University of Pretoria etd – Vlok, P-J (2006)

Dynamic Residual Life Estimation of Industrial Equipment based on Failure Intensity Proportions

By
Pieter-Jan Vlok

Submitted in partial fulfillment of the requirements for the degree
Philosophiae Doctor

in the
Department of Industrial and Systems Engineering

in the
Faculty of Engineering, Built Environment and Information Technology

at the
UNIVERSITY OF PRETORIA
Advisor: Prof. S.J. Claasen
October, 2001

University of Pretoria etd – Vlok, P-J (2006)

Acknowledgements

I would like to express my sincere gratitude towards the following people and organizations:

(i) ABB South Africa, especially Mr. Toni de Sousa, for accommodating and partially sponsoring my research.

(ii) Prof. Schalk Claasen for his supervision and guidance.

(iii) Prof. Andrew Jardine and Dr. Dragan Banjevic of the University of Toronto for their comments, advice and support.

(iv) Family and friends for their encouragement and continued interest in my research.

(v) Our heavenly Father for strength and perseverance.

The Author
October, 2001

University of Pretoria etd – Vlok, P-J (2006)

OPSOMMING

Dinamiese Oorblywende Lewe Skatting van Industriële Toerusting Gebaseer op Falingsintensiteit Verhoudings

Deur
Pieter-Jan Vlok
Promotor: Prof. S.J. Claasen

Philosophiae Doctor

Departement van Bedryfs- en Sisteemingenieurswese
Fakulteit van Ingenieurswese, Bou-omgewing en Inligtingstegnologie
Universiteit van Pretoria

Daar is 'n wêreldwye strewe na optimering van instandhoudingsbesluitneming in 'n meer mededingende vervaardigingsindustrie. Voorkomende instandhouding is dikwels die mees georganiseerde en koste effektiewe strategie om te volg, maar 'n besluit moet steeds geneem word oor die tydstip waarop die voorkomende instandhouding gedoen word. Gebruiksgebaseerde instandhoudingsbesluitneming is tot 'n groot mate geoptimeer deur statistiese analise van falingsdata, terwyl voorspellende voorkomende instandhouding (toestandsmonitering) geoptimeer word deur van meer gesofistikeerde tegnologie gebruik te maak. Baie min werk is egter al gedoen om die voordele van hierdie twee denkwyses te kombineer. Hierdie proefskrif het ontstaan na 'n besef van die moontlike verbetering in instandhoudingspraktyk deur gebruiksgebaseerde instandhoudingsoptimeringstegnieke te kombineer met hoë tegnologie toestandsmonitering.

In hierdie proefskrif word 'n benadering ontwikkel waarmee oorblywende lewe van industriële toerusting dinamies geskat word deur statistiese falingsanalise en gesofistikeerde toestandsmoniteringstegnieke te kombineer. Die benadering is gebaseer op falingsintensiteitverhoudings wat bereken word uit historiese oorlewingstye en die dienooreenkomstige diagnostiese inligting verkry uit toestandsmoniteringsresultate. Gekombineerde Proporsionele Intensiteitsmodelle (PIMe) vir nie-herstelbare en herstelbare stelsels, wat die meeste konvensionele verbeterings op PIMe as spesiale gevalle bevat, asook numeriese metodes om die regressie koeffisiënte te bepaal, is ontwikkel.

University of Pretoria etd – Vlok, P-J (2006)

Saam met die oorblywende lewe skatting benadering, is 'n gebruikersvriendelike grafiese metode waarmee oorblywende lewe skattings vertoon kan word, ontwikkel. Hierdie metode is natuurlik selfs vir onervare data analiste maklik verstaanbaar. Die oorblywende lewe skatting benadering is toegepas op 'n tipiese datastel verkry van 'n Suid-Afrikaanse industrie en resultate is vergelyk met resultate verkry van 'n soortgelyke, bestaande instandhoudingsbesluitnemingstegniek. Die vergelyking toon aan dat die benadering ontwikkel in hierdie proefskrif relevant en prakties is en volgens sekere kriteria marginaal beter is as die genoemde bestaande instandhoudingsbesluitnemingstegniek.

SLEUTELWOORDE: Oorblywende lewe, Proporsionele gevaar, Falingsintensiteit, Voorwaardelike gemiddelde

University of Pretoria etd – Vlok, P-J (2006)

SUMMARY

Dynamic Residual Life Estimation of Industrial Equipment based on Failure Intensity Proportions

By
Pieter-Jan Vlok
Advisor: Prof. S.J. Claasen

Philosophiae Doctor

Department of Industrial and Systems Engineering
Faculty of Engineering, Built Environment and Information Technology
University of Pretoria

There is a world-wide drive to optimize maintenance decisions in an increasingly competitive manufacturing industry. Preventive maintenance is often the most organized and cost efficient strategy to follow, but a decision still has to be made on the optimal instant to perform preventive maintenance. Use based preventive maintenance decisions have been optimized through statistical analysis of failure data while predictive preventive maintenance (condition monitoring) has been optimized by utilizing more sophisticated technology. Very little work has however been done to combine the advantages of the two schools of thought. This thesis originated from a realization of the potential improvement in maintenance practice by combining use based preventive maintenance optimization techniques with high technology condition monitoring.

In this thesis an approach is developed to estimate residual life of industrial equipment dynamically by combining statistical failure analysis and sophisticated condition monitoring technology. The approach is based on failure intensity proportions determined from historic survival time information and corresponding diagnostic information such as condition monitoring. Combined Proportional Intensity Models (PIMs) for non-repairable and repairable systems, containing the majority of conventional PIM enhancements as special cases, with numerical optimization techniques to solve for the regression coefficients, are derived.

In addition to the residual life estimation approach, a user-friendly graphical method with which residual life estimates can be presented was also developed. This method is natural

University of Pretoria etd – Vlok, P-J (2006)

and easy to comprehend, even by inexperienced data analysts.

The residual life estimation approach is applied to a typical data set from a South African industry and results are compared to those obtained from a similar, established maintenance decision support tool. This comparison showed that the approach developed in this thesis is relevant, practical and marginally better than the established decision support tool for certain criteria.

KEYWORDS: Residual life, Proportional hazards, Failure intensity, Conditional expectation

University of Pretoria etd – Vlok, P-J (2006)

Table of Contents

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

List of Figures

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

List of Tables

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

University of Pretoria etd – Vlok, P-J (2006)

List of Abbreviations

AFTM	Accelerated Failure Time Model
AHM	Additive Hazard Model
AIM	Additive Intensity Model
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
AROCOF	Average Rate of Occurrence of Failure (See ROCOF)
BAO	Bad-as-old
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BOWN	Better than old but worse than new
BPP	Branching Poisson Process
CM	Condition Monitoring
CMF	Cumulative Mean Function
CMMS	Computerized Maintenance Management System
EHRM	Extended Hazard Regression Model
FOM	Force of Mortality
GAN	Good-as-new
HFD	High Frequency Domain
HPP	Homogeneous Poisson Process
IID	Independent and Identically Distributed
KS	Kolmogorov-Smirnov
LCC	Life Cycle Cost
LLP	Log-linear Process
LNF	Lifted Noise Floor
MLE	Maximum likelihood estimates
MRL	Mean residual life
MTBR	Mean Time Between Renewals
NDT	Non-destructive testing
NHPP	Non-homogeneous Poisson Process
OEM	Original Equipment Manufacturers
PAR	Proportional Age Reduction
PAS	Proportional Age Setback
PDF	Probability Density Function
PHM	Proportional Hazards Model

University of Pretoria etd – Vlok, P-J (2006)

PIM	Proportional Intensity Model
PLP	Power-law Process
PMIM	Proportional Mean Intensity Model
POM	Proportional Odds Model
PWP	Prentice Williams Peterson
RCM	Reliability Centered Maintenance
RLE	Residual Life Estimation
ROCOF	Rate of Occurrence of Failure, i.e. the time derivative of an expected number of failures
ROOF	Repair only on failure
RV	Random variable
TPM	Total Productive Maintenance
TPMX	Transition Probability Matrix
TTT	Total Time on Test
URL	Useful Remaining Life
WO	Worse than old
WRP	Weibull Renewal Process

University of Pretoria etd – Vlok, P-J (2006)

Notation

ν	Additive functional term
ι_τ	Average intensity
$B(t)$	Backward recurrence time, i.e. $t - T_{N(t)}$
$\hat{c}_j$	Correlation coefficient of lag j
R	Linear correlation coefficient
C_1	Cost of minimal repair
C_2	Cost of system replacement
C_f	Cost of unexpected failure maintenance
C_p	Cost of planned preventive maintenance
$F_X(x)$	Cumulative density function (Unreliability function)
D_x	Derivative with respect to x
C	Event indicator, i.e. $C = 0$ in case of suspension and $C = 1$ in case of failure
$M(t)$	Expected number of failures up to time t, i.e. $E[N(t)]$, for a situation modeled by the full intensity, $\iota(t)$
$M_u(t)$	Expected number of failures up to time t, i.e. $E[N(t)]$, for a situation modeled by the unconditional intensity, $\iota_u(t)$
$W(\mathbb{D})$	Expected time until replacement in the decision-model of Makis and Jardine, regardless whether preventive action or failure
$E[]$	Expected value of a function
τ	Factor that acts additively on x or t in g to represent a time jump or time setback that could be system copy- and stratum-specific
ψ	Factor that acts multiplicatively on x or t in g to result in an acceleration or deceleration of time that could be system copy- and stratum-specific
$h_X(x)$	Force of Mortality
$W(t)$	Forward recurrence time, i.e. $T_{N(t)+1} - t$
g	Fully parametric baseline function used in a combined PIM that could be system copy- and stratum-specific
t	Global time
H_t	History or filtration of a process
Δ	Inspection interval
$\{N(t), t \geq 0\}$	Integer valued counting process

