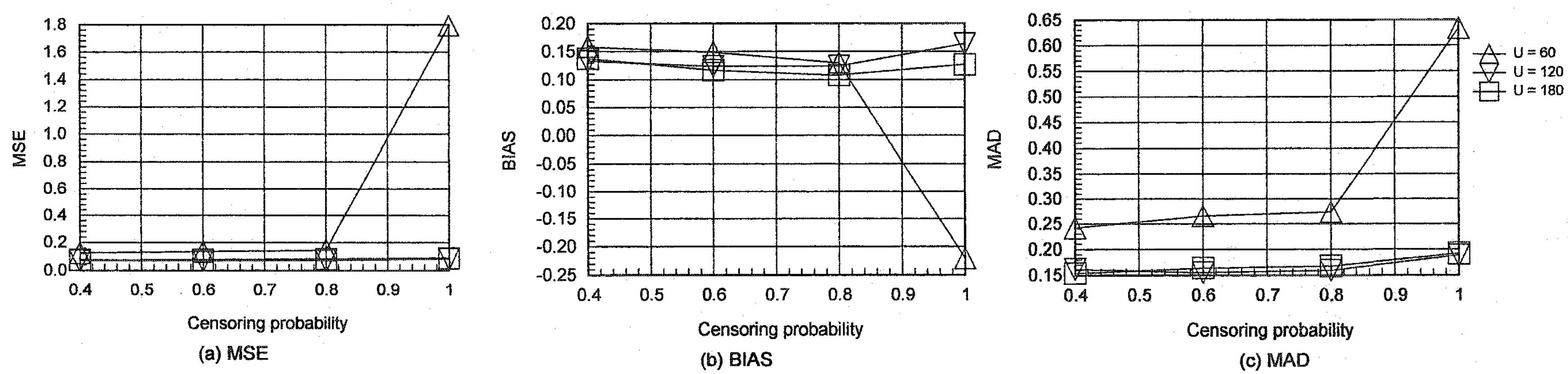


Fig A.I.1 Experimental units effect of the PWP-GT model ($\delta = 1.8$)

Table A. I. 2 Experimental units effect of the PWP-GT model ($\delta = 2.0$)

	U = 60			U = 120			U = 180		
	MSE	BIAS	MAD	MSE	BIAS	MAD	MSE	BIAS	MAD
$P_c = 0.4$	0.17465	0.17962	0.26974	0.10789	0.15956	0.16452	0.10950	0.17793	0.19784
$P_c = 0.6$	0.18666	0.16955	0.29295	0.10932	0.14829	0.18587	0.11578	0.14400	0.19210
$P_c = 0.8$	0.19294	0.15996	0.29812	0.11142	0.14281	0.18920	0.12272	0.12980	0.19181
$P_c = 1.0$	2.14182	-0.21764	0.71010	0.12494	0.18721	0.22697	0.12772	0.14455	0.21804

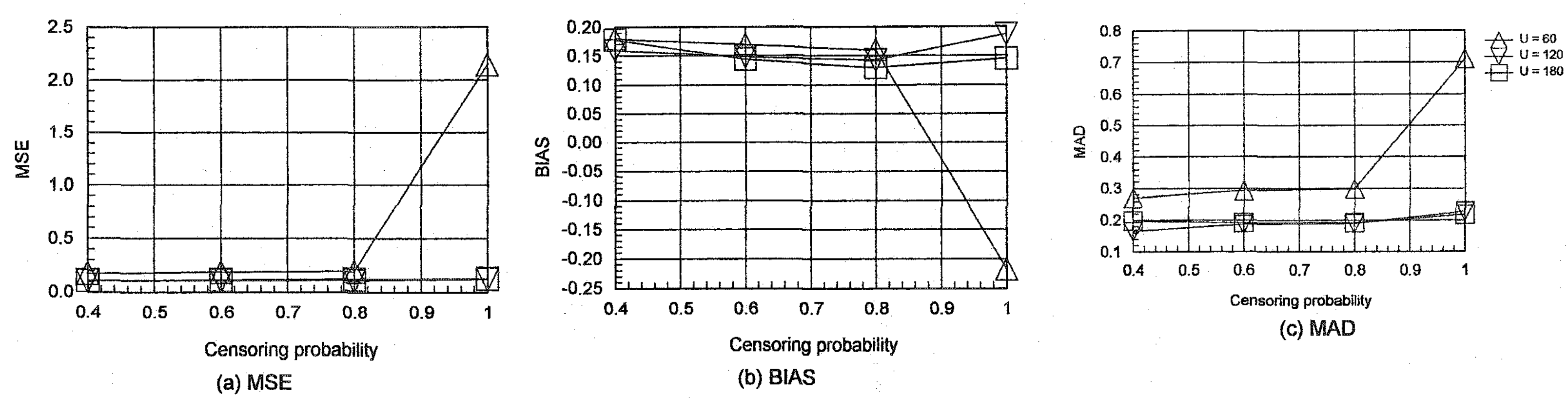
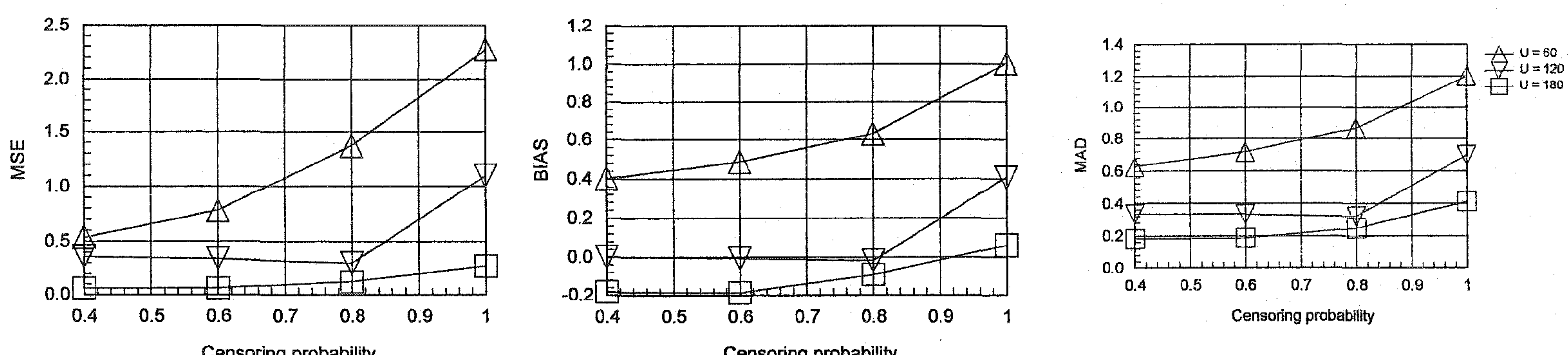


Fig A.I.2 Experimental units effect of the PWP-GT model ($\delta = 2.0$)

Table A.I.3 Experimental units effect of the PWP-GT model ($\delta = 0.5$)

	U = 60			U = 120			U = 180		
	MSE	BIAS	MAD	MSE	BIAS	MAD	MSE	BIAS	MAD
$P_c = 0.4$	0.53551	0.40407	0.62746	0.36198	0.00465	0.33390	0.06075	-0.18178	0.18178
$P_c = 0.6$	0.77886	0.48680	0.71696	0.34196	-0.00555	0.33308	0.06217	-0.18696	0.18696
$P_c = 0.8$	1.36832	0.62978	0.86481	0.30120	-0.01706	0.31954	0.12328	-0.09188	0.24669
$P_c = 1.0$	2.27255	1.00221	1.20024	1.09745	0.40857	0.69729	0.27329	0.06185	0.40939

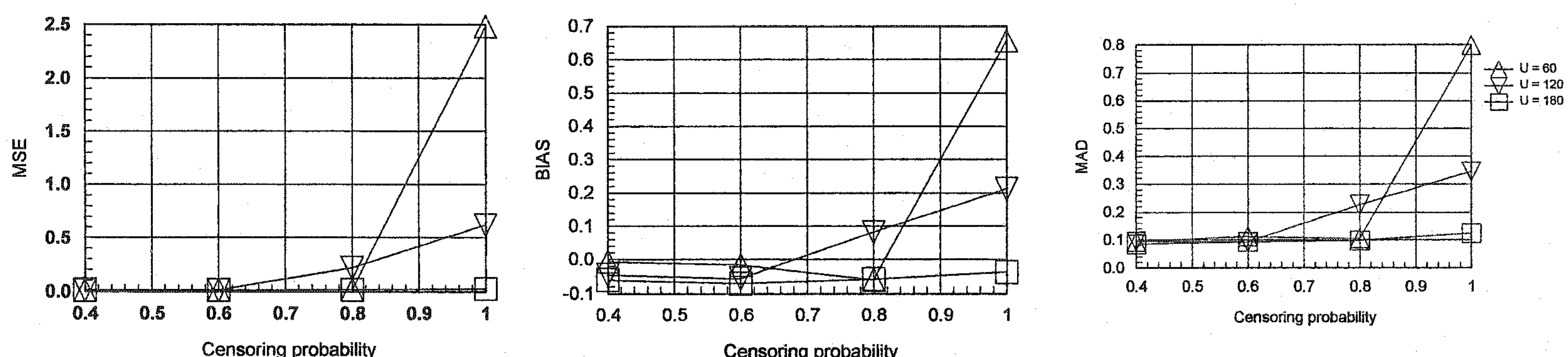


(a) MSE (b) BIAS (c) MAD

Fig A.I.3 Experimental units effect of the PWP-GT model ($\delta = 0.5$)

Table A.I.4 Experimental units effect of the PWP-GT model ($\delta = 0.8$)

	U = 60			U = 120			U = 180		
	MSE	BIAS	MAD	MSE	BIAS	MAD	MSE	BIAS	MAD
$P_c = 0.4$	0.01291	-0.00612	0.09325	0.01563	-0.04467	0.09571	0.01140	-0.05950	0.08415
$P_c = 0.6$	0.01670	-0.01597	0.11158	0.01477	-0.05590	0.09200	0.01306	-0.07045	0.09304
$P_c = 0.8$	0.01871	-0.05920	0.10437	0.21895	0.08223	0.22646	0.01753	-0.05754	0.09754
$P_c = 1.0$	2.47916	0.65585	0.79796	0.62236	0.21174	0.34510	0.02774	-0.03674	0.12327



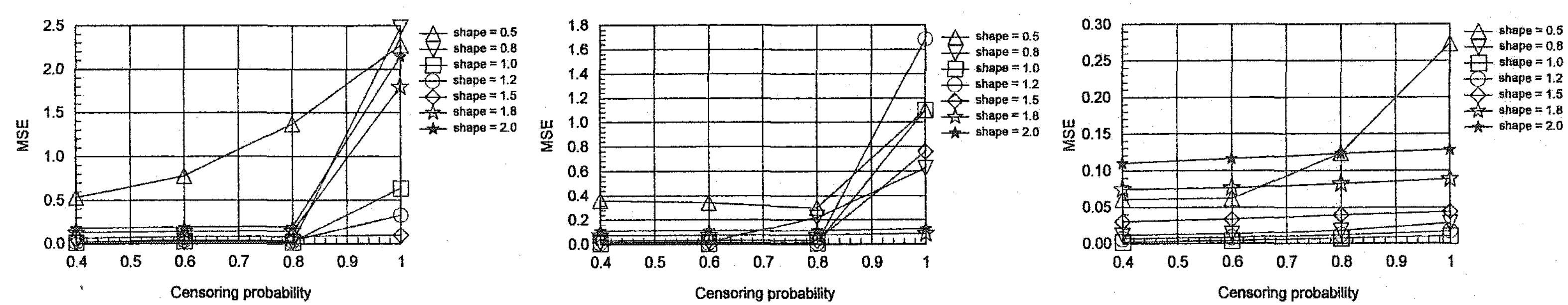
(a) MSE (b) BIAS (c) MAD

Fig A.I.4 Experimental units effect of the PWP-GT model ($\delta = 0.8$)

I.2 Shape parameter (singular type)

Table A.I.5 Shape parameter effect of the PWP-GT model (MSE)

	P_c	$\delta=0.5$	$\delta=0.8$	$\delta=1.0$	$\delta=1.2$	$\delta=1.5$	$\delta=1.8$	$\delta=2.0$
U=60	0.4	0.53551	0.01291	0.01143	0.01989	0.06218	0.12377	0.17465
	0.6	0.77886	0.01670	0.02750	0.03993	0.07480	0.13549	0.18666
	0.8	1.36832	0.01871	0.01590	0.03159	0.08060	0.14486	0.19294
	1.0	2.27255	2.47916	0.63632	0.32410	0.09508	1.79361	2.14182
U=120	0.4	0.36198	0.01563	0.00590	0.00864	0.02970	0.0715	0.10789
	0.6	0.34196	0.01477	0.00455	0.00811	0.02931	0.07106	0.10932
	0.8	0.30120	0.21895	0.00923	0.00894	0.03006	0.07195	0.11142
	1.0	1.09745	0.62236	1.09905	1.68430	0.75671	0.08326	0.12494
U=180	0.4	0.06075	0.01140	0.00195	0.00615	0.03006	0.07494	0.10950
	0.6	0.06217	0.01306	0.00361	0.00806	0.03304	0.07669	0.11578
	0.8	0.12328	0.01753	0.00719	0.01176	0.03924	0.08241	0.12272
	1.0	0.27329	0.02774	0.00984	0.01642	0.04380	0.08858	0.12772

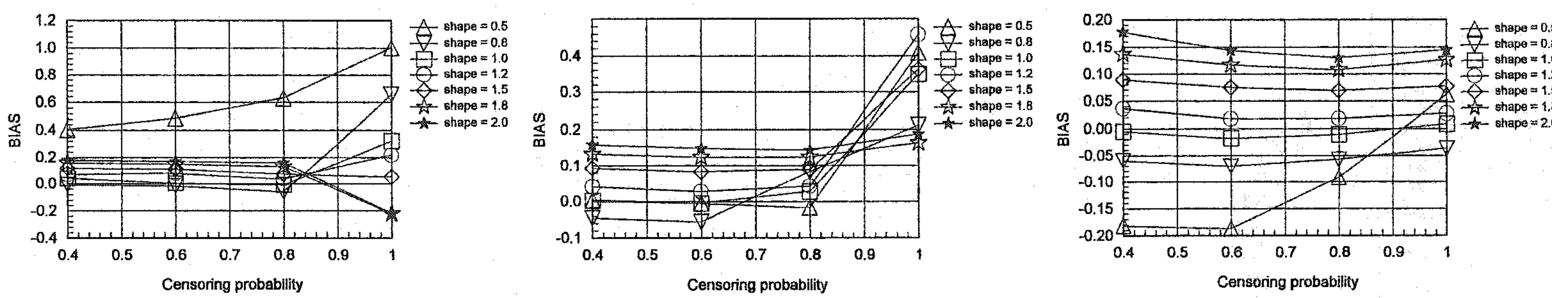


(a) U=60 (b) U=120 (c) U=180

Fig A.I.5 Shape parameter effect of the PWP-GT model (MSE)

Table A.I.6 Shape parameter effect of the PWP-GT model (BIAS)

	P_c	$\delta = 0.5$	$\delta = 0.8$	$\delta = 1.0$	$\delta = 1.2$	$\delta = 1.5$	$\delta = 1.8$	$\delta = 2.0$
U=60	0.4	0.40407	-0.00612	0.04178	0.06894	0.11742	0.15846	0.17962
	0.6	0.48680	-0.01597	0.00454	0.07725	0.11172	0.14878	0.16955
	0.8	0.62978	-0.05920	-0.00877	0.03604	0.07272	0.12972	0.15996
	1.0	1.00221	0.65585	0.32151	0.21679	0.05251	-0.22053	-0.21764
U=120	0.4	0.00465	-0.04467	0.00232	0.04091	0.09289	0.13317	0.15956
	0.6	-0.00555	-0.05590	-0.00506	0.02710	0.08247	0.12388	0.14829
	0.8	-0.01706	0.08223	0.02708	0.04122	0.08892	0.12368	0.14281
	1.0	0.40857	0.21174	0.34912	0.45836	0.36274	0.16373	0.18721
U=180	0.4	-0.18178	-0.05950	-0.00516	0.03653	0.08892	0.13741	0.17793
	0.6	-0.18696	-0.07045	-0.01788	0.01795	0.07461	0.11684	0.14400
	0.8	-0.09188	-0.05754	-0.01083	0.01806	0.06877	0.10765	0.12980
	1.0	0.06185	-0.03674	0.00855	0.02747	0.07652	0.12651	0.14455

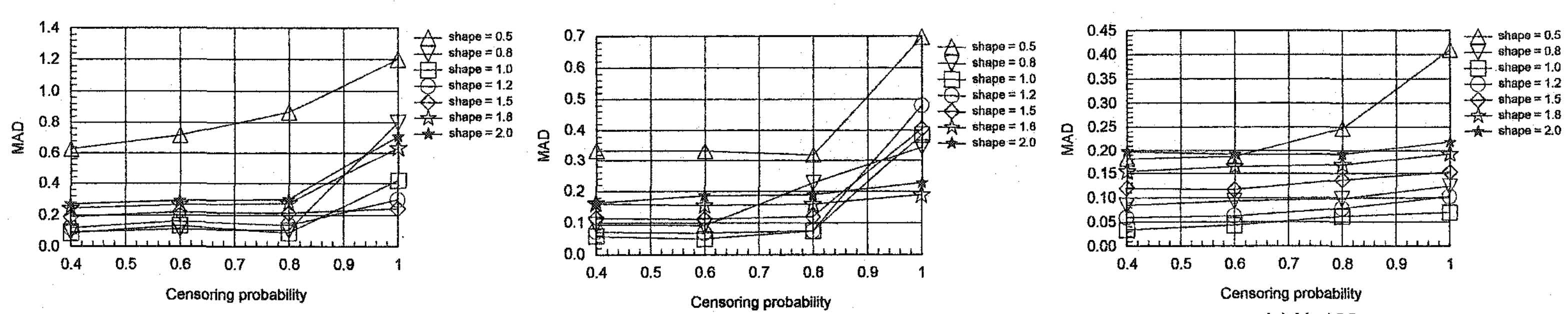


(a) U=60 (b) U=120 (c) U=180

Fig A.I.6 Shape parameter effect of the PWP-GT model (BIAS)

Table A.I.7 Shape parameter effect of the PWP-GT model (MAD)

	P_c	$\delta = 0.5$	$\delta = 0.8$	$\delta = 1.0$	$\delta = 1.2$	$\delta = 1.5$	$\delta = 1.8$	$\delta = 2.0$
U=60	0.4	0.62746	0.09325	0.08594	0.11685	0.19005	0.24172	0.26974
	0.6	0.71696	0.11158	0.13398	0.16187	0.22042	0.26557	0.29295
	0.8	0.86481	0.10437	0.08845	0.13645	0.21307	0.27344	0.29812
	1.0	1.20024	0.79796	0.41818	0.29794	0.24110	0.63335	0.71010
U=120	0.4	0.33390	0.09571	0.05852	0.07307	0.11596	0.1624	0.16452
	0.6	0.33308	0.09200	0.04979	0.06802	0.11322	0.15561	0.18587
	0.8	0.31954	0.22646	0.07464	0.07453	0.11935	0.15971	0.18920
	1.0	0.69729	0.34510	0.38596	0.48052	0.40576	0.18841	0.22697
U=180	0.4	0.18178	0.08415	0.03339	0.05888	0.11935	0.15393	0.19784
	0.6	0.18696	0.09304	0.04288	0.06244	0.11632	0.16411	0.19210
	0.8	0.24669	0.09754	0.06113	0.07786	0.13615	0.16802	0.19181
	1.0	0.40939	0.12327	0.06957	0.10220	0.15275	0.19201	0.21804



(a) U=60 (b) U=120 (c) U=180

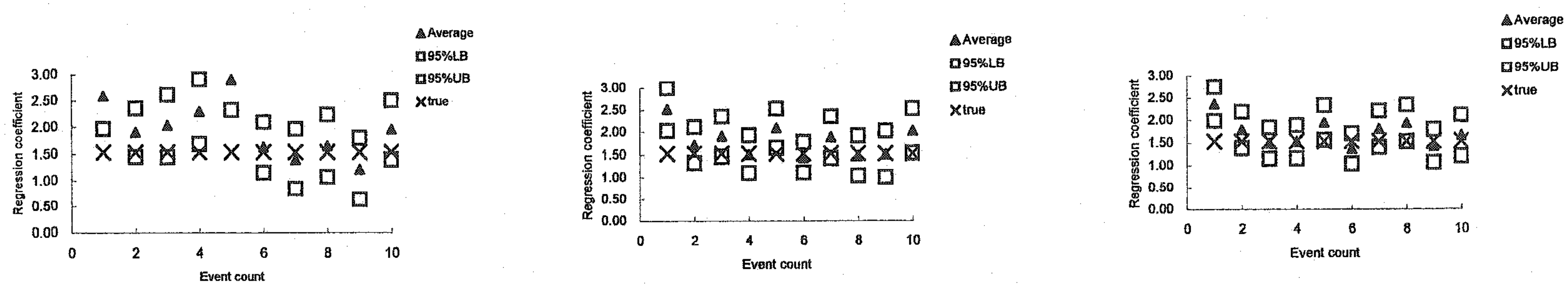
Fig A.I.7 Shape parameter effect of the PWP-GT model (MAD)

I.3 95% C.I. charts (major and minor types)

Table A.I.8 PWP-GT estimates and 95% C.I., $\delta = 1.5$

	n	Major events				Minor events				
		Average	s.d.[a]	95%LB	95%UB	Average	s.d.[a]	95%LB	95%UB	True
U=120	1	2.58752	0.31974	1.96084	3.21420	2.24535	0.16561	1.92076	2.56993	1.53506
	2	1.90205	0.23166	1.44800	2.35610	1.85753	0.14640	1.57060	2.14446	1.53506
	3	2.02619	0.29791	1.44229	2.61009	1.82541	0.14304	1.54505	2.10577	1.53506
	4	2.29329	0.31280	1.68021	2.90638	1.53341	0.13538	1.26807	1.79874	1.53506
	5	2.89702	0.29562	2.31761	3.47642	1.93824	0.15197	1.64040	2.23609	1.53506
	6	1.61944	0.23976	1.14951	2.08936	1.25598	0.12700	1.00706	1.50490	1.53506
	7	1.39773	0.28355	0.84198	1.95349	1.53782	0.12929	1.28443	1.79122	1.53506
	8	1.64866	0.29580	1.06891	2.22841	1.48733	0.13079	1.23099	1.74367	1.53506
	9	1.21178	0.29575	0.63213	1.79144	1.31116	0.12999	1.05638	1.56594	1.53506
	10	1.94658	0.28336	1.39121	2.50195	1.46138	0.12975	1.20708	1.71568	1.53506
U=180	1	2.51532	0.24797	2.02931	3.00133	2.53602	0.14466	2.25249	2.81955	1.53506
	2	1.71797	0.20008	1.32581	2.11012	1.69407	0.11464	1.46937	1.91877	1.53506
	3	1.90476	0.22730	1.45926	2.35026	1.75554	0.11554	1.52910	1.98199	1.53506
	4	1.51389	0.21023	1.10183	1.92594	1.57951	0.11145	1.36107	1.79795	1.53506
	5	2.08875	0.22428	1.64916	2.52833	1.67383	0.11360	1.45118	1.89648	1.53506
	6	1.43896	0.17146	1.10291	1.77501	1.23985	0.10491	1.03422	1.44547	1.53506
	7	1.88826	0.23571	1.42627	2.35025	1.54954	0.11029	1.33337	1.76571	1.53506
	8	1.47523	0.22286	1.03843	1.91202	1.69800	0.11860	1.46555	1.93045	1.53506
	9	1.51054	0.26039	1.00018	2.02090	1.26669	0.10447	1.06194	1.47145	1.53506
	10	2.03049	0.25222	1.53615	2.52484	1.58336	0.11146	1.36490	1.80182	1.53506
U=240	1	2.36312	0.19532	1.98029	2.74595	2.43992	0.12149	2.20180	2.67803	1.53506
	2	1.78564	0.20421	1.38541	2.18587	1.73332	0.09858	1.54011	1.92652	1.53506
	3	1.49069	0.17422	1.14923	1.83215	1.69289	0.10036	1.49619	1.88960	1.53506
	4	1.52187	0.18601	1.15730	1.88644	1.77231	0.09649	1.58319	1.96142	1.53506
	5	1.93965	0.19773	1.55211	2.32720	1.57420	0.09510	1.38780	1.76059	1.53506
	6	1.36698	0.16779	1.03811	1.69585	1.25973	0.09168	1.08005	1.43941	1.53506
	7	1.80017	0.20638	1.39567	2.20467	1.49603	0.09292	1.31391	1.67815	1.53506
	8	1.92484	0.20929	1.51464	2.33504	1.60654	0.09818	1.41410	1.79898	1.53506
	9	1.42738	0.18411	1.06654	1.78822	1.24945	0.09082	1.07145	1.42745	1.53506
	10	1.65260	0.23046	1.20090	2.10430	1.60690	0.09976	1.41137	1.80243	1.53506

[a] s.d. is derived from the composite variance (seed numbers: 539, 255, and 59)