University of Pretoria etd – Vlok, P-J (2006)

ι	Intensity of a process (also called *full* intensity or *conditional* intensity)
ι	Intensity or conditional intensity
T_i	i^{th} arrival time
X_i	i^{th} interarrival time
L	Likelihood
x	Local time
ι_u	Mean intensity or unconditional intensity
S	Mean sojourn time of a system in a particular state
λ	Multiplicative functional term that acts on g
$N(t)$	Number of failures recorded in the interval $(0, t]$
d_s	Number of events observed in stratum s
q	Number of observed events
n	Number of parts in a system
w	Number of system copies
n^*	Optimal number of minimal repairs before system replacement
I^*	Optimal system replacement time under the minimal repair assumption
$\rho_1(t)$	Peril rate, i.e. the ROCOF of an NHPP, modeled by a log-linear process
$\rho_2(t)$	Peril rate, i.e. the ROCOF of an NHPP, modeled by a power-law process
$\Pr[]$	Probability
$f_X(x)$	Probability density function
$Q(\mathbb{D})$	Probability that failure replacement will occur in Makis and Jardine's cost optimization decision-model
ζ	Random variable that acts as a frailty in a combined PIM and that could be system copy- and stratum-specific
$R_X(x)$	Reliability function, i.e. $1 - F_X(x)$
$\mu(x, \boldsymbol{\theta})$	Residual life of a non-repairable system
$\mu(t, \boldsymbol{\theta})$	Residual life of a repairable system
$\mathbb{D}$	Threshold risk level
$v(t)$	Time derivative of an expected number of failures, i.e. ROCOF
$\boldsymbol{\theta}$	Vector containing all form parameters of a PIM
$\boldsymbol{z}$	Vector containing covariates that may be time-dependent, i.e. $\boldsymbol{z}(x)$ in the non-repairable case and $\boldsymbol{z}(t)$ in the repairable case
G	Warning level function

University of Pretoria etd – Vlok, P-J (2006)

Chapter 1

Problem Statement

1.1 Introduction

Maintenance engineering is one of the fasted growing engineering disciplines in the world. Industry has only started to realize the importance of maintenance in the early 1980's and, ever since, there was no turning back the rapid development in the theory of maintenance. This theory is also more readily accepted by maintenance practitioners in industry as the mindset with regards to maintenance changes and greater successes are achieved by formal maintenance programs.

As is the case with most engineering disciplines, there is a drive in the field of maintenance engineering to optimize methodologies and practices. The maintenance fraternity has realized that the use of formalized maintenance models and tactics alone are not necessarily the optimal way to maintain equipment. One aspect of formal maintenance that needs optimization is decision making in life-limiting maintenance strategies, i.e. preventive maintenance, because of enormous losses industries are suffering due to a waste of residual life of equipment.

Preventive maintenance practitioners* have mostly reasoned along one of two schools of thinking. The first is to take action (replacement, repair or overhaul) based purely on an item's age as measured in time, miles, tons processed or any other convenient process parameter. The second is to assess the condition of an item through diagnostic measurements, which may include vibration monitoring, results of oil analysis, thermographic profiles, pressure, temperature, etc. This second viewpoint is referred to as predictive maintenance. Coetzee (1997) compiled a maintenance strategy tree that serves as a concise summary of possible

*Preventive maintenance, contrary to popular believe, is not necessarily the optimal maintenance strategy to apply. Any strategy's technical and economical feasibility should be determined before it is implemented. A methodology such as Reliability Centered Maintenance (RCM) or Total Productive Maintenance (TPM) should lead maintenance practitioners to the correct strategy.

University of Pretoria etd – Vlok, P-J (2006)

maintenance strategies. See Figure 1.1.

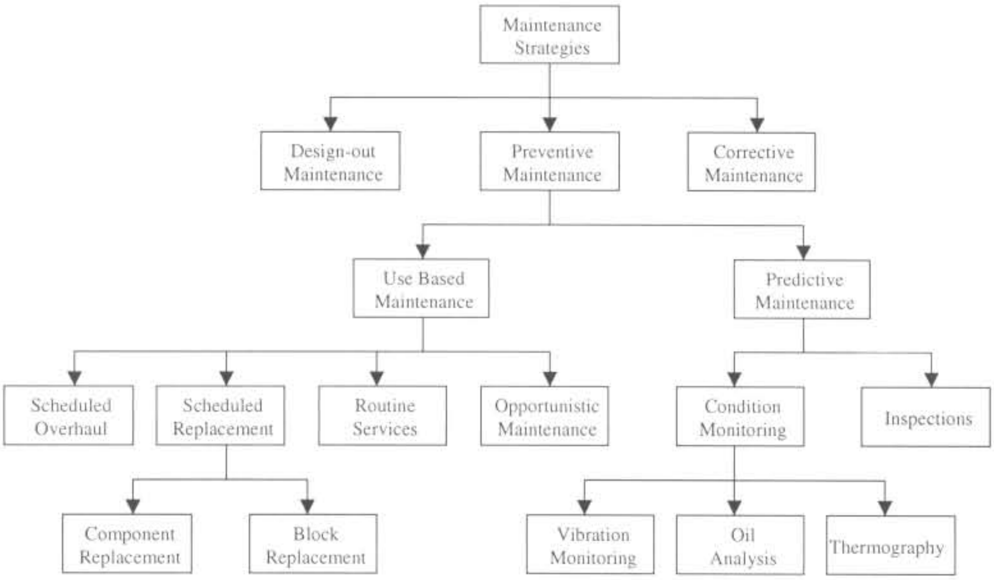

Figure 1.1: Maintenance Strategy Tree

Preventive maintenance is performed for one reason only: to prevent unexpected failure, which is in most cases considerably more expensive than planned preventive action. Unexpected failures often involve costly secondary damage to equipment, production losses, late delivery penalties, overtime labor costs and even loss of life. Preventive action is usually inexpensive relative to corrective action because of the planned nature of this type of action that eliminates many of the unwanted cost factors associated with unexpected failure.

In the case of use based maintenance, action is taken (by definition) only when an item has reached a certain age [†]. The time at which action is taken should be chosen in such a way that acceptably little residual life is wasted in the process but also such that the risk of unexpected failure does not rise unacceptably high. Optimization of use based maintenance thus involves a tradeoff between the waste of residual life and the risk of suffering an unexpected failure.

Predictive maintenance technologies, on the contrary, strive through continual [‡] assessment of an item's condition to warn those concerned of an imminent failure shortly before occurrence of failure. With advanced technology available at present, this seems to be a much more elegant approach than use based action. Closer investigation reveals that this is not necessarily true because, even with the advanced technology, there are still numerous unknowns

[†] *Time* will be used consistently to refer to an item's age but it should be emphasized that any convenient process parameter may be used

[‡] The word *continual* may be replaced by *continuous* in some cases

University of Pretoria etd – Vlok, P-J (2006)

that cannot be eliminated and a tradeoff has to be made again. In this case the tradeoff is between the accuracy of the technology utilized to perform the condition assessment and the risk of running into an unexpected failure.

Both use based maintenance and predictive maintenance procedures have been optimized individually but very little work has been done to combine the advantages of the two schools of thought to produce an optimal solution. In this thesis, an methodology will be developed to merge the advantages of the two approaches into one approach that can be used as an authoritative decision making tool.

1.2 Conventional use based maintenance optimization

Lawrence (1999) studied mathematical use based optimization techniques in maintenance and concluded that most models address one of three questions,

(i) How often should a component be replaced?
(ii) How many spare parts should be kept in stock?
(iii) How should maintenance tasks be scheduled?

This section (and thesis) addresses point (i).

Many authors agree that the only scientific way to optimize use based maintenance strategies is through statistical analysis of event data. In this section, conventional optimization techniques are discussed, i.e. optimization through statistical models without covariates [§] or discontinuities. This field is poorly understood by maintenance practitioners, mainly because of the confusing terminology found in the literature. It is thus very important to define clear notation before any further discussion on optimization of used based maintenance through statistical modeling.

1.2.1 Terminology

The terminology of Ascher and Feingold (1984) will be used in this thesis. Ascher and Feingold's book was specifically written with the objective to clear some of the confusion in the field of statistical failure analysis. First of all it is important to distinguish between different types of items:

[§]Covariates are often also referred to as *explanatory variables*.

University of Pretoria etd – Vlok, P-J (2006)

(i) *Part.* An item that is never disassembled and is discarded after first failure ¶.
(ii) *Socket.* A space that, at any given time, holds a part of a given type.
(iii) *System.* A collection of two or more sockets with their associated parts that is interconnected to perform a specific function(s).
(iv) *Non-repairable system.* A system that is discarded the first time it ceases to perform satisfactory, i.e. after first failure.
(v) *Repairable system.* A system that, after failure, can be restored to perform all of its function by any method other than complete replacement of the system.

After a system is repaired it could be in one of the following states:

(i) As good as new (GAN).
(ii) As bad as old (BAO).
(iii) Better than old but worse than new (BOWN).
(iv) Worse than old (WO).

The GAN and BAO assumptions are by definition the backbone of conventional statistical failure data analysis. Models with covariates or discontinuities are required to model BOWN or WO situations.