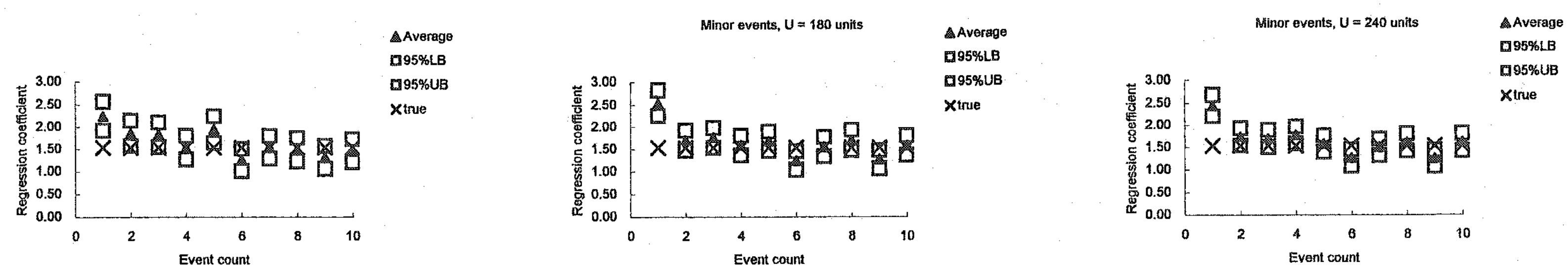


(a) U=120

(b) U=180

(c) U=240

Fig A.I.8 PWP-GT model on major events ($\delta = 1.5$)



(a) U=120

(b) U=180

(c) U=240

Fig A.I.9 PWP-GT model on minor events ($\delta = 1.5$)

Table A.I.9 PWP-GT estimates and 95% C.I., $\delta = 1.0$

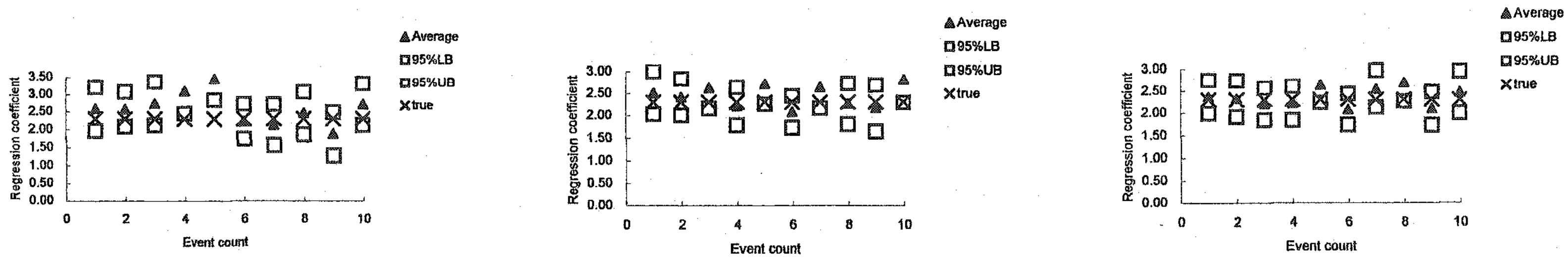
	n	Major events				Minor events				
		Average	s.d.[a]	95%LB	95%UB	Average	s.d.[a]	95%LB	95%UB	True
U=120	1	2.58752	0.31974	1.96084	3.21420	2.24535	0.16561	1.92076	2.56993	2.30259
	2	2.58464	0.25241	2.08993	3.07935	2.50245	0.17513	2.15920	2.84571	2.30259
	3	2.74223	0.31661	2.12170	3.36277	2.50695	0.17528	2.16341	2.85049	2.30259
	4	3.09574	0.32978	2.44939	3.74209	2.30248	0.16703	1.97511	2.62985	2.30259
	5	3.45394	0.31333	2.83982	4.06805	2.65093	0.18891	2.28067	3.02118	2.30259
	6	2.23321	0.25092	1.74141	2.72501	1.92020	0.14533	1.63536	2.20505	2.30259
	7	2.13522	0.29320	1.56056	2.70988	2.27701	0.15346	1.97623	2.57778	2.30259
	8	2.46468	0.31123	1.85469	3.07467	2.25997	0.15901	1.94831	2.57162	2.30259
	9	1.88085	0.30619	1.28073	2.48096	1.99446	0.15349	1.69362	2.29529	2.30259
	10	2.71827	0.30172	2.12691	3.30963	2.24122	0.16225	1.92322	2.55923	2.30259
U=180	1	2.51532	0.24798	2.02929	3.00135	2.53602	0.14466	2.25249	2.81955	2.30259
	2	2.41438	0.21360	1.99574	2.83303	2.31893	0.13440	2.05551	2.58234	2.30259
	3	2.62405	0.24152	2.15068	3.09742	2.45064	0.14566	2.16516	2.73613	2.30259
	4	2.20970	0.22113	1.77629	2.64310	2.29765	0.12312	2.05635	2.53895	2.30259
	5	2.71702	0.23688	2.25275	3.18129	2.32770	0.13338	2.06628	2.58912	2.30259
	6	2.08090	0.18228	1.72364	2.43817	1.90136	0.12022	1.66573	2.13699	2.30259
	7	2.64872	0.24976	2.15920	3.13823	2.37139	0.13480	2.10719	2.63560	2.30259
	8	2.26270	0.23660	1.79898	2.72641	2.43710	0.13627	2.17003	2.70418	2.30259
	9	2.16391	0.26789	1.63885	2.68896	1.92413	0.12037	1.68821	2.16006	2.30259
	10	2.81512	0.26840	2.28907	3.34118	2.35544	0.13998	2.08110	2.62979	2.30259
U=240	1	2.36312	0.19532	1.98029	2.74595	2.43992	0.12149	2.20180	2.67803	2.30259
	2	2.32677	0.21163	1.91198	2.74155	2.36321	0.11530	2.13722	2.58919	2.30259
	3	2.20676	0.18589	1.84243	2.57110	2.41187	0.12072	2.17526	2.64848	2.30259
	4	2.21981	0.19645	1.83478	2.60485	2.31046	0.11613	2.08285	2.53807	2.30259
	5	2.64113	0.20824	2.23298	3.04927	2.24010	0.11235	2.01989	2.46030	2.30259
	6	2.07843	0.17766	1.73023	2.42663	2.04818	0.10678	1.83890	2.25746	2.30259
	7	2.54038	0.21650	2.11604	2.96472	2.31013	0.11192	2.09077	2.52948	2.30259
	8	2.69026	0.22104	2.25702	3.12350	2.44848	0.11945	2.21437	2.68259	2.30259
	9	2.09593	0.19231	1.71901	2.47285	1.99932	0.10549	1.79256	2.20608	2.30259
	10	2.46961	0.24314	1.99306	2.94616	2.44808	0.12546	2.20217	2.69398	2.30259

[a] s.d. is derived from the composite variance (seed numbers: 539, 255, and 59)

Table A.I.10 PWP-GT estimates and 95% C.I., $\delta = 0.8$

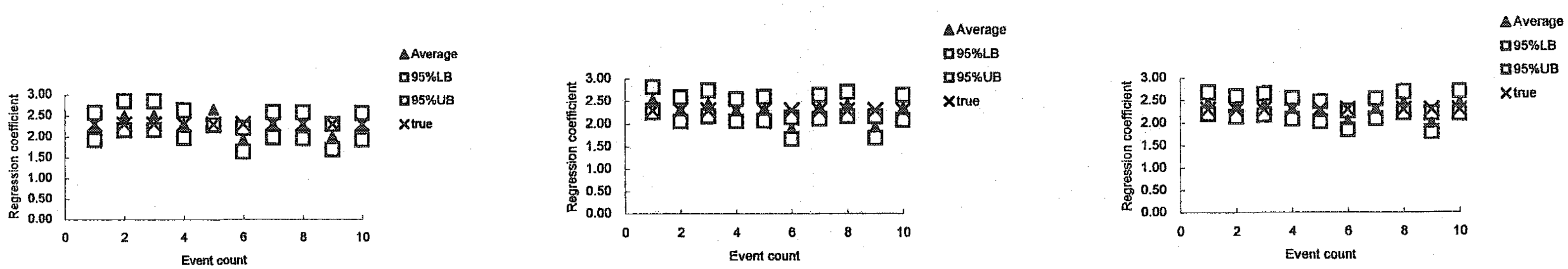
	n	Major events				Minor events				
		Average	s.d.[a]	95%LB	95%UB	Average	s.d.[a]	95%LB	95%UB	True
U=120	1	2.58752	0.31974	1.96084	3.21420	2.24535	0.16561	1.92076	2.56993	2.87823
	2	2.84789	0.26201	2.33435	3.36143	2.79625	0.19079	2.42230	3.17020	2.87823
	3	3.12894	0.33221	2.47782	3.78006	2.90549	0.19988	2.51374	3.29724	2.87823
	4	3.58258	0.34151	2.91322	4.25193	2.77956	0.18626	2.41450	3.14462	2.87823
	5	3.86098	0.33289	3.20852	4.51344	3.13797	0.21830	2.71011	3.56584	2.87823
	6	2.70284	0.26366	2.18608	3.21960	2.40634	0.16431	2.08429	2.72839	2.87823
	7	2.66441	0.30632	2.06403	3.26479	2.78499	0.17670	2.43867	3.13131	2.87823
	8	3.09079	0.33590	2.43244	3.74913	2.89694	0.20185	2.50133	3.29255	2.87823
	9	2.39306	0.31600	1.77371	3.01242	2.47120	0.17246	2.13319	2.80921	2.87823
	10	3.14562	0.31639	2.52550	3.76574	2.68603	0.18655	2.32040	3.05166	2.87823
U=180	1	2.51532	0.24798	2.02929	3.00135	2.53602	0.14466	2.25249	2.81955	2.87823
	2	2.75430	0.22113	2.32089	3.18772	2.62193	0.14559	2.33659	2.90727	2.87823
	3	3.02777	0.25258	2.53272	3.52281	2.84671	0.15608	2.54080	3.15262	2.87823
	4	2.66767	0.23099	2.21495	3.12039	2.72492	0.14967	2.43158	3.01826	2.87823
	5	3.14575	0.24939	2.65695	3.63455	2.78261	0.15236	2.48399	3.08122	2.87823
	6	2.54858	0.19372	2.16889	2.92827	2.37786	0.13645	2.11042	2.64530	2.87823
	7	3.14738	0.26536	2.62728	3.66747	2.95984	0.16282	2.64072	3.27895	2.87823
	8	2.80749	0.25216	2.31327	3.30170	2.96837	0.16315	2.64861	3.28813	2.87823
	9	2.64770	0.27587	2.10700	3.18841	2.39486	0.13567	2.12895	2.66078	2.87823
	10	3.31396	0.28218	2.76090	3.86702	2.83575	0.16241	2.51743	3.15406	2.87823
U=240	1	2.36312	0.19532	1.98029	2.74595	2.43992	0.12149	2.20180	2.67803	2.87823
	2	2.61543	0.21680	2.19052	3.04034	2.68586	0.12663	2.43767	2.93405	2.87823
	3	2.61614	0.19530	2.23336	2.99892	2.81260	0.13500	2.54801	3.07719	2.87823
	4	2.66876	0.20585	2.26530	3.07223	2.76388	0.13190	2.50536	3.02240	2.87823
	5	3.03030	0.21580	2.60735	3.45326	2.68327	0.12696	2.43443	2.93210	2.87823
	6	2.56637	0.18730	2.19926	2.93348	2.52565	0.12123	2.28805	2.76325	2.87823
	7	3.11057	0.22812	2.66346	3.55767	2.88848	0.13197	2.62982	3.14713	2.87823
	8	3.24040	0.23345	2.78285	3.69795	2.99919	0.14216	2.72057	3.27781	2.87823
	9	2.55987	0.20048	2.16694	2.95280	2.47484	0.11985	2.23994	2.70974	2.87823
	10	2.97263	0.25594	2.47101	3.47425	2.93079	0.14721	2.64226	3.21932	2.87823

[a] s.d. is derived from the composite variance (seed numbers: 539, 255, and 59)



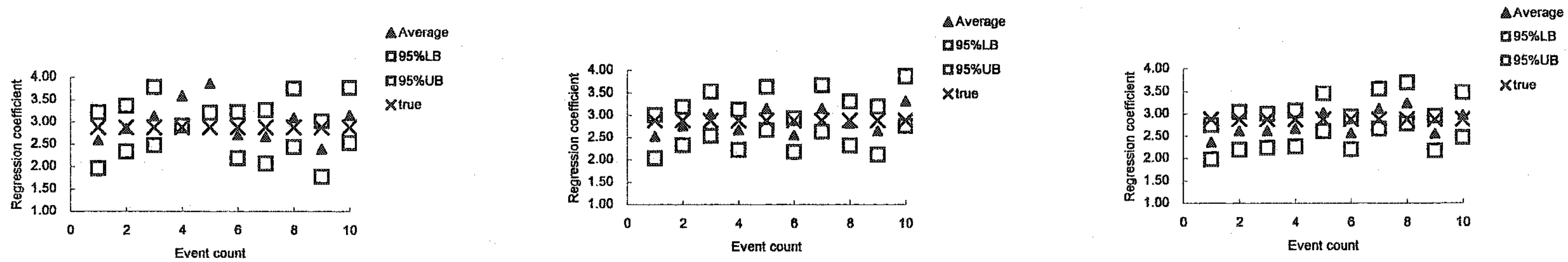
(a) U=120 (b) U=180 (c) U=240

Fig A.I.10 PWP-GT model on major events ($\delta = 1.0$)



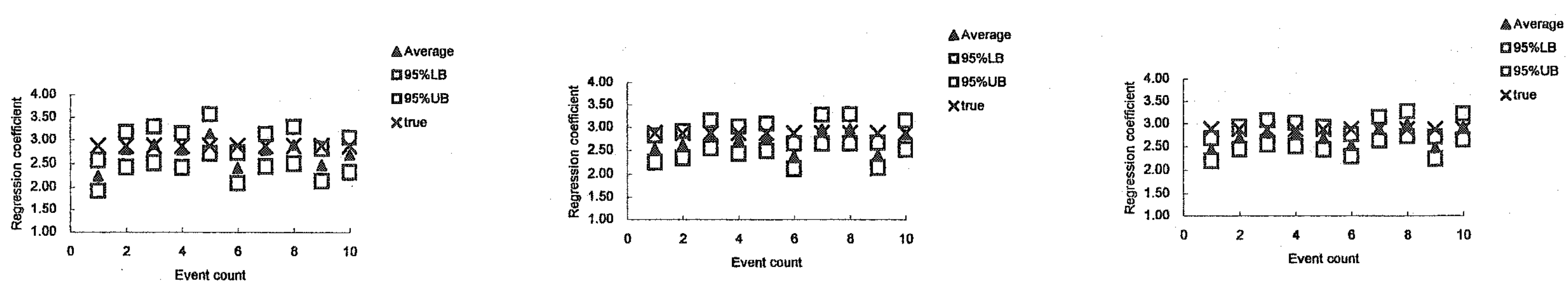
(a) U=120 (b) U=180 (c) U=240

Fig A.I.11 PWP-GT model on minor events ($\delta = 1.0$)



(a) U=120 (b) U=180 (c) U=240

Fig A.I.12 PWP-GT model on major events ($\delta = 0.8$)



(a) U=120 (b) U=180 (c) U=240

Fig A.I.13 PWP-GT model on minor events ($\delta = 0.8$)

I.4 Experimental units (major and minor types)

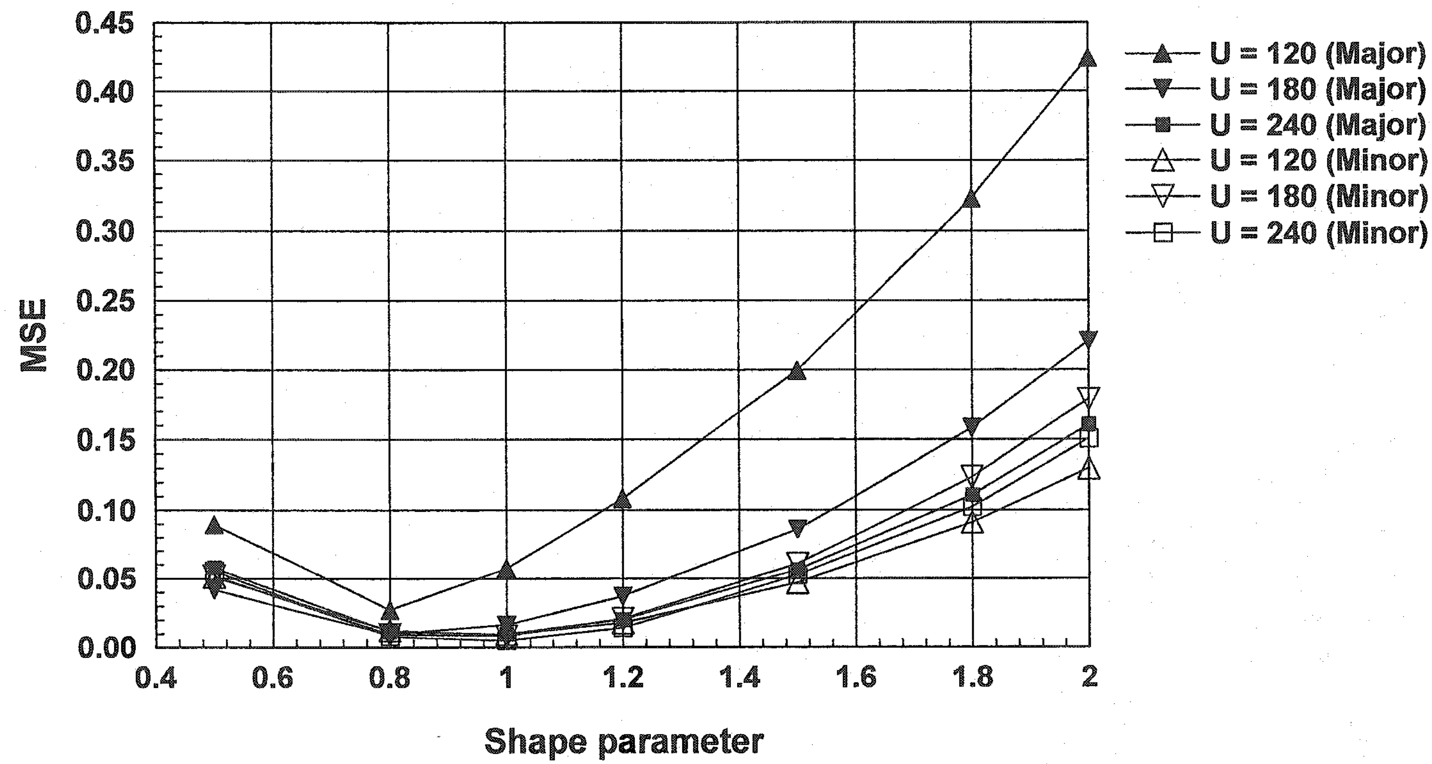


Fig A.I.14 Experimental units effect of the PWP-GT model (MSE)

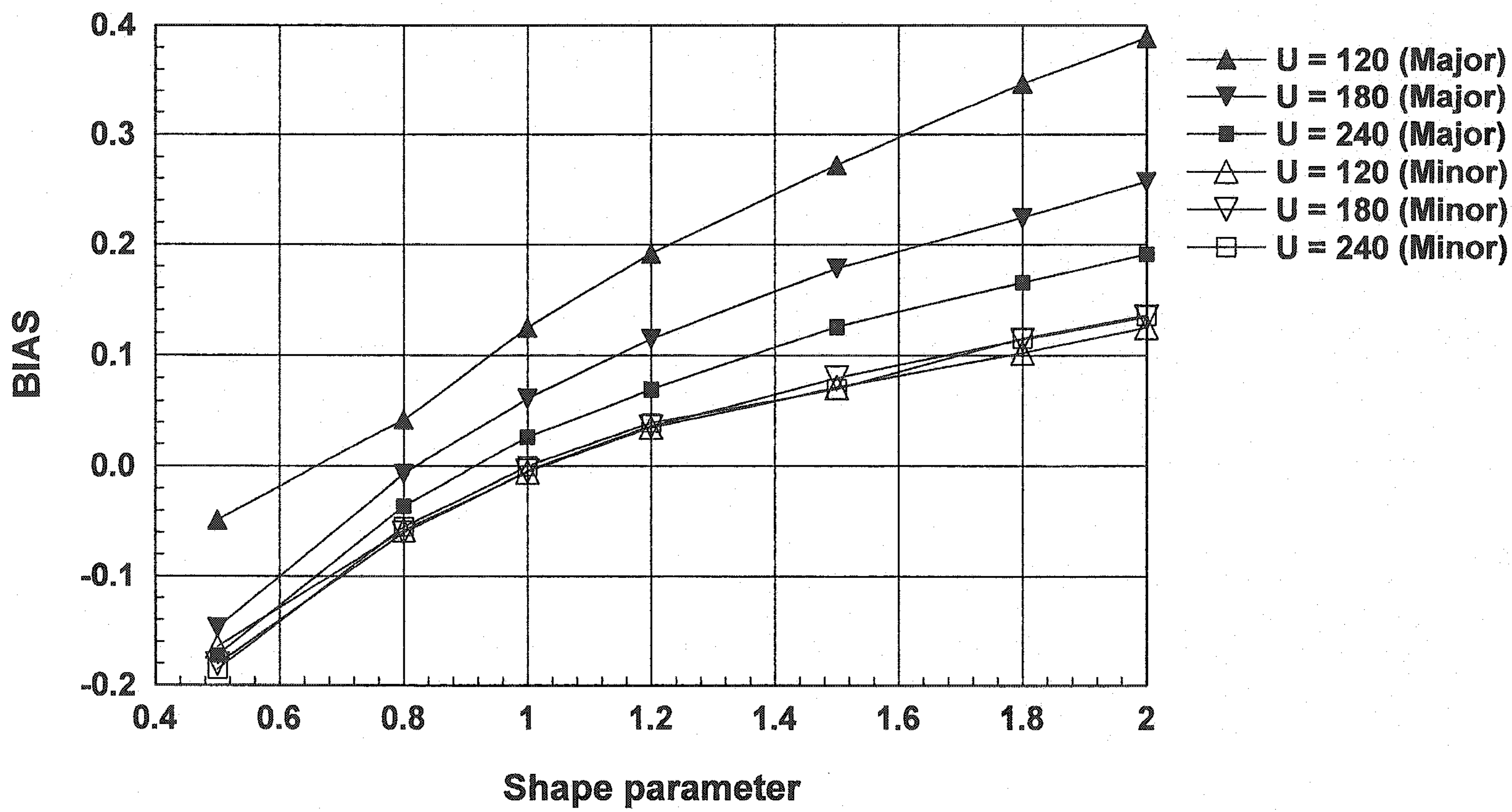


Fig A.I.15 Experimental units effect of the PWP-GT model (BIAS)

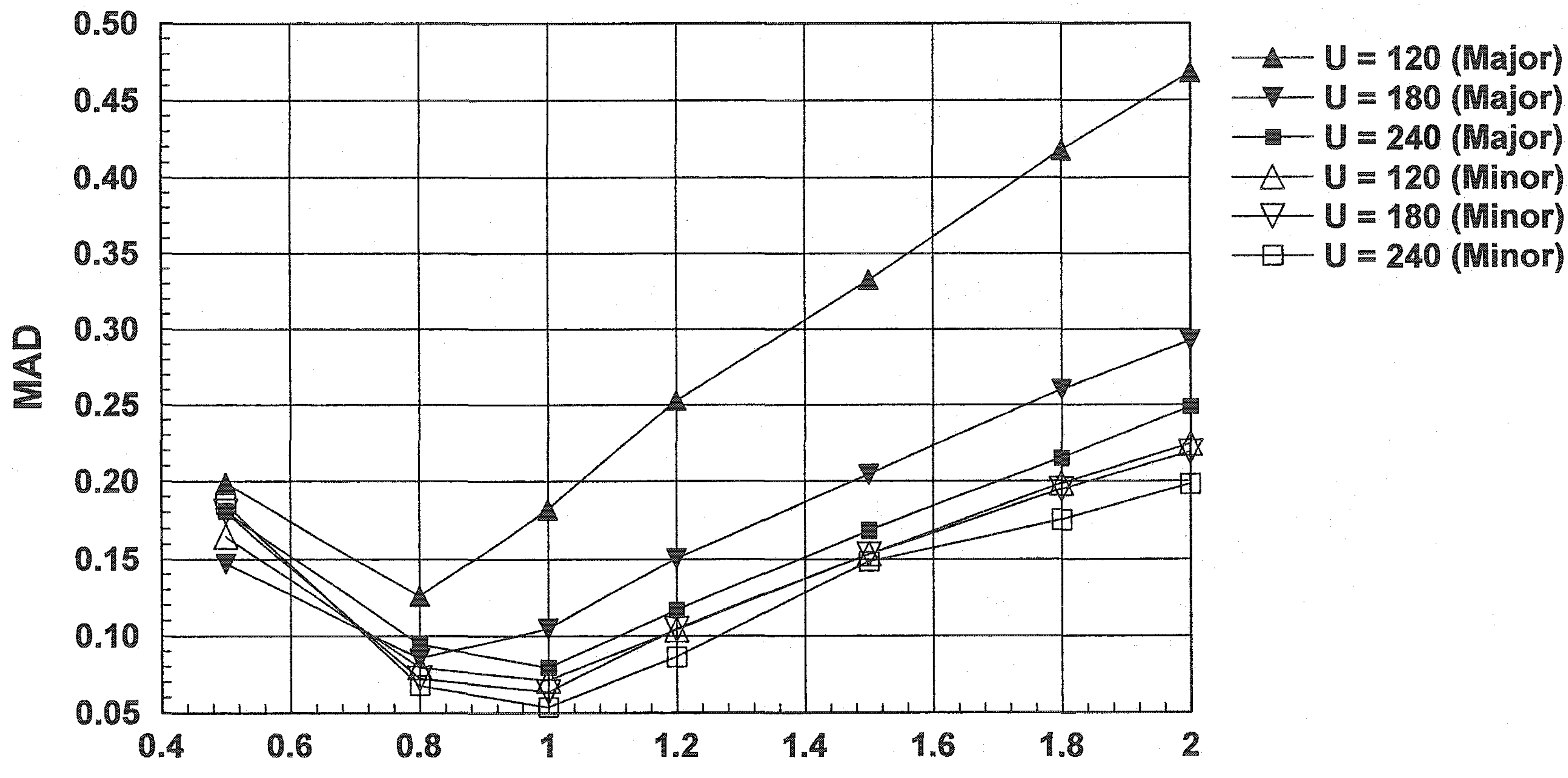
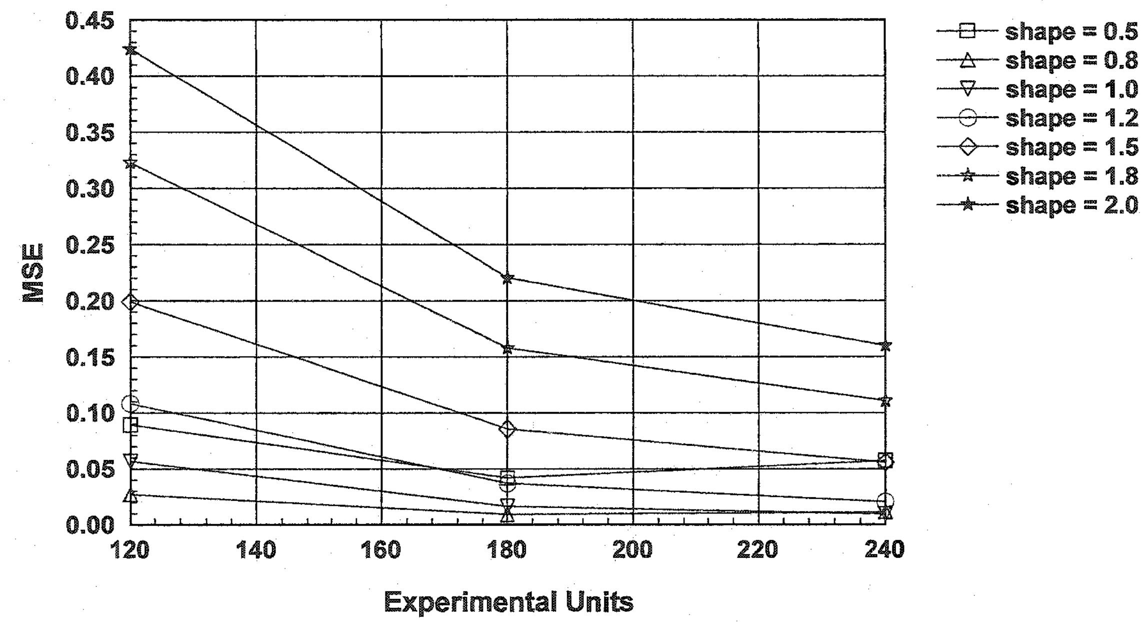