It is also very important to define appropriate time scales to measure life times of items. See Figure 1.2 for an example sample path of a failure process.

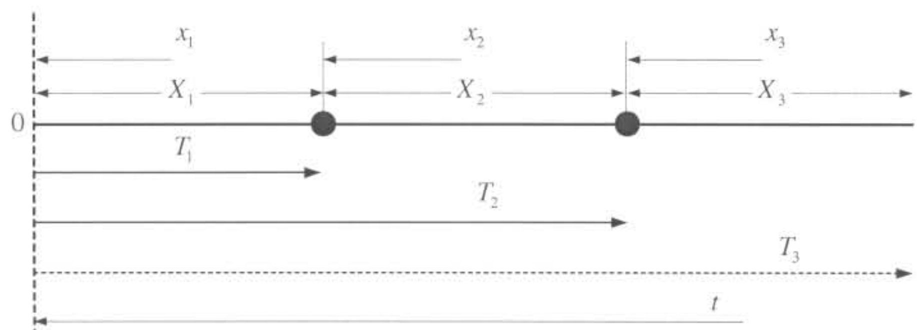

Figure 1.2: Example sample path of a failure process (Dots denote failures)

In Figure 1.2, X_i, $i = 1, 2, 3...$, refers to the *interarrival* time between the $(i-1)^{\text{th}}$ failure and i^{th} failure. X_i is a random variable (RV) with $X_0 \equiv 0$. This is referred to as *local* time and is

¶An item has failed when it no longer performs according to certain preset standards. This does not necessarily imply complete destruction.

University of Pretoria etd – Vlok, P-J (2006)

convenient to use when analyzing non-repairable systems. The real variable x_i measures the time elapsed since the most recent failure. T_i, $i = 1, 2, 3, ...$, measures time from 0 to the i^{th} failure time. T_i is also called the *arrival* time to the i^{th} failure and is mostly used to analyze repairable systems. This time scale is referred to as *global* time.

Clearly, $T_k \equiv X_1 + X_2 + ... + X_k$. From this, a RV $N(t)$ can be defined as the maximum value of k for which $T_k \leq t$, i.e. $N(t)$ is the number of failures that occur during $(0, t]$. $N(t), t \geq 0$ is the integer valued *counting process* that includes information on both the number of failures in $(0, t]$, $N(t)$, and the instants of occurrence, $T_1, T_2, ...$

Another important concept used in survival data modeling is that of the *backward recurrence* time, $B(t)$. It is defined as the time from the arbitrary time t to the immediately preceding failure, i.e. $B(t) \equiv t - T_{N(t)}$. Similarly, is the *forward recurrence* time, $W(t)$, defined as $W(t) \equiv T_{N(t)+1} - t$.

1.2.2 Selecting an appropriate model type

The process of selecting the correct model type for a particular data set is totally ignored in many applications of statistical failure analysis theory. Ascher and Feingold (1984) have constructed an outline of this process based on fundamental statistics. See Figure 1.3.

Some comments will be made on Figure 1.3:

(i) *Chronologically ordered X_i's.* It is extremely important to keep data in chronological order when starting with the process of deciding on the model type. Very often, failure data is reordered by magnitude which makes the process appear to follow, for example, an exponential distribution according to Ascher and Hansen (1998).

(ii) *Trend testing.* A number of techniques exist to recognize trends in data. Graphical techniques include (a) plotting cumulative failure times versus cumulative time on linear paper (Nelson (1982)); (b) estimating the average rate of occurrence of failure (ROCOF, see Section A.3) in successive time periods; and (c) Duane plots as introduced by Duane (1964).

Mathematical tests generally suitable to identify trends in data include De Laplace (1773) (commonly referred to as Laplace's test), Bartholomew (1955), Cox (1955), Bartholomew (1956a), Bartholomew (1956b), Bates (1955), Boswell (1966), Cox and Lewis (1966), Boswell and Brunk (1969), Lorden and Eisenberger (1973) and Saw (1975). More recent examples are Bain, Engelhardt, and Wright (1985), Lawless and Thiagarajah (1996), Martz and Kvam (1996) and Vaurio (1999). Laplace's test is regarded as the most reliable test and is used most often because it produces useful results even for small samples and its result is easily interpreted. Laplace's test is discussed in

University of Pretoria etd – Vlok, P-J (2006)

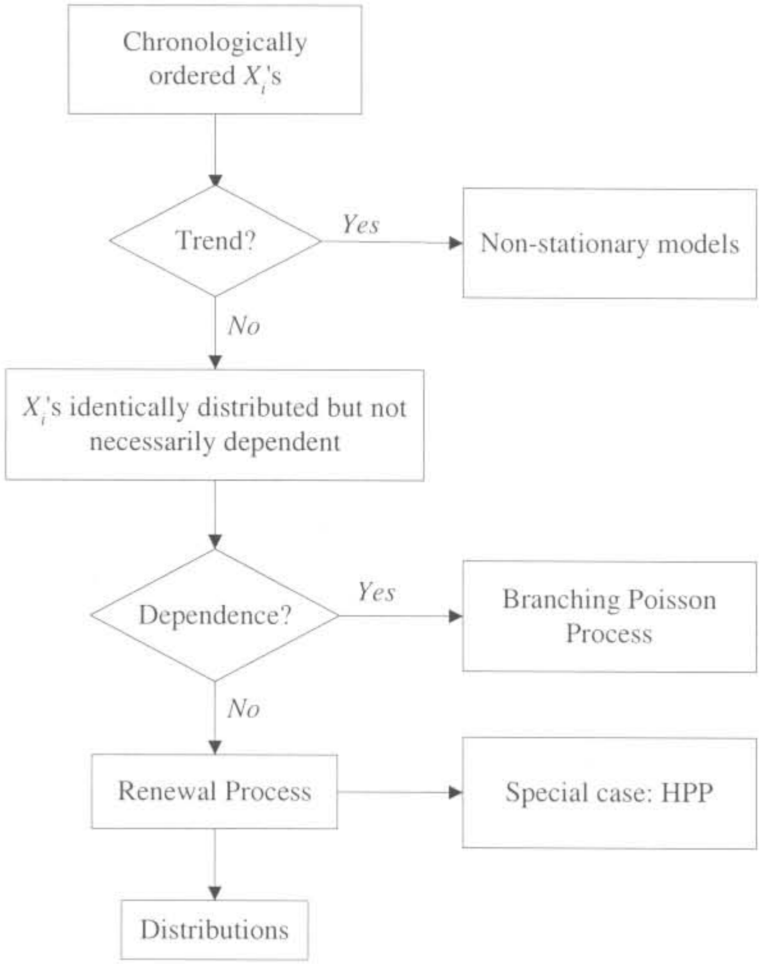

Figure 1.3: Statistical failure analysis of successive interarrival times of a system. (Adapted from Ascher and Feingold (1984)).

Section A.1.

(iii) *Non-stationary models.* Non-stationary models should be used to model data with a definite trend. The Non-homogeneous Poisson Process (NHPP) is used extensively for this purpose. Countless examples of the application of the NHPP are found in the literature, including Kumar and Westberg (1996b), Vineyard, Amoako-Gyampah, and Meredith (1999), Rhodes, Halloran, and Longini (1996), Percy, Kobbacy, and Ascher (1998), Newby (1993) and Lawless (1987). The NHPP is defined and described in Section A.3.3. Although the NHPP is used most often to model repairable systems' failure behaviour, there are also some other fundamentally different non-stationary models suitable for this application. See for example Cozzolino (1968), Singpurwalla (1978) and McWilliams (1979). These approaches were never very popular and are seldom cited in the literature.

University of Pretoria etd – Vlok, P-J (2006)

Differential equations are also suitable to model non-stationary point processes in special cases. Schafer, Sallee, and Torrez (1975) have summarized a few differential equation models for repairable systems. Another approach occasionally used to model repairable systems' failure behaviour is time series models such as the Auto Regressive (AR) model (see Chatfield (1980)) and the Box-Jenkins Auto Regressive Integrated Moving Average (ARIMA) model (see Wals and Bendell (1987)). The Box-Jenkins model has been used on a few occasions to model software reliability. See for example Burtschy, Albeanu, Boros, Popentiu, and Nicola (1997) and Chatterjee, Misra, and Alam (1997).

(iv) *Testing for dependence.* Although testing for dependence of interarrival times are of extreme importance in reliability modeling, it is almost always ignored. Two reasons for this are (1) the need for large sample sizes; and (2) the complexity of interpreting dependency tests. Cox and Lewis (1966) propose a very natural technique to test for dependency by simply calculating the sample correlation coefficient of lag j, i.e. $\hat{c}_j$. Thus, the correlation between X_i and X_{i+j} is calculated for $i = 1, 2, ..., m - j$ and $1 < i + j \leq m$ where m is the total number of observed events.

(v) *Branching Poisson Process (BPP).* The BPP is described in Section A.3.4. "*The BPP potentially has wide applicability to reliability problems*" according to Ascher and Feingold (1984). However, no practical application of the BPP was found in the literature. This could be because of large data set requirements and that the reliability fraternity still has not accepted and understood a model like the NHPP.

(vi) *Renewal Process.* A renewal process describes an item that, after a failure, is simply replaced by a new item with the same characteristics, so that the life distribution of the item is enough to deduce all the properties of the item. Although it is very important to recognize renewal situations, it is seldom realistic for true life systems. Parts or non-repairable systems do, however, sometimes behave according to renewal processes. Some notes on renewal theory are presented in Section A.2.1.