Fig A.I.16 Experimental units effect of the PWP-GT model (MAD)

I.5 Shape parameter (major and minor types)



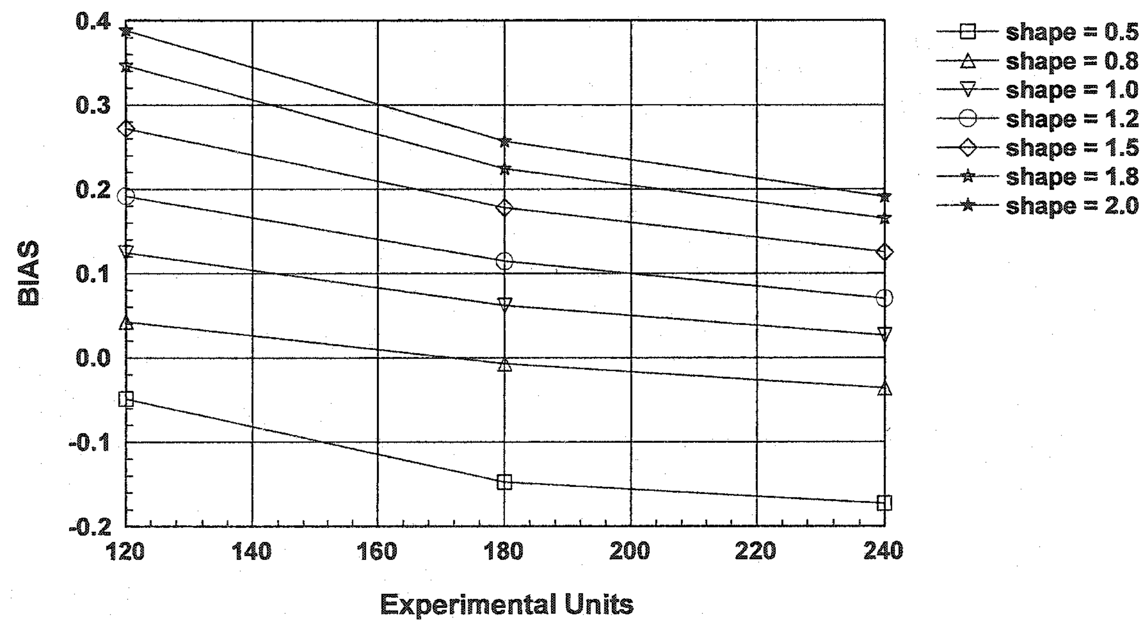


Fig A.I.17 Shape parameter effect of the PWP-GT model on major events

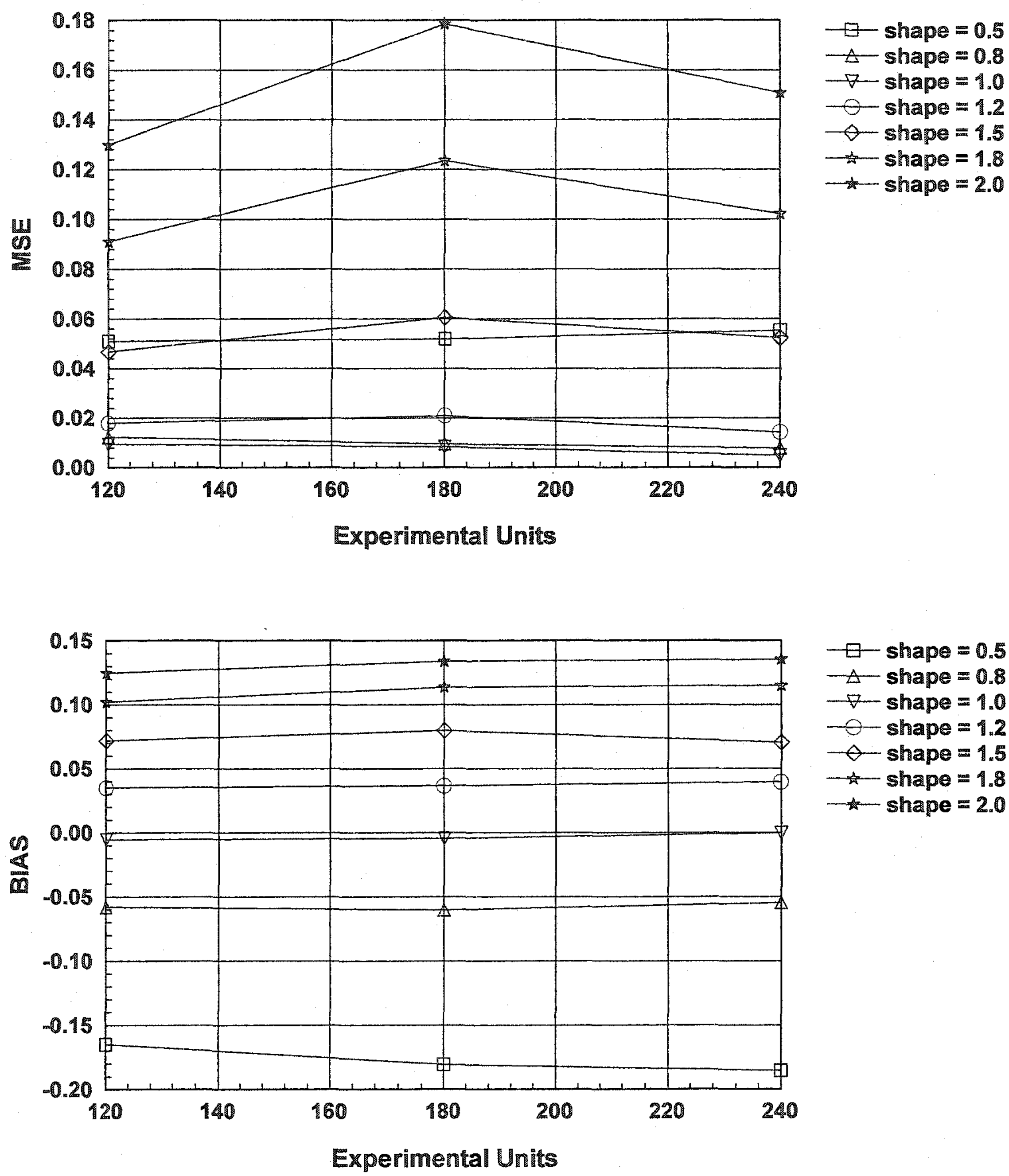
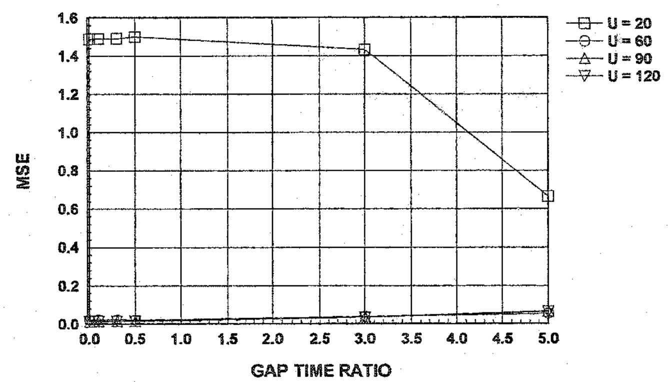


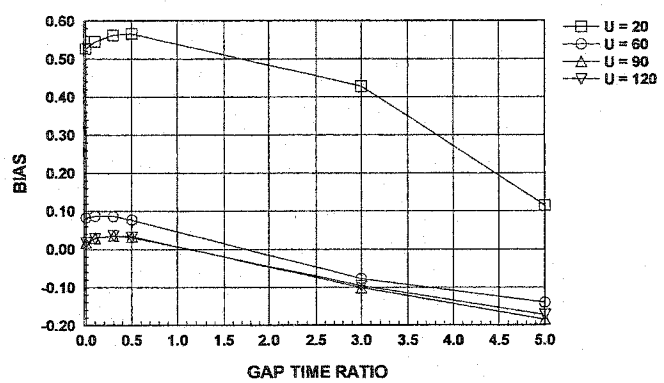
Fig A.I.18 Shape parameter effect of the PWP-GT model on minor events

Appendix II (Semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals)

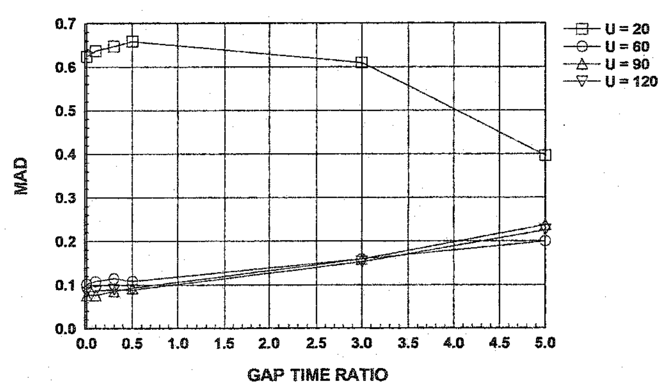
II.1 Experimental units



(a) MSE

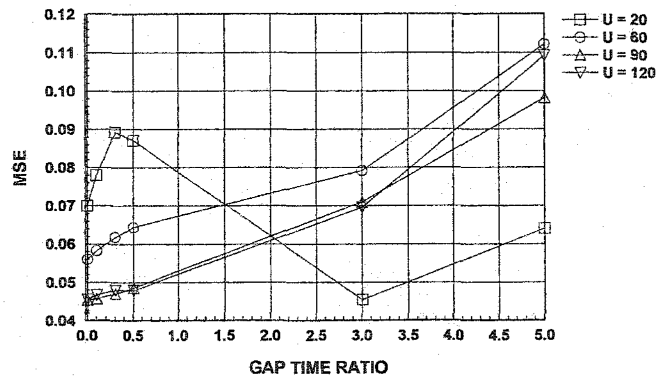


(b) BIAS

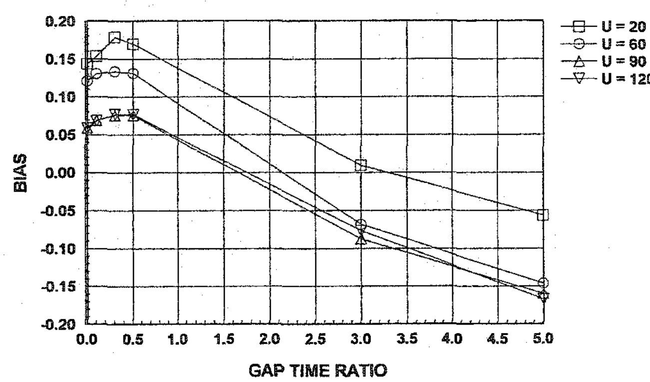


(c) MAD

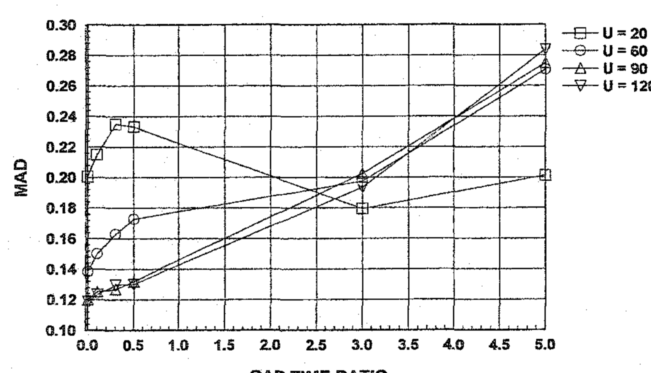
Fig A.II.1 Experimental units effect of the PWP-GT model ($\delta = 1.2$)



(a) MSE

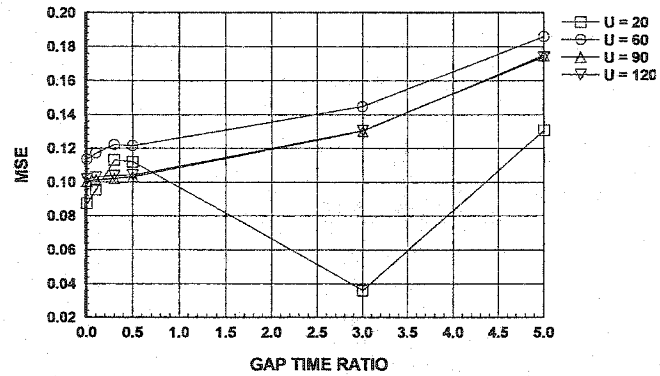


(b) BIAS

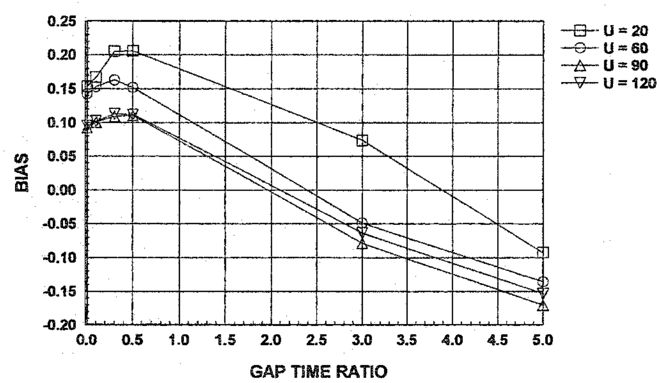


(c) MAD

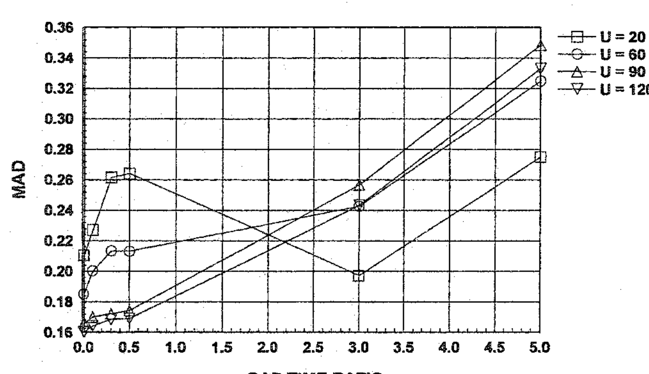
Fig A.II.2 Experimental units effect of the PWP-GT model ($\delta = 1.5$)



(a) MSE

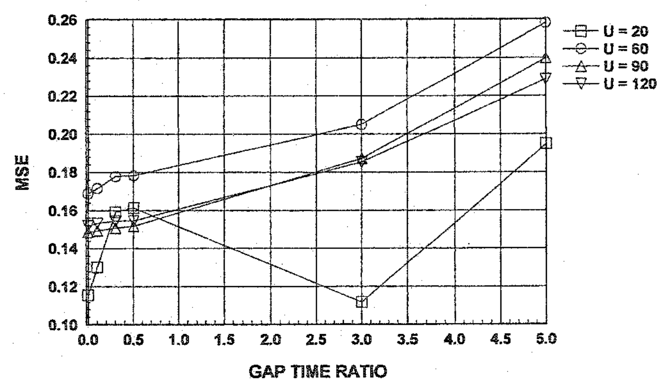


(b) BIAS

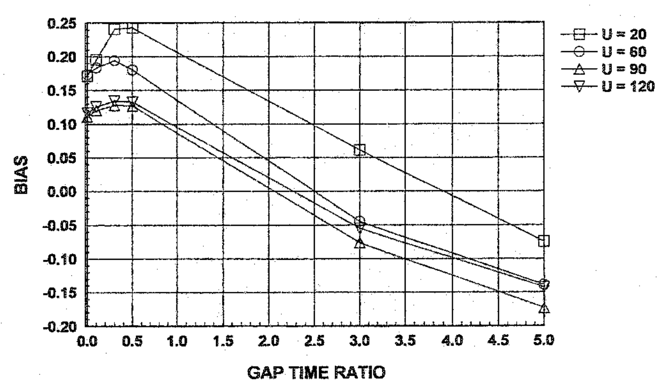


(c) MAD

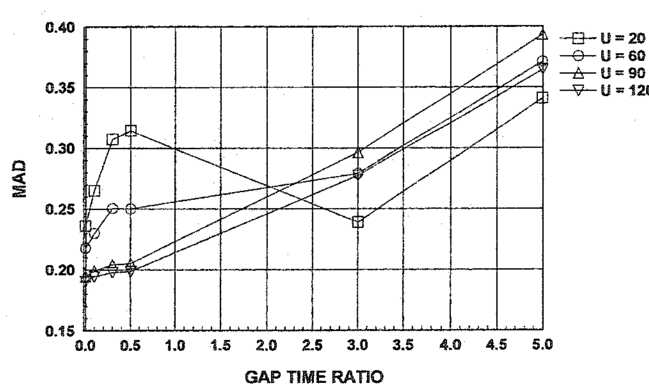
Fig A.II.3 Experimental units effect of the PWP-GT model ($\delta = 1.8$)



(a) MSE

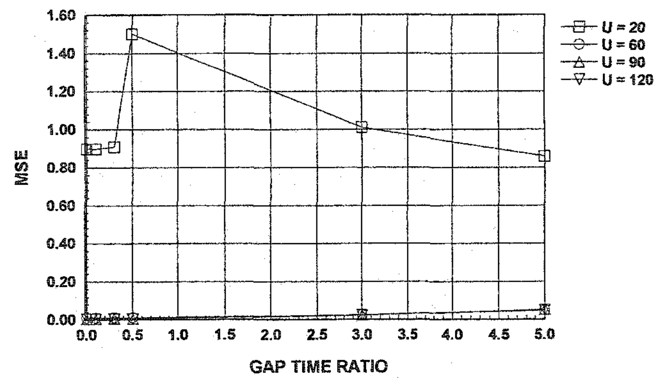


(b) BIAS

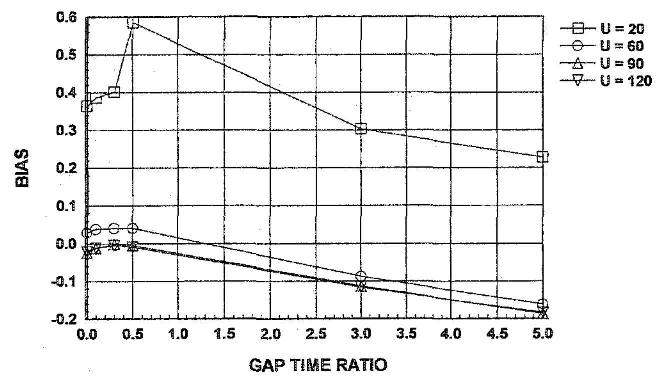


(c) MAD

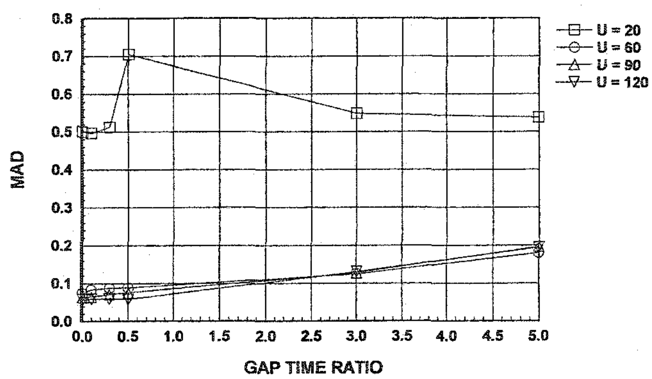
Fig A.II.4 Experimental units effect of the PWP-GT model ($\delta = 2.0$)



(a) MSE

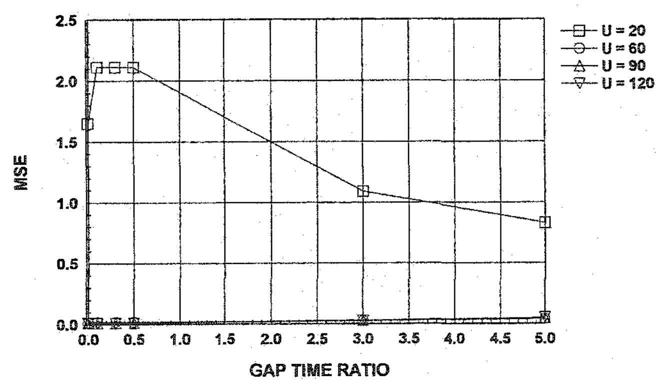


(b) BIAS

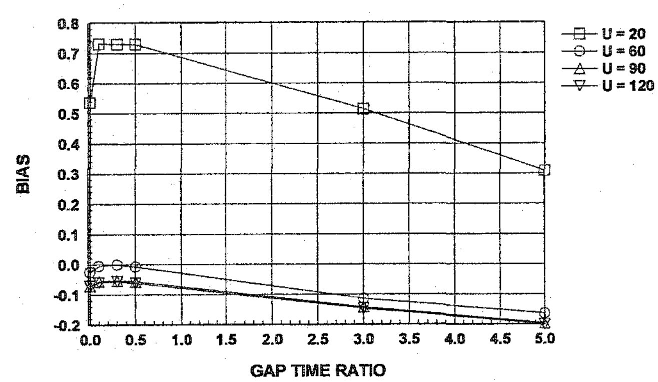


(c) MAD

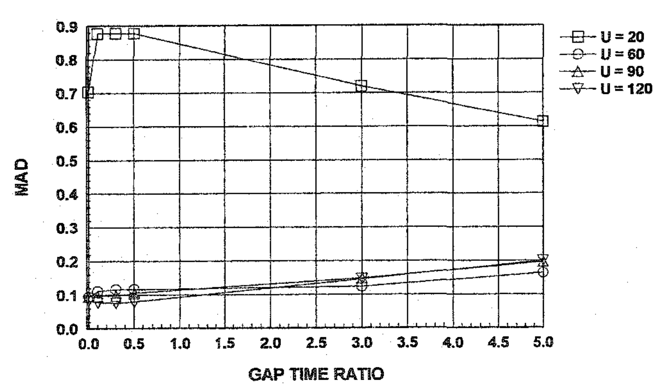
Fig A.II.5 Experimental units effect of the PWP-GT model ($\delta = 1.0$)



(a) MSE

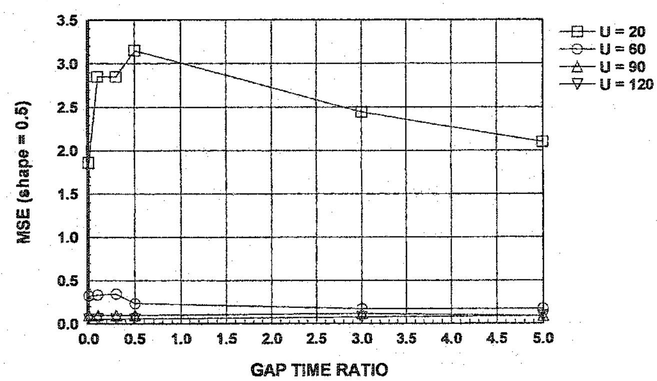


(b) BIAS

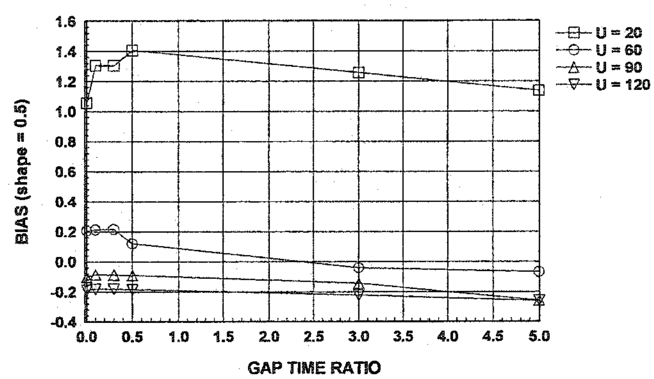


(c) MAD

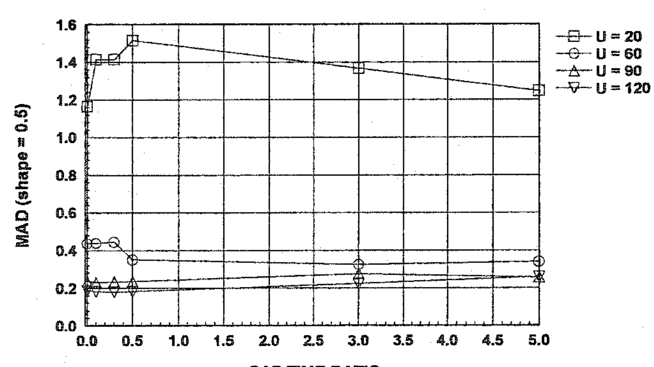
Fig A.II.6 Experimental units effect of the PWP-GT model ($\delta = 0.8$)