(vii) *Homogeneous Poisson Process (HPP).* The details regarding the HPP are discussed in Section A.3.2. It is given as a special case of a renewal process in Figure 1.3 because it is numerically equivalent to the FOM of a renewal processes being represented by an exponential distribution. Other than this property, there is no relationship between the HPP and a renewal process.

(viii) *Distributions.* Distributions typically used to model renewal processes are presented as part of the discussion on renewal theory in Section A.2.2.

The outline in Figure 1.3 can be seen as a road map to the correct model-type and should always be used in failure data analysis. Guidelines for the appropriate selection of regression models are presented in by Kumar and Westberg (1996b) and are considered in Chapter 2.

University of Pretoria etd – Vlok, P-J (2006)

1.2.3 Statistical models in conventional failure time data analysis

In conventional failure time data analysis it is either assumed that an item is totally renewed after maintenance (GAN), i.e. perfect maintenance was done [‖] or that the item is in the same condition after maintenance as it was shortly before failure (BAO), i.e. minimal repair was done. The GAN property is modeled by zeroing an item's Force of Mortality (FOM) after renewal while the BAO assumption is represented by equating an item's *intensity* shortly before and shortly after failure. These concepts are introduced in the sections to follow.

1.2.3.1 Renewal models

Suppose the interarrival times of a system follow a distribution $f_X(x)$ with cumulative distribution $F_X(x)$. $F_X(x)$ is referred to as the *unreliability* function since it gives the probability of failure up to a certain age x, i.e. $F_X(x) = \Pr[X \leq x]$. Similarly, the *reliability* function, $R_X(x)$, is defined as $R_X(x) = \Pr[X \geq x]$ or $R_X(x) = 1 - F_X(x)$, i.e. the probability of survival up to age x. From this it is possible to define the force of mortality (FOM) or hazard rate of an item that gives the probability of failure within a short time, provided that the item survived up to that time, i.e. $h_X(x) = \Pr[x < X \leq x + \mathrm{d}x | X > x]$. The FOM can also be expressed as,

$$h_X(x) = \frac{f_X(x)}{1 - F_X(x)} \tag{1.1}$$

The FOM is further known as the *full intensity* or *conditional intensity* of the failure process of a non-repairable system. These concepts are defined in detail in Section 2.2. The FOM is often erroneously described as a conditional probability density function. The FOM is clearly not a conditional PDF because,

$$R_X(x) = e^{-\int_0^x h_X(\tau)\mathrm{d}\tau} \tag{1.2}$$

and since $R_X(\infty) = 0$ it implies that

$$\lim_{x \to \infty} \int_0^x h_X(\tau)\mathrm{d}\tau = \infty \tag{1.3}$$

For an increasing FOM, an item has an increasing probability to fail as time progresses and use based preventive renewal will be a definite option to consider, although cost will be the decisive factor. Preventive renewal will usually only be used if the total cost of a failure is considerably higher than the total cost of preventive actions. If equation (1.1) yields a constant risk, the component is said to have a random shock failure pattern because the risk of failure of the component remains the same throughout the item's life. Corrective

‖This could imply complete replacement.

University of Pretoria etd – Vlok, P-J (2006)

renewal will be the first option to consider for this case, i.e. a Repair Only On Failure (ROOF) strategy. A ROOF strategy will also most probably be used for a component with a decreasing FOM, since the probability of component failure becomes less as time increases. It should be kept in mind, however, that condition monitoring could be used for any shape of the FOM. The GAN assumption implies that the FOM is zeroed after every failure. Figure 1.4 illustrates this concept.

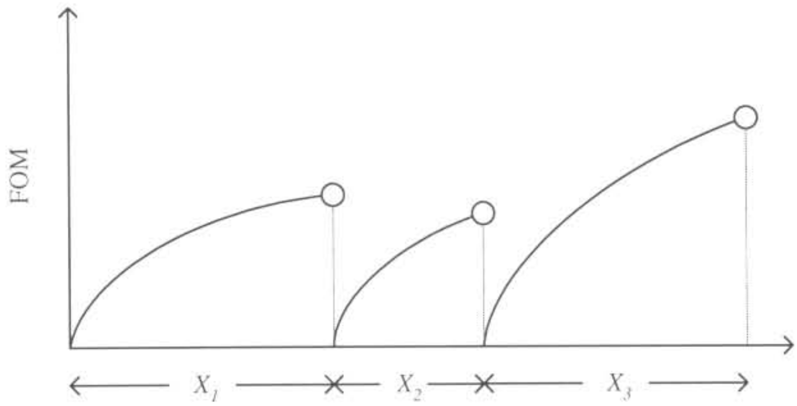

Figure 1.4: Illustration of the GAN assumption

Because of the assumption that interarrival times are part of an underlying distribution, only independent and identically distributed data sets can be used in renewal theory. This requirement is often totally ignored in the literature. In cases where the IID assumption holds, the Weibull distribution is usually most suitable to describe the data set because of its flexibility. Other distributions favored by analysts include the exponential, log-normal, log-logistic and normal distributions. Section A.2.2 gives more information about these distributions.

1.2.3.2 Models for repairable systems

For repairable systems it is assumed that the intensity of the failure process is equal shortly before and shortly after failure. To continue the discussion it is necessary to introduce the concept of intensity (also known as *full intensity* or *conditional intensity*) briefly at this point. This is done in detail in Section 2.2. The intensity of a counting process is generally defined as:

$$\iota(t) \equiv \lim_{\Delta t \to 0} \frac{\Pr\{N(t+\Delta t) - N(t) \geq 1 | H_t\}}{\Delta t} \tag{1.4}$$

where $N(t)$ is the observed number of failures in $(0, t]$ and H_t is the history up to, but not including, time t. Thus, $\iota(t)\Delta t$ is, for a small Δt, the approximate probability of an event in $[t, t+\Delta t)$, given the process history.

University of Pretoria etd – Vlok, P-J (2006)

In conventional repairable systems modeling it is assumed that processes are orderly, i.e. simultaneous failures cannot occur, and also stationary, which implies $\iota(t) = v(t)$, where $v(t)$ is the so called Rate of OCcurrence of Failure (ROCOF), given by

$$v(t) = \frac{d}{dt} E\{N(t)\} \tag{1.5}$$

The above mentioned simplifications make the NHPP a very suitable candidate for modeling the ROCOF ** of repairable systems. The following forms are encountered most frequently: (1) $\rho_1 = \exp(\Gamma + \Upsilon t)$ (log-linear) and (2) $\rho_2 = \kappa\beta t^{\beta-1}$ (power-law) or even a constant ROCOF. A few authors that used these models are Balakrishnan (1995), Shin, Lim, and Lie (1996), Hokstad (1997), Jensen (1990), Ledoux and Rubino (1977), Kobbacy, Percy, and Fawzi (1994) and Hasser, Dietrich, and Szidarovszky (1995). Figure 1.5 illustrates the BAO assumption for an item.

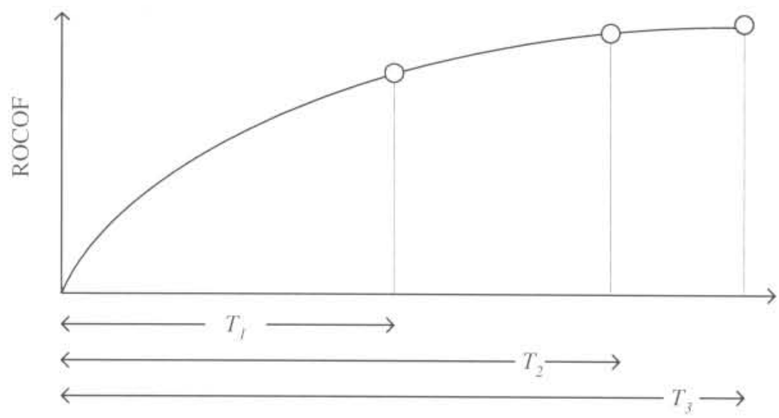

Figure 1.5: Illustration of the BAO assumption

Even though the BAO assumption is much more realistic than the GAN assumption, it could still be a very limited approach according to Ascher and Feingold (1984), since in practice the highest probability of failure is often directly after maintenance.

1.2.4 Conventional replacement/repair cost optimization models

Conventional cost optimization models strives to minimize long term operational cost of equipment by using the statistical models mentioned above. This optimum is often referred to as the minimum Life Cycle Cost (LCC) of an item. The term LCC could be somewhat confusing in this context since it is commonly used in capital replacement studies where the total cost of ownership is taken into account, including operation and maintenance cost, the time value of money, depreciation, etc. To be consistent with the majority of literature in

**The ROCOF of an NHPP is referred to as the *peril* rate and is denoted by $\rho(t)$.

University of Pretoria etd – Vlok, P-J (2006)

this field, the term LCC will also be used in this document even though only operational costs are considered. In this section, some examples of different approaches are presented to explain the concept.

1.2.4.1 Optimization models for renewal situations

Here, the risk of wasting residual life is balanced with the risk of suffering an expensive unexpected failure in terms of cost. At the point of balance, the LCC per unit time will be a minimum. The costs involved are C_p, the cost of preventive replacement (or renewal) and C_f, the cost of unexpected failure.