(a) MSE



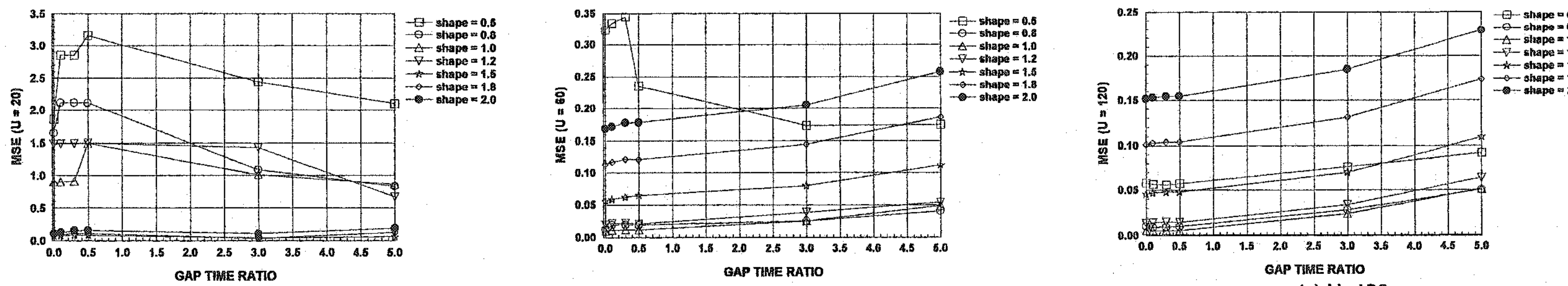
(b) BIAS



(c) MAD

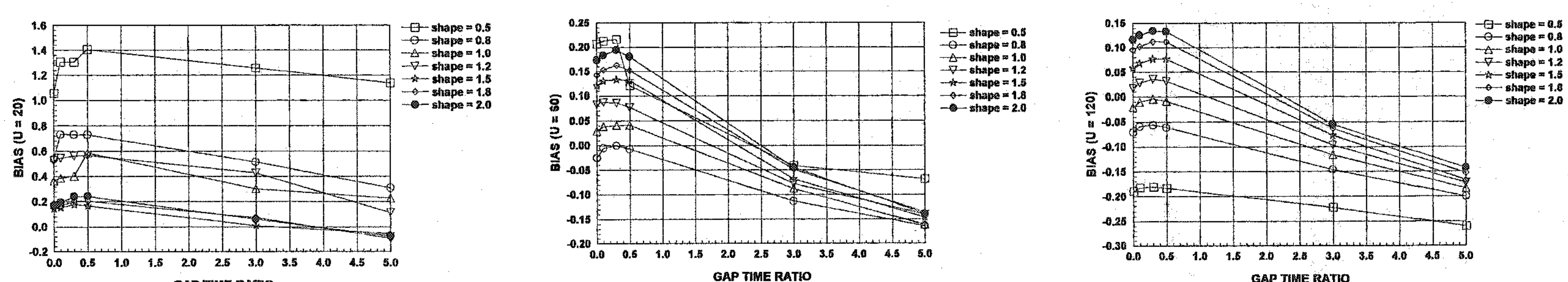
Fig A.II.7 Experimental units effect of the PWP-GT model ($\delta = 0.5$)

II.2 Shape parameter



(a) U=20 (b) U=60 (c) U=120

Fig A.II.8 Shape parameter effect of the PWP-GT model (MSE)



(a) U=20 (b) U=60 (c) U=120

Fig A.II.9 Shape parameter effect of the PWP-GT model (BIAS)

Appendix III (Program for semi-parametric proportional intensity models robustness for right-censored recurrent failure data)

```
DATA POWERLAWB;
      RETAIN SEED 539;
      FORMAT T Y 16.2;

      DO ITEM = 1 TO 60;

            P=RANUNI(SEED);
            IF P < 0.4 THEN CENSOR=0;
            ELSE CENSOR=1;

            F=FLOOR(10*RANUNI(SEED))+1;
            T = 0;
            M = 0;

            TSTART=0;

            DO FAILURE = 1 TO 10;

                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 1.0;
                  IF ITEM <= 30 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  IF(FAILURE>F & CENSOR=0)THEN STATUS=0;
                  ELSE STATUS=1;

                  IF STATUS=1 THEN DO;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  TSTOP=T;
                  OUTPUT;
                  TSTART=TSTOP;

                  END;

                  IF STATUS=0 THEN DO;
                  T=0;
                  Y=0;
                  TSTOP=TSTART;
                  OUTPUT;
                  END;

            END;

      END;

DATA TBFB;
      SET POWERLAWB;
      DROP  M NU DELTA;
PROC PRINT DATA=TBFB;
```

```
      TITLE1 'RIGHT CENSORING DATA OF TIME BETWEEN FAILURES';

DATA CENSOR;
      SET TBFB;
      DROP X P F;
      IF STATUS=1 THEN DELETE;
PROC PRINT DATA=CENSOR;
      TITLE'CENSOR';

DATA UNCENSOR;
      SET TBFB;
      DROP X P F;
      IF STATUS=0 THEN DELETE;
PROC PRINT DATA=UNCENSOR;
TITLE 'UNCENSOR';

DATA CENSOR_AG;
      SET TBFB;
      IF TSTART=TSTOP THEN DELETE;
PROC PRINT DATA=CENSOR_AG;
      TITLE'CENSOR_AG';

DATA CENSOR_PWP(DROP=LSTATUS);
      RETAIN LSTATUS;
      SET TBFB;
      BY ITEM;
      IF FIRST.ID THEN LSTATUS=1;
    IF (STATUS=0 AND LSTATUS=0) THEN DELETE;
    LSTATUS=STATUS;
PROC PRINT DATA=CENSOR_PWP;
      TITLE'CENSOR_PWP';

DATA CENSOR_WLW;
      SET TBFB;
PROC PRINT DATA=CENSOR_WLW;
      TITLE'CENSOR_WLW';

PROC PHREG DATA=CENSOR_AG;
      MODEL (TSTART,TSTOP)* STATUS(0)= CLASS;
      TITLE1' ANDERSEN-GILL SUMMARY';


DATA CENSOR_PWP1;
      SET CENSOR_PWP;
      IF FAILURE<11;
      CLASS1=CLASS*(FAILURE=1);
      CLASS2=CLASS*(FAILURE=2);
      CLASS3=CLASS*(FAILURE=3);
      CLASS4=CLASS*(FAILURE=4);
      CLASS5=CLASS*(FAILURE=5);
      CLASS6=CLASS*(FAILURE=6);
      CLASS7=CLASS*(FAILURE=7);
      CLASS8=CLASS*(FAILURE=8);
      CLASS9=CLASS*(FAILURE=9);
      CLASS10=CLASS*(FAILURE=10);

PROC PHREG DATA=CENSOR_PWP1;
```

```
      MODEL Y * STATUS(0)= CLASS1-CLASS10;
      STRATA FAILURE;
      TITLE1' PWP-GAP TIME SUMMARY';
      OUTPUT OUT=SURL_EST_PWP_GAP SURVIVAL=SURL_EST_PWP_GAP;
PROC SORT;
      BY FAILURE CLASS1-CLASS10 Y;

PROC PHREG DATA=CENSOR_PWP1;
      MODEL TSTOP * STATUS(0)= CLASS;
      TITLE1' PWP-TOTAL TIME SUMMARY';
      OUTPUT OUT=SURL_EST_PWP_TOTAL SURVIVAL=SURL_EST_PWP_TOTAL;
PROC SORT;
      BY CLASS TSTOP;


DATA CENSOR_WLW1;
      SET CENSOR_WLW;
      IF FAILURE<11;
      CLASS1=CLASS*(FAILURE=1);
      CLASS2=CLASS*(FAILURE=2);
      CLASS3=CLASS*(FAILURE=3);
      CLASS4=CLASS*(FAILURE=4);
      CLASS5=CLASS*(FAILURE=5);
      CLASS6=CLASS*(FAILURE=6);
      CLASS7=CLASS*(FAILURE=7);
      CLASS8=CLASS*(FAILURE=8);
      CLASS9=CLASS*(FAILURE=9);
      CLASS10=CLASS*(FAILURE=10);

PROC PHREG DATA=CENSOR_WLW1;
      MODEL TSTOP * STATUS(0)=CLASS;
      TITLE1' WEI-LIN-WEISSFELD SUMMARY';
      OUTPUT OUT=SURL_EST_WLW SURVIVAL=SURL_EST_WLW;
PROC SORT;
      BY CLASS TSTOP;

RUN;
```

Appendix IV (Program for robustness of semi-parametric proportional intensity models for right-censored recurrent failure data from a stationary counting process)

```
DATA POWERLAWB;
      RETAIN SEED 539;
      FORMAT T Y 16.2;

      DO ITEM = 1 TO 180;

            P=RANUNI(SEED);
            IF P < 0.4 THEN CENSOR=0;
            ELSE CENSOR=1;

            F=FLOOR(4*RANUNI(SEED))+1;
            T = 0;
            M = 0;

            TSTART=0;

            DO FAILURE = 1 TO 4;

                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 1.0;
                  IF ITEM <= 90 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  IF(FAILURE>F & CENSOR=0)THEN STATUS=0;
                  ELSE STATUS=1;

                  IF STATUS=1 THEN DO;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  TSTOP=T;
                  OUTPUT;
                  TSTART=TSTOP;

                  END;

                  IF STATUS=0 THEN DO;
                  T=0;
                  Y=0;
                  TSTOP=TSTART;
                  OUTPUT;
                  END;

            END;

      END;

DATA TBFB;
      SET POWERLAWB;
      DROP  M NU DELTA;
PROC PRINT DATA=TBFB;
```

```
      TITLE1 'RIGHT CENSORING DATA OF TIME BETWEEN FAILURES';

DATA CENSOR;
      SET TBFB;
      DROP X P F;
      IF STATUS=1 THEN DELETE;
PROC PRINT DATA=CENSOR;
      TITLE'CENSOR';

DATA UNCENSOR;
      SET TBFB;
      DROP X P F;
      IF STATUS=0 THEN DELETE;
PROC PRINT DATA=UNCENSOR;
TITLE 'UNCENSOR';

DATA CENSOR_AG;
      SET TBFB;
      IF TSTART=TSTOP THEN DELETE;
PROC PRINT DATA=CENSOR_AG;
      TITLE'CENSOR_AG';

DATA CENSOR_PWP(DROP=LSTATUS);
      RETAIN LSTATUS;
      SET TBFB;
      BY ITEM;
      IF FIRST.ID THEN LSTATUS=1;
    IF (STATUS=0 AND LSTATUS=0) THEN DELETE;
    LSTATUS=STATUS;
PROC PRINT DATA=CENSOR_PWP;
      TITLE'CENSOR_PWP';

DATA CENSOR_WLW;
      SET TBFB;
PROC PRINT DATA=CENSOR_WLW;
      TITLE'CENSOR_WLW';

PROC PHREG DATA=CENSOR_AG;
      MODEL (TSTART,TSTOP)* STATUS(0)= CLASS;
      STRATA FAILURE;
      TITLE1' ANDERSEN-GILL SUMMARY';


DATA CENSOR_PWP1;
      SET CENSOR_PWP;
      IF FAILURE<5;
      CLASS1=CLASS*(FAILURE=1);
      CLASS2=CLASS*(FAILURE=2);
      CLASS3=CLASS*(FAILURE=3);
      CLASS4=CLASS*(FAILURE=4);

PROC PHREG DATA=CENSOR_PWP1;
      MODEL Y * STATUS(0)= CLASS1-CLASS4;
      STRATA FAILURE;
      TITLE1' PWP-GAP TIME SUMMARY';
      OUTPUT OUT=SURL_EST_PWP_GAP SURVIVAL=SURL_EST_PWP_GAP;
PROC SORT;
```

```
        BY FAILURE CLASS1-CLASS4 Y;

PROC PHREG DATA=CENSOR_PWP1;
        MODEL TSTOP * STATUS(0)= CLASS1-CLASS4;
        STRATA FAILURE;
        TITLE1' PWP-TOTAL TIME SUMMARY';
        OUTPUT OUT=SURL_EST_PWP_TOTAL SURVIVAL=SURL_EST_PWP_TOTAL;
PROC SORT;
        BY FAILURE CLASS1-CLASS4 TSTOP;


DATA CENSOR_WLW1;
        SET CENSOR_WLW;
        IF FAILURE<5;
        CLASS1=CLASS*(FAILURE=1);
        CLASS2=CLASS*(FAILURE=2);
        CLASS3=CLASS*(FAILURE=3);
        CLASS4=CLASS*(FAILURE=4);

PROC PHREG DATA=CENSOR_WLW1;
        MODEL TSTOP * STATUS(0)=CLASS1-CLASS4;
        STRATA FAILURE;
        TITLE1' WEI-LIN-WEISSFELD SUMMARY';
        OUTPUT OUT=SURL_EST_WLW SURVIVAL=SURL_EST_WLW;
PROC SORT;
        BY FAILURE CLASS1-CLASS4 TSTOP;

RUN;
```

Appendix V (Program for semi-parametric proportional intensity models robustness for recurrent failure data with overhaul intervals)

```
DATA GENRATEA;
      RETAIN SEED 539;
      FORMAT T T1 Y 16.2;
      DO ITEM = 1 TO 20;
            T = 0;
            T1 = 0;
            M = 0;
            RATIO = 5.0;
            F=FLOOR(10*RANUNI(SEED))+1;
            DO FAILURE = 1 TO F;
                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 0.5;
                  IF ITEM <= 10 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  OUTPUT;
                  T1=T;
            END;

            DO FAILURE = F+1;
                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 0.5;
                  D = RATIO * Y;
                  T1=T1+D;
                  T=T1;
                  IF ITEM <= 10 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  OUTPUT;
                  T1=T;
            END;

            DO FAILURE = F+2 TO 10;
                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 0.5;
                  IF ITEM <= 10 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  OUTPUT;
```

```
                T1=T;
            END;

        END;

DATA TBF;
        SET GENRATEA;
        DROP SEED D;
        IF FAILURE>10 THEN DELETE;
PROC PRINT DATA=TBF;
TITLE1 'SIMULATED TIME BETWEEN FAILURES';

PROC PHREG DATA=TBF;
        MODEL (T1,T)=CLASS;
        TITLE'THE ANDERSEN-GILL SUMMARY';


DATA TBF_GT;
        SET TBF;
        IF FAILURE<11;
        CLASS1=CLASS*(FAILURE=1);
        CLASS2=CLASS*(FAILURE=2);
        CLASS3=CLASS*(FAILURE=3);
        CLASS4=CLASS*(FAILURE=4);
        CLASS5=CLASS*(FAILURE=5);
        CLASS6=CLASS*(FAILURE=6);
        CLASS7=CLASS*(FAILURE=7);
        CLASS8=CLASS*(FAILURE=8);
        CLASS9=CLASS*(FAILURE=9);
        CLASS10=CLASS*(FAILURE=10);


PROC PHREG DATA=TBF_GT OUTEST=BETA_TT;
        MODEL T = CLASS;
        TITLE'THE WEI-LIN-WEISSFELD SUMMARY';

PROC SORT DATA=TBF_GT ;
        BY FAILURE DESCENDING Y;
PROC PHREG DATA=TBF_GT OUTEST=BETA_GT;
        MODEL Y = CLASS1-CLASS10;
        STRATA FAILURE;
        TITLE1' THE PWP-GAP TIME SUMMARY';

RUN;
```

Appendix VI (Covariate proportional intensity modeling for recurrent data of two failure types (major and minor))

```
DATA GENRATEA;
      RETAIN SEED 539;
      FORMAT T T1 Y 16.2;
      DO ITEM = 1 TO 120;
            T = 0;
            T1 = 0;
            M = 0;
            ZM=0;
            ZN=0;
            F1=FLOOR(10*RANUNI(SEED))+1;

            DO FAILURE = 1 TO 10;
                  RETAIN M 0;
                  X = RANUNI(SEED);
                  DELTA = 1.2;
                  IF ITEM <= 60 THEN NU = 0.001;
                  ELSE NU = 0.01;
                  IF NU = 0.001 THEN CLASS = 0;
                  ELSE CLASS = 1;
                  T = ((T**DELTA)-(LOG(X)/NU))**(1/DELTA);
                  Y = T-M;
                  M = T;
                  OUTPUT;
                  T1=T;
            END;
      END;


DATA TBF;
      SET GENRATEA;
      DROP SEED M NU DELTA X;
      IF CLASS=1 & FAILURE^=F1 THEN ZN=1;
      IF CLASS=1 & FAILURE=F1 THEN ZM=1;

PROC PRINT DATA=TBF;
      TITLE1 'SIMULATED MINOR-MAJOR EVENTS';

PROC PHREG DATA=TBF;
      MODEL (T1,T)=ZM ZN;
      TITLE'THE ANDERSEN-GILL SUMMARY';


PROC SORT DATA=TBF;
      BY FAILURE;
PROC PHREG DATA=TBF OUTEST=BETA_WLW;
      MODEL T = ZM ZN;
      TITLE'THE WEI-LIN-WEISSFELD SUMMARY';

PROC SORT DATA=TBF;
      BY FAILURE DESCENDING Y;
PROC PHREG DATA=TBF OUTEST=BETA_PWP_GT;
      MODEL Y = ZM ZN;
      BY FAILURE;
      TITLE1' THE PWP-GAP TIME SUMMARY';
```

```
PROC PRINT DATA=BETA_PWP_GT;
      TITLE'BETA_PWP_GT';

RUN;
```

Baseline hazard function
The baseline hazard function in the proportional hazards (PH) model can be set as known or arbitrary depending on whether a parametric or semi-parametric method is applied. The baseline hazard function determines the intercept of the PH function, while the regression coefficient β decides the slope of the PH function.

Censoring probability
The ratio (probability) of the sample units that contain censored times to total sample units is defined as censored probability (P_c) in this research.

CMTBF
The abbreviation for cumulative mean time between failures defined as the mean time between failures per event. Mathematically, $CMTBF$ is derived from the time t divided by the expected number of failure events in $(0,t]$ in a discrete time system (Patrick (1991) and Ascher and Feingold (1984)).

Conditional method
The PWP-GT and PWP-TT both utilize the concept of the condition method. The intensity function for n^{th} event is determined based on the past history (in terms of the failure times, event count, etc.). Participants that have experienced $(n-1)^{st}$ event are qualified to contribute to the n^{th} event intensity function estimation.

Covariate
Covariates, introduced from the PH and PI models, featured as regression factors in lifetime or recurrent data analysis, and also termed as explanatory variables or concomitant variables. Covariates can be time-variant or constant throughout the observation time.

Multi-dimensional covariate modeling
Covariate modeling used to handle recurrent data with multiple failure types.

Cox-based regression methods:
Referred to as the PWP-GT, PWP-TT, AG, and WLW models. The Cox-based regression methods employ a partial maximum likelihood function to estimate the PI function.

Andersen-Gill method
The AG method employs the counting process concept to estimate the PI function. The dataset only contains uncensored data.

Prentice-Williams-Peterson method
The PWP method employs the conditional method to estimate the PI function. There are two ways to perform the PWP model depending on the time frame, The PWP-GT is on a local time scale and the PWP-TT is on a global time scale. The dataset contains uncensored data (failure times) and the first censored time.

Wei-Lin-Weissfeld method
The WLW method employs the marginal method to estimate the PI function. The dataset contains full records (both failure and censored times).

Discontinuous risk-free-intervals
The concept of the discontinuous risk-free-intervals relaxes the assumption of zero repair time.

Hazard function
The hazard function is defined as the probability density function (p.d.f.) divided by the survival function.

HPP
Homogeneous Poisson process (HPP) is a sequence of independent and identically distributed exponential random variables (Ascher and Feingold (1984)).

Independent increment
The intensity function is not affected by other time increments, and thus is memoryless.

Information matrix
The second derivative of maximum likelihood function with respect to parameters utilized in the parametric Lawless method

IMTBF
The abbreviation for instantaneous mean time between failures, defined as the derivative of failure time with respect to the expected number of failures (Patrick (1991) and Ascher and Feingold (1984)).

Intensity function
The instantaneous rate of event occurrence for a point process in a continuous time compared to a hazard rate in a discrete time.

Lawless method
A parametric method that assumes the true underlying process is known. The Lawless method employs the maximum likelihood method and the Newton-Raphson iterative method to estimate the relevant parameters.

Left-censoring
Left-censoring observations occur when the failure time data is incomplete, and truncated from the left, due to the loss of the historical data.

Major and minor failure types
The counting process with a mixed events stream is composed of two failure types (major and minor).

Major overhaul period
The assumption of zero repair time can be relaxed using the discontinuous risk-free-intervals modeling. The traditional method neglects the repair time in repairable systems reliability.

Marginal method
Unlike the conditional method, all subjects (including censored subjects) have equal likelihood to contribute to the intensity function for the n^{th} event.

Multiple event types
Multiple k event types are modeled in the multivariate proportional intensity function $\lambda_k(t;z_k)$, where each intensity function performs an independent analysis for each failure type.

NHPP
Non-homogeneous Poisson Process (NHPP) is a nonstationary counting process with intensity function $\lambda(t)$, where t is a time variable. The number of events in any interval Δt is the integration of the intensity function along with the time interval Δt.

PHREG
The syntax in the SAS program utilized to perform the regression analysis in the Cox PH model or the accelerated failure time model

Product-Limit method
The Product-Limit method is a non-parametric estimator of a survival function $\hat{S}(t)$, which is defined as (Lawless (1982)): $\hat{S}(t) =$ { Number of observations $\geq t$ } /n, where n denotes sample size. The survival function $\hat{S}(t)$ is a step function, which decreases by 1/n after each observed lifetime in the PH model. In the case of censored data, the survival function is modified as the Kaplan-Meier estimate (Lawless (1982)).

Proportional hazards (PH) function
The PH model deals with single event data (lifetime data), while the PI model is designated to handle recurrent data. The proportionality property is that the hazard functions of any two individuals are proportional to each other (Lawless (1982)).

Proportional intensity (PI) function
The PI model is an extension of PH model when the data contains more than one occurrence. The proportionality property of a PI model follows the PH model.

ROCOF
ROCOF is an abbreviation form for rate of occurrence of failures. Mathematically, ROCOF is defined as the instantaneous rate of change of the expected number of failures in a continuous time (Patrick (1991) and Ascher and Feingold (1984)).

Repairable systems
Systems are designed to be repairable after each failure in the system and the system can be restored to a certain degree between as-good-as-new and as-bad-as-old.

Replacement
Systems are designed to be non-repairable, and the system will be replaced after each failure.

Right-censoring
The unit is removed from observation after a certain time point or number of failures. Leemis (1995) listed a few cases of right-censoring: cost consideration, high reliability products, the death of a patient, losing contact with a patient, etc.

Risk-free-intervals
The risk-free-intervals concept originates from the hospitalization in a clinical study. When a patient is admitted to the hospital for drug treatments, the period of the hospitalization is considered as a risk-free-interval. Likewise, for the reliability engineering application, the system is not at risk when a major overhaul is taken place.

Risk interval:
Risk interval defines the duration when a subject is at risk of having an event given under a time scale (Kelly and Lim (2000)).

Risk interval- total time
The total time (global time) is the duration starting at the beginning of the experiment. The clock resets to zero as an event occurs. The risk interval of the total time scale can be expressed as $(0, t_n)$, where n denotes the event number.

Risk interval- gap time
The gap time (local time) is the duration starting at the end of the previous event. The clock resets to zero as an event occurs. The risk interval of the gap time scale can be expressed as $(0, t_n - t_{n-1})$, where n denotes the event number.

Risk interval- counting process
A subject is not considered at risk for n^{th} event until the end of the $(n-1)^{st}$ event. The risk interval of a counting process can be expressed as (t_{n-1}, t_n), where n denotes the event number. Note that the clock does not reset to zero as an event occurs.

Risk set:
The risk set contains the individuals that are at risk for the n^{th} event.

Risk set- Unrestricted/ common baseline hazard
The risk set is determined regardless the event number, which means the subjects have equal likelihood to contribute to the n^{th} intensity function and share the same baseline hazard.

Risk set- Semi-restricted/ event-specific baseline hazard
Event-specific baseline intensity allows the individuals that have experienced the $(n-1)^{th}$ event to contribute to the n^{th} event intensity function. However, the semi-restricted concept tolerates the censored individuals as in the risk set regardless of any censoring.

Risk set- Restricted/ event-specific baseline hazard
The baseline hazard changes stratum by stratum defined by the event count, also termed as restricted baseline intensity. The subject is not at risk of contributing to the n^{th} event until the subject has experienced the $(n-1)^{st}$ event.

Sample size
There are two expressions of defining the sample size in this study: the number of units and recurring events (failure count) for each sample unit.

Score vector
The first derivative of the maximum likelihood function with respect to the parameter utilized in the Lawless parametric method