The principle of these models is fairly simple to understand. Suppose a component is always replaced at time X_p or at failure time X, whichever comes first. The total cycle cost is then given by $C_pR_X(X_p) + C_f[1 - R_X(X_p)]$. If it is assumed that it takes a time units to perform preventive action and b time units to perform corrective maintenance, the expected duration of the component's life is $(X_p+a)R_X(X_p)+(X+b)[1-R_X(X_p)]$. Division yields the following relation for component cost per unit time (if the replacement rule is followed):

$$C(X_p) = \frac{C_pR_X(X_p) + C_f[1 - R_X(X_p)]}{(X_p + a)R_X(X_p) + (\int_0^{X_p} x \cdot f_X(x)dx + b)[1 - R_X(X_p)]} \tag{1.6}$$

The minimum cost is found where $dC(X_p)/dx = 0$. (See Jardine (1973) for details). For example, suppose a data set is described by a Weibull distribution with $\beta = 2.5$ and $\eta = 200$. Also, assume C_p = R 5 000 (with a = 2h) and C_f = R 20 000 (with b = 8h), then equation (1.6) will yield the graph in Figure 1.6. This graph shows that there is a clear optimum at around 111 days, i.e. R 77 per unit time .

Using the same methodology as above, a relation can be derived to optimize availability instead of cost. It is also possible to calculate the optimum preventive replacement frequency for component-blocks rather than for single components.

Many authors have made some minor refinements to the conventional optimization models for components, often to adapt to data constraints. Overviews of these refined models can be found in Sherif and Smith (1981), Aven and Dekker (1997), Aven and Bergman (1986), Dekker (1995), Zijlstra (1981), Sherwin (1999), Van Noortwijk (2000) and Schäbe (1995). A noteworthy extension of these models, is the model of Ran and Rosenlund (1976) in which the time value of money is taken into account. This model is obviously only useful in cases where equipment is expected to survive for several years.

University of Pretoria etd – Vlok, P-J (2006)

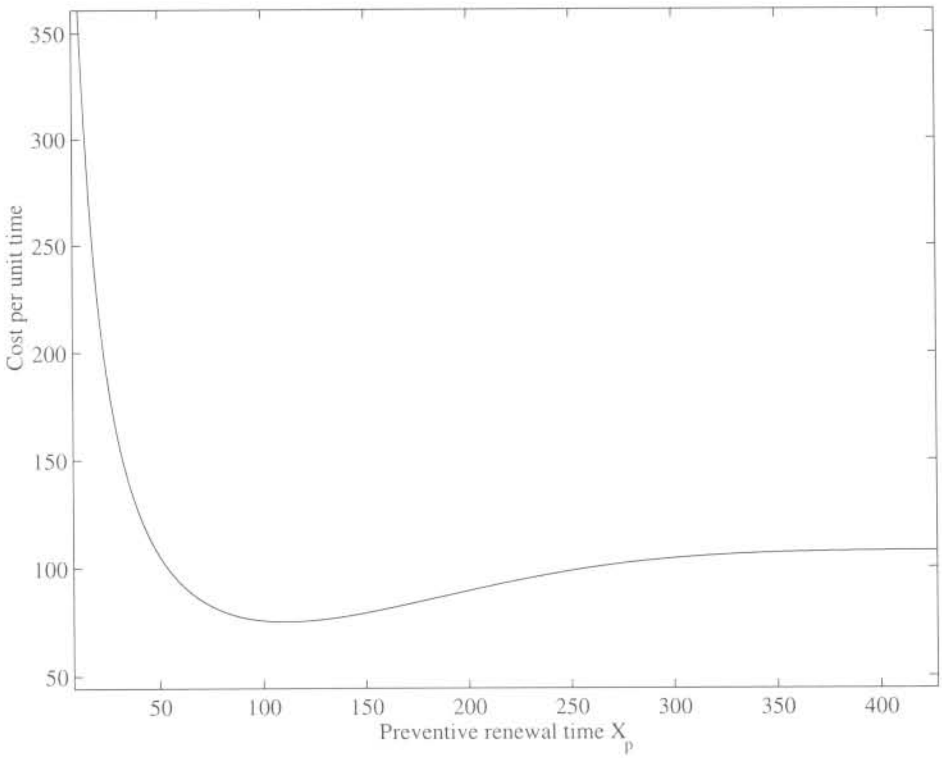

Figure 1.6: LCC of an item renewed after X_p time units or at failure (if $X < X_p$)

1.2.4.2 Optimization models for repairable systems

For repairable systems, a decision has to be made between minimal repair or complete replacement of a system when it has failed. The cost of minimal replacement, C_1, is expected to be considerably less than that of system replacement, C_2. It is hence required to calculate the number of minimal repairs that should be allowed before system replacement with the objective to minimize the LCC. The optimal solution can be expressed in terms of the number of minimal repairs, n, or as time, I. Ascher and Feingold (1984) showed that if the power-law process, $\rho_2 = \kappa\beta t^{\beta-1}$, is used to model the ROCOF of a process modeled by an NHPP, the optimal replacement time will be,

$$I^* = \left[\frac{C_2}{C_1 \cdot (\beta - 1) \cdot \kappa}\right]^{1/\beta} \tag{1.7}$$

and the optimal number of minimal repairs before complete replacement is:

$$n^* = \frac{C_2}{C_1 \cdot (\beta - 1)} \tag{1.8}$$

where I^* and n^* are the optimal solutions. Suppose $\beta = 1.7$ and $\kappa = 0.0015$ with $C_1 = \text{R}500$ and $C_2 = \text{R}8,000$, then $I^* \approx 288$ at R 67 per unit time and $n^* \approx 22$ at R 65 per unit time. Figures 1.7 and 1.8 show these results graphically. The costs per unit time resulting from the two policies above are often very similar except in situations where $C_1 \approx C_2$ (which is rare).

University of Pretoria etd – Vlok, P-J (2006)

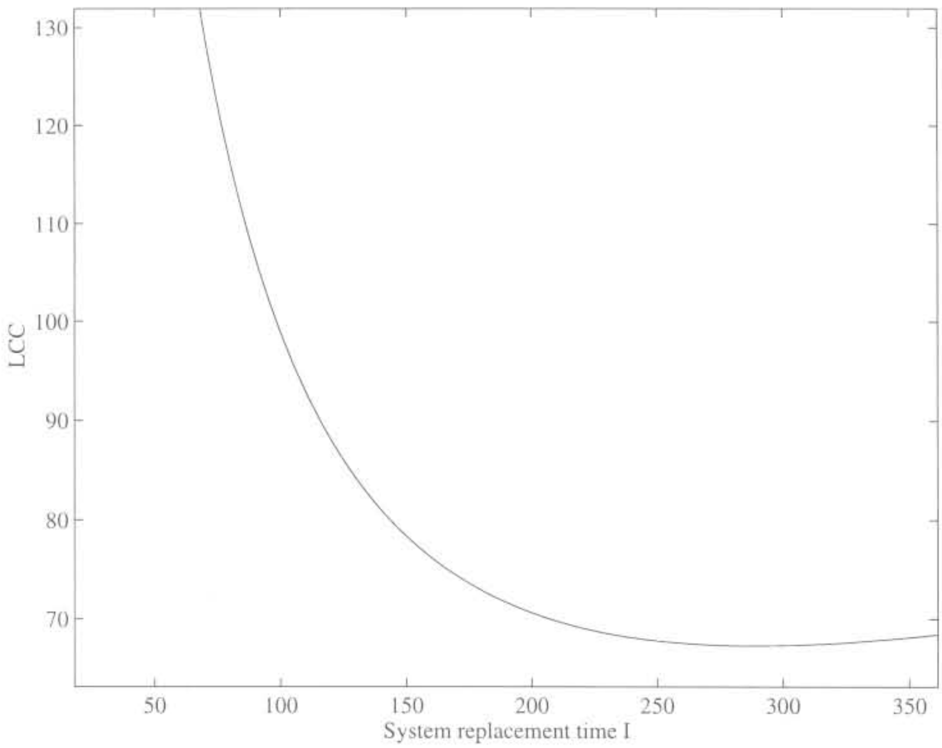

Figure 1.7: LCC of a system minimally repaired up to I* time units

Several minor improvements to minimal repair/replacement policies have been proposed since the introduction of conventional statistical failure analysis. For some early references see Barlow and Hunter (1960), Ross (1969), Morimura (1970) and Park (1979). More recent examples include Stadje and Zuckerman (1991), Yeh (1991), Lam and Yeh (1994), Hsu (1999), Sheu (1999) and Lim and Park (1999).

1.2.5 Shortcomings of conventional approaches

Limitations of conventional approaches do not so much lie in the techniques themselves but rather in the underlying assumptions. The renewal (GAN) assumption is probably the most unrealistic of the two assumptions discussed above. Deterioration of a system may influence the lifetimes of future components in certain sockets severely, even if components are completely renewed/replaced. Renewal theory deals with an important data type however, and certainly has its place in theory even though it is seldom practical. The minimal repair (BAO) assumption is much more realistic than the GAN assumption but still not completely practical. Human interference to improve the condition of a system is often the greatest cause of maintenance - a fact that the BAO assumption does not take into account.

Many authors have proposed models with discontinuities to incorporate the BOWN or WO

University of Pretoria etd – Vlok, P-J (2006)

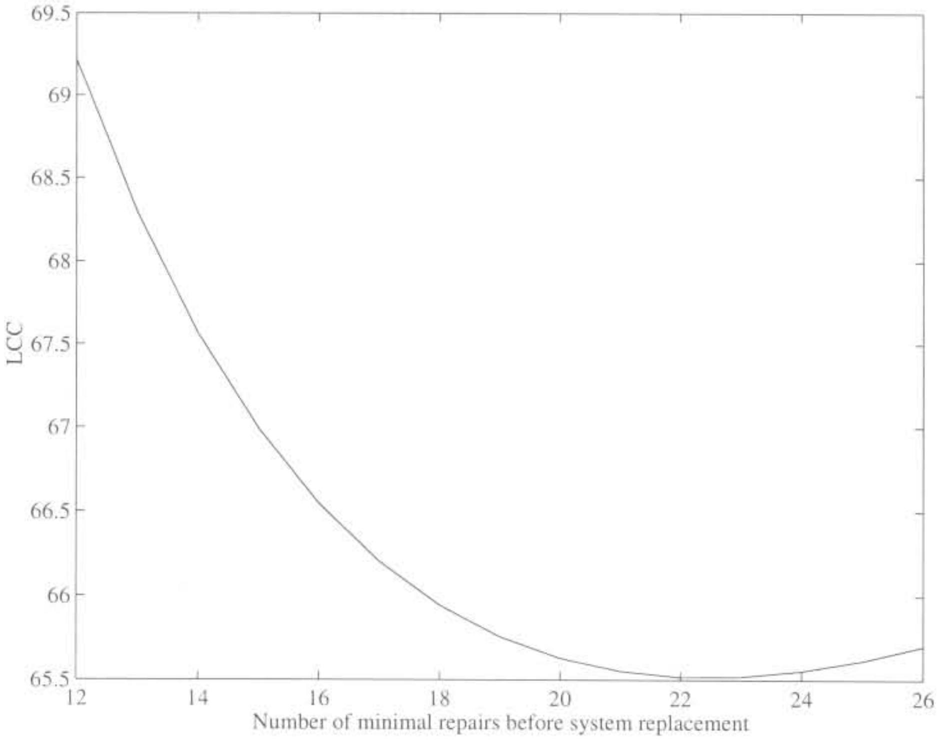

Figure 1.8: LCC of a system minimally repaired up to n failures

situations. These models are major improvements on the conventional approaches although seldom utilized in practice, mainly because of complexity. Models with discontinuities are discussed in Chapter 2.

The biggest shortcoming of models used for conventional failure time data analysis is their inability to include concomitant information in analyses. Diagnostic information recorded during lifetimes, such as Condition Monitoring (CM) results, is not included in models which certainly limits the accuracy of predictions immensely. Regression models solves this problem to a great extent because diagnostic information can be included in the form of covariates. Regression models are described extensively in chapters to follow.

A further serious disadvantage of conventional approaches is the long term nature of replacement/repair policies. Costs only converge to the statistical optimum after a few lifetimes and the minimum LCC approach is often rejected due to the impatience of maintenance practitioners. It is often also very difficult to estimate realistic values for C_p, C_f, C_1 and C_2 for use in the policies outlined in the previous section. The LCC is further poorly understood in industry and the optimum is commonly interpreted as a prediction of time to failure.

University of Pretoria etd – Vlok, P-J (2006)

1.3 Preventive maintenance optimization through Condition Monitoring

Condition Monitoring has become increasingly popular in recent times. One reason for this is the present affordability of specialized condition monitoring equipment and the perception that advanced technology can solve all maintenance problems. A further reason is that maintenance strategy setting methodologies, such as RCM, recommend an on-condition task as default strategy, provided that the task is technically and economically feasible. (See Nolan and Heap (1978) for details on RCM). These and other factors contribute to a (often erroneous) drive towards condition monitoring in industry.

An item's condition can be assessed much better at present than a few years ago with technology of the day incorporated into techniques such as vibration analysis, oil analysis and thermography. This however does not imply that these techniques are perfect. A general investigation into typical condition monitoring practices revealed several shortcomings, which are discussed below.

1.3.1 Alarm trigger setting

CM techniques assess an item's condition in its present operating state and a maintenance decision has to be made based on the observed diagnostic information. This implies than limits have to be set for measured parameters and once one or more of the limits are exceeded (triggered), preventive action should take place. This may seem simple, but setting appropriate benchmarks is no trivial procedure.

Original Equipment Manufacturers (OEM's) often give guidelines as to what is acceptable operating conditions for equipment in terms of temperature, vibration, oil debris, etc. These guidelines usually form the basis of benchmarks although it is normally very conservative for obvious reasons. Initial benchmarks can then only be optimized through a trial and error approach that may be very expensive.

Many algorithms / techniques have been proposed by researchers in the various CM fields to determine optimal benchmarks in a process to eliminate trial and error approaches. These algorithms have one common underlying principle: to learn from observed diagnostic measurements taken in the past and then estimate optimal benchmarks in a scientific manner for a piece of equipment currently in operation. The most successful of these techniques is neural networks. Neural networks have a large appeal to many researchers due to their great closeness to the structure of the brain, a characteristic not shared by other modeling techniques. In an analogy to the brain, an entity made up of interconnected neurons, neural networks are made up of interconnected processing elements called units, which respond in parallel to

University of Pretoria etd – Vlok, P-J (2006)

a set of input signals given to each. The unit is the equivalent of its brain counterpart, the neuron.

A neural network consists of four main parts:

(i) Processing units, where each processing unit has a certain activation level at any point in time.
(ii) Weighted interconnections between the various processing units which determine how the activation of one unit leads to input for another unit.
(iii) An activation rule which acts on the set of input signals at a unit to produce a new output signal, or activation.
(iv) Optionally, a learning rule that specifies how to adjust the weights for a given input/output pair.

Time failure data with CM information can be used as processing units to estimate and teach neural networks and additional data can then be used as inputs to predict future outputs. Recent attempts to apply neural networks in the reliability modelling field include Shyur and Luxhoj (1995), Rawicz and Girling (1994) and Lakey (1993). Neural networks have not made much ground in the field of reliability because of its general complexity, large data set requirements and its inability to eliminate insignificant observations.

Setting appropriate alarms for CM parameters is no easy task and CM techniques are seldom optimal from implementation. This is a significant shortcoming in the field of condition monitoring.

1.3.2 Significance of observed parameters

CM techniques use several parameters to assess an item's condition. This may be a frequency spectrum in vibration monitoring, a range of temperatures in thermography, the quantity of various foreign elements in an oil sample, etc. In some instances different CM techniques are combined to estimate equipment reliability. The reason for using more than one parameter is because it is very seldom obvious which parameter is the best indicator of approaching failure and no general technique exists in contemporary CM to isolate significant parameters. The inability of CM techniques to isolate significant parameters is closely related to the alarm trigger limit issues outlined in the previous section.

University of Pretoria etd – Vlok, P-J (2006)

1.3.3 Lack of commitment towards CM

In general, there is a lack of commitment towards condition monitoring in the South African industry. In many cases, expensive CM equipment is used as the flagship of maintenance departments although inspections are done very irregularly and not recorded properly. Often the information supplied by CM is totally disregarded when a decision has to be made and experience or intuition is relied on. Even if CM information is considered, the final decision is frequently left to the discretion of technicians involved with the equipment.

It does not matter how technologically advanced CM is, if it is not practiced correctly, meaningful results are impossible to obtain. This is a maintenance management issue that is not directly addressed in this thesis.

1.4 Combining use based preventive maintenance optimizing techniques with CM technology

From the discussions above it follows that use based preventive maintenance optimization techniques complement CM technology extremely well. A technique that combines these strategies would have enormous potential. The solution lies in statistical regression models since this type of model allows for concomitant information with time to event data - in this context the concomitant information could be diagnostic information recorded by CM techniques.

Several regression models have been applied in reliability to estimate the risk of failure of an item and most of these models are discussed in the next chapter. Only the Proportional Hazards Model (PHM) is discussed in this section as an introduction to regression models, but also because this is the only regression model for which a scientific preventive maintenance decision model exist.

1.4.1 Proportional Hazards Modeling

The PHM was introduced by Cox (1972) and was considered to be a total revolution in survival analysis. This model was intended for the field of biomedicine but became increasingly popular in reliability modeling over the past two decades. The model uses a baseline hazard rate and allows a functional term containing covariates to act multiplicatively on the baseline hazard rate (or FOM), i.e.

$$h(x, \boldsymbol{z}) = h_0(x) \cdot \lambda(x, \boldsymbol{z}(x)) \tag{1.9}$$

University of Pretoria etd – Vlok, P-J (2006)

where h_0 is the baseline FOM, λ is the functional term and $\boldsymbol{z}$ is a vector of covariates which may be time-dependent. Kumar and Klefsjo (1993) summarized the assumptions of the PHM as follows:

(i) Event data is IID.

(ii) All influential covariates are included in the model.

(iii) The ratio of any two FOMs as determined by any two sets of time-independent covariates $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ associated with a particular item has to be constant with respect to time, i.e. $h(x, \boldsymbol{z}_1) \propto h(x, \boldsymbol{z}_2)$. For time-dependent covariates, this assumption is not defined.

The exponential function is used most often for the functional term. This leads to a semi-parametric model. It is possible to calculate the semi-parametric model without making any assumption on the baseline hazard rate but this only yield relative risks. In reliability, the absolute risk is usually required and the model is hence parameterized by specifying some parametric FOM for the baseline, for example the Weibull FOM, i.e.

$$h(x, \boldsymbol{z}) = \frac{\beta}{\eta} \cdot \left(\frac{x}{\eta}\right)^{\beta-1} \cdot \exp\left(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)\right) \quad (1.10)$$

where β and η are the Weibull shape and scale parameters respectively and $\boldsymbol{\gamma}$ is a vector of regression coefficients. The influence of the functional term results in an improved estimate of an item's FOM. Figure 1.9 illustrates this concept.

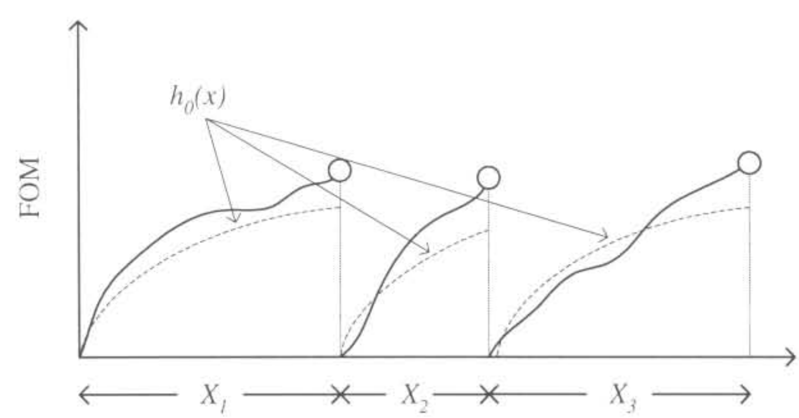

Figure 1.9: Illustration of the PHM with time-dependent covariates

The PHM has been applied successfully in diverse reliability applications because of the improved estimate of the FOM, including modeling component failures in a light water reactor plant by Booker et al. (1981), marine gas turbine and ship sonar by Ascher (1983), motorrettes by Dale (1985), aircraft engines by Jardine and Anderson (1988), high speed train

University of Pretoria etd – Vlok, P-J (2006)

brake discs by Bendell et al. (1986), sodium sulfur cells by Ansell and Ansell (1987), surface controlled subsurface safety valves by Lindqvist et al. (1988) and machine tools by Mazzuchi and Soyer (1989). Other authors that published applications of the PHM in reliability include Jardine et al. (1989), Leitao and Newton (1989), Love and Guo (1991a) and Love and Guo (1991b).

The biggest criticism of the PHM is the fact that it is by definition only applicable for IID data. This shortcoming can be addressed by allowing for imperfect repair in the covariates, but this does not solve the problem completely and in some cases can even worsen the situation. Some authors, for example Kumar (1996), have applied the PHM with reasonable success on repairable systems, despite the requirement of IID data.

1.4.2 Decision making with the PHM

Estimating the optimum maintenance instant that will result in the minimum LCC of an item, based on the FOM as determined by the PHM, is no trivial procedure since the FOM is now dependent on time and the values of covariates. This implies that the optimum LCC instant must be specified in terms of risk and not in terms of a process parameter, such as time, as was described in Section 1.2.4.1.

Two attempts to calculate the optimal maintenance instant for a system with the PHM were found in the literature. The first was by Kumar and Westberg (1996a) that used the PHM together with Total Time on Test (TTT) plotting to estimate the optimum maintenance frequency. This paper was not very case-orientated and is, as far is known, the only of its kind. The second consists of a series of publications by, amongst others, Makis and Jardine. These authors have developed a technique for calculating the minimum LCC in terms of a system's risk as determined by the PHM. Makis and Jardine (1991) and Makis and Jardine (1992) proposed a semi-Markov approach to calculate the minimum LCC where covariate behavior is predicted by semi-Markov chains. Makis and Jardine's technique was then refined in several publications to follow, the most important being Banjevic, Ennis, Braticevic, Makis, and Jardine (1997) and Jardine, Banjevic, and Makis (1997).

Makis and Jardine's optimization technique produces a result that looks very similar to Figure 1.6, except that the cost is expressed as a function of $h(x, \boldsymbol{z})$. It is then required to allow an item to operate until the optimum risk level (as opposed to time) is reached before preventive action is taken. Figure 1.10 illustrates the policy in two dimensions with imaginary inspection data. The figure shows how the optimal risk is influenced by both time and the observed level of covariates.

Examples of successful applications of this replacement policy include Vlok (1999) and Jardine, Banjevic, and Makis (1997).

University of Pretoria etd – Vlok, P-J (2006)

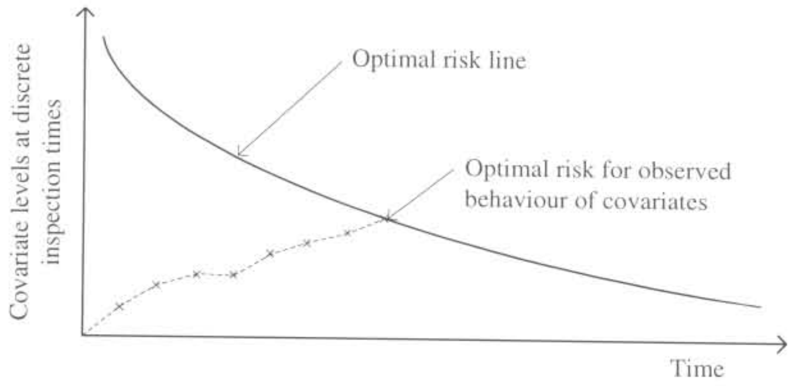

Figure 1.10: Illustration of the optimal policy with imaginary covariate levels

1.4.3 Shortcomings of PHM cost optimization

Even though the PHM cost optimization approach is an improvement on conventional techniques, there are still two major shortcomings:

(i) The PHM has an underlying assumption that data is IID. This limiting assumption can be overcome to a certain extend by including covariates that describe an item's failure history. The fact remains, however, that the model is not entirely suitable for repairable systems data and repairable systems data is expected much more often than IID data.

(ii) The minimum LCC cost approach for the PHM is a long term type approach, as is the case with conventional analysis techniques. This approach is also not accepted well amongst maintenance practitioners because the minimum is only reached after a few lifetimes of which some could be expensive unexpected failures.

Both the above-mentioned shortcomings should be addressed in order to make a truly valuable contribution to the field of statistical failure analysis.

1.5 Residual life

Maintenance would be a trivial affair if the exact times to failure of items were known. This would imply that no residual life is wasted and that expensive unexpected failures are totally eliminated. Although this view seems only feasible in a perfect world, the approach is certainly meritorious in a world striving for perfection. A scientific technique with the ability to estimate the residual life of equipment will be of great advantage to the field of maintenance engineering.

University of Pretoria etd – Vlok, P-J (2006)

In CM, some empiric methods exist to estimate residual life. These methods are seldom generalized and are often only meaningful after many iterations. The methods also differ from situation to situation, even for nominally similar items, which makes it very risky to use a particular method to predict residual life.

For conventional renewal analysis, residual life estimation (in principle) only goes as far as conditional expectation or mean. If a distribution, $f_X(x)$, describes an item's failure history, the residual life, $\mu(x)$, can be calculated by

$$\mu(x) \equiv E[X - x | X \geq x] = \frac{\int_x^\infty (\tau - x) f_X(\tau) d\tau}{R_X(x)} = \frac{\int_x^\infty R_X(\tau) d\tau}{R_X(x)} \tag{1.11}$$

The simple statistical mean life, μ, of the item is given by

$$\mu = \int_0^\infty x \cdot f_X(x) dx \tag{1.12}$$

It is important to note that any differentiable $\mu(x)$ has to satisfy $\mu(x) \geq -1$ because of the identity

$$h_X(x) = \frac{\frac{d\mu(x)}{dx} + 1}{\mu(x)} \tag{1.13}$$

as described by Muth (1977). Ghai and Mi (1999) discussed mean residual life and its association with the FOM in detail. Other authors that worked on this subject include Tang, Lu, and Chew (1999), Baganha, Geraldo, and Pyke (1999) and Guess and Prochan (1988).

Equally little work has been done on estimating the residual life of items with conventional repairable systems theory. Calabria, Guida, and Pulcini (1990) proposed a point estimation procedure for future failure times of a repairable system modeled by a NHPP with a power intensity law. Suppose a repairable system has suffered n failures and it is required to estimate the $(n+m)^{\text{th}}$ failure, where $m \geq 1$ and $m \in \mathbb{Z}$. The Maximum Likelihood Estimate (MLE) of the expected value of the m^{th} future failure is given by

$$E_m = (n-1) \cdot \gamma \int_{t_n}^\infty \sum_{j=1}^m C_j \frac{n+j-1}{n} \left[1 + \frac{n+j-1}{n} \gamma \ln(t_{n+m}/t_n)\right]^{-n} dt_{n+m} \tag{1.14}$$

where $C_j = \prod_{i \neq j}^m (n+i-1)/(i-j)$. Schäbe (1995) followed a similar approach, as did Reinertsen (1996).

The theory above shows that conventional statistical failure analysis only yields a *mean* residual life estimate. This fact makes the use of residual life estimates very unpopular and unreliable in practice. A dynamic residual life estimate is required to be useful in practice, i.e. a technique that will adjust estimates based on certain observed influences. Statistical models that have the ability to incorporate concomitant information immediately seem to be a possible solution even though very few publications on this subject exist. Zahedi (1991)

i 15935978
b15358288

University of Pretoria etd – Vlok, P-J (2006)

proposed a proportional mean remaining life model analogous to the PHM where a baseline survivor function is influenced by a functional term containing covariates. No publication was found where this model was applied on real life survival data, however. Other contributions to multivariate residual life estimation include Nair and Nair (1989), Arnold and Zahedi (1988) and Zahedi (1985). In neither of these publications, practical illustrations of the theory were presented.

1.6 Problem statement

There is a need to optimize preventive maintenance decisions in today's ever increasingly competitive market. At present there are three established means for doing this namely, conventional statistical failure analysis, condition monitoring and Proportional Hazards Modeling. The following shortcomings were identified for the respective techniques:

(A) *Conventional failure analysis*

A-1. Only allows for the GAN or BAO assumption, which is extremely limiting.

A-2. Lack of ability to include concomitant information in the analyses.

A-3. Requires fixed estimates for C_p and C_f, which often varies for every failure.

A-4. The long term nature of optimal replacement/repair policies is often rejected by maintenance practitioners because unexpected failures are regarded as unacceptable.

(B) *Condition Monitoring*

B-1. It is very difficult to set optimal initial alarm trigger settings for CM techniques.

B-2. No scientific technique exists with which the significance of CM parameters can be calculated.

B-3. There is a general lack of managerial commitment to CM.

(C) *Proportional Hazards Modeling*

C-1. Assumes data to be IID.

C-2. The only replacement decision model found for the PHM is also based on costs and requires a few lifetimes before it converges to the minimum cost.

This thesis aims at improving all nine shortcomings listed above. It is proposed that this objective can be reached as follows:

University of Pretoria etd – Vlok, P-J (2006)

(i) *Development of a combined Proportional Intensity Model (PIM) with the ability to address all the model-related shortcomings mentioned above*

It is proposed that a combined PIM, one for non-repairable and one for repairable systems, is developed that will include the majority of conventional PIM enhancements as special cases (including the PHM) to be able to model most of the typical wear-out/deterioration patterns found amongst industrial equipment. Such a PIM would be able to accommodate discontinuities in the failure intensity and to adapt to discontinuities or to scalings in its time scale which will be ideal for the WO and BOWN scenarios. By developing the combined PIM, shortcomings A-1, A-2, B-2 and C-1 will be addressed.

(ii) *Development of an algorithm to calculate residual life of an item based on the combined PIM*

A flexible and adaptive combined PIM will theoretically lead to a close representation of reality and hence realistic estimates of the residual life, provided that the future behavior of covariates can be estimated with relative high certainty. This could be a challenging task since very little work has been done in this field and the numerical implementation of the theory is fairly complicated. Successful completion of this goal would solve shortcomings A-3, A-4, B-1, B-2 and C-2.

(iii) *Comprehensible presentation of results*

To make a truly practical contribution to the field of reliability modeling, results produced by this study should be presented in a user-friendly and comprehensible manner. This step is required to address shortcoming B-3.

1.7 Thesis outline

In Chapter 2, a literature survey of advanced failure intensity models (including PIMs) in survival analysis is done. Terminology used in failure intensity models is defined and different models are categorized and evaluated. Chapter 2 serves as the foundation for the development of the combined PIMs in Chapter 3. In Chapter 3 the combined PIMs are derived and it is illustrated how these models can be reduced to most conventional PIMs. Parameter estimation techniques based on maximum likelihood are also discussed in Chapter 3. In Chapter 4 conventional techniques are applied to the combined PIMs to estimate residual life. Confidence bounds on estimates are also discussed. Chapter 5 contains a case study in which the theory developed in this thesis is applied to a typical data set from a South African industry. Results are compared to results obtained from a maintenance decision support tool similar to the residual life approach. In Chapter 6 the findings of this thesis are summarized with some recommendations for future research.

University of Pretoria etd – Vlok, P-J (2006)

Chapter 2

Advanced Failure Intensity Models

2.1 Introduction

Advanced failure intensity models are in this thesis defined as mathematical representations of failure processes that require more than standard distributions or 2-parameter counting process models to capture their characteristics. This chapter deals with advanced failure intensity models found in the literature.

Chapter 2 starts off with a discussion of the concept of *intensity* with specific reference to non-repairable and repairable situations. The importance of the difference between these situations cannot be overemphasized even though it is frequently ignored in statistical failure analysis. A clear notation with regards to intensities is defined in Section 2.2 and used throughout this thesis. Deviations from the notation are explicitly indicated. Different model classes are identified and relevant models are discussed. Some acclaimed applications of advanced failure intensity concepts are also considered . For most models, the likelihood or partial likelihood are derived or presented without describing the estimation of regression parameters. Parameter estimation techniques are considered in Chapter 3.

The chapter ends with a summary of the advantages and disadvantages of the models considered.

2.2 Intensity Concepts

The concept of intensity was introduced briefly in Section 1.2.3.2. In this section, the concept is explained in detail since all reliability models discussed in this chapter strive to represent the intensity of a certain failure process. It is assumed throughout the thesis that all failure processes considered are orderly, i.e. simultaneous failures cannot occur on the same item.

University of Pretoria etd – Vlok, P-J (2006)

This is a reasonable assumption according to most authors, e.g. Hokstad (1997) and Lawless (1987), and not much generality is sacrificed.

Let $N(t)$ denote the number of failures an item has experienced in the interval $(0, t]$. The *unconditional* intensity (i.e., the rate of failure events) of the process at any instant in time, t, is then given by

$$\begin{aligned} \iota_u(t) &= \lim_{\Delta t \to 0} \frac{\Pr[\text{Failure occurs in } [t, t+\Delta t)]}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{E[\Delta N(t)]}{\Delta t} \end{aligned} \tag{2.1}$$

where $\Delta N(t)$ represents the increment $N(t + \Delta t) - N(t)$. Because it is assumed that the process is orderly, the following basic relation for counting processes applies:

$$\iota_u(t) = \frac{\mathrm{d}M(t)}{\mathrm{d}t} \tag{2.2}$$

with $M(t) = E[N(t)]$. The time derivative of the expected number of failures as in (2.2), is referred to as the rate of occurrence of failure (ROCOF) and will often be denoted by $\rho(t)$ (instead of $\iota_u(t)$), for convenience. From (2.2) it follows directly that the cumulative number of failures up to time t is equal to the cumulative unconditional intensity, i.e.

$$M_u(t) = E[N(t)] = \int_0^t \iota_u(u)\mathrm{d}u \tag{2.3}$$

Additional information about the failure process is often recorded with the times to failure. The additional information is referred to as the *history*, H_t, or *filtration* of the process. History is recorded in the form of covariates and could be any quantification of an influence on the failure process. From Martingale theory (see Hokstad (1997)) it follows that H_t is the σ-algebra generated by $N(s)$, $s \leq t$, starting from a probability space (Ω, H_t, P) that defines the stochastic process, $N(t) = N(t, \omega)$, with $\omega \in \Omega$. Hence it is possible to define the *full intensity* (also referred to simply as *intensity* or *conditional intensity*), $\iota(t, H_t) = \iota(t|H_t)$, which is the conditional rate of occurrence of events, given the state of H_t. Thus, $\iota(t, H_t)\Delta t$ is the probability of an event to occur in $[t, t + \Delta t)$, i.e.

$$\iota(t, H_t) = \lim_{\Delta t \to 0} \frac{E[\Delta N(t)|H_t)}{\Delta t} \tag{2.4}$$

The complete intensity as defined in (2.4) provides a general framework for modeling failure event processes because the effect of maintenance activities can be recorded in H_t. Conventional failure process modeling concepts such as the FOM and ROCOF are also special cases of (2.4).

Similar to (2.3), it is possible to define a cumulative intensity process, i.e.

$$M(t, H_t) = \int_0^t \iota(u, H_t)\mathrm{d}u \tag{2.5}$$

University of Pretoria etd – Vlok, P-J (2006)

where $M(t, H_t)$ is the *compensator* in Martingale theory. Both $\iota(t, H_t)$ and $M(t, H_t)$ are denoted as *predictable* which means that for a given H_t, the values of $\iota(t, H_t)$ and $M(t, H_t)$ are known but the value of $N(t)^*$ not yet.

It is important to note that $\iota_u(t)$ is a mean function of $\iota(t, H_t)$, averaged over all possible sample paths. Suppose $N(t, \omega)$ is a specific realization of the process of $N(t)$ where $\omega \in \Omega$ in the probability space (Ω, H_t, P). Here, N is not only a function of ω for a fixed value of t but also a function of t for a fixed ω (called the sample path of N). Taking the the mean over the sample space, Ω, yields

$$E[\Delta N(t)] = \int_{\Omega} \mathrm{d}N(t, \omega)\mathrm{d}P(\omega) = i_u(t) \tag{2.6}$$

Similarly, $E[\Delta N(t)|H_t]$, and thus $\iota(t, H_t)$, is found as the conditional mean.

The last intensity concept to define is that of *average intensity*. Average intensity is simply the average of $M_u(t)$ or $M(t, H_t)$ over an interval $[0, \tau]$, i.e. $\iota_{u_\tau} = M_u(\tau)/\tau$ or $\iota_\tau = M(\tau)/\tau$. The concept of average intensity is not encountered frequently in the literature but is not without interest. Bodsberg and Hokstad (1995) have shown that the average intensity concept is very useful in modeling dormant failures.

Table 2.1: Summary of failure intensity concepts

	Failure Intensity Concept		
	Intensity	**Mean Intensity**	**Average Intensity**
Alternative term	Conditional intensity	Unconditional intensity	-
Symbol	ι	ι_u	ι_τ
Definition	$\lim_{\Delta t \to 0} \frac{E[\Delta N(t)\|H_t)}{\Delta t}$	$\lim_{\Delta t \to 0} \frac{E[\Delta N(t)]}{\Delta t}$	$t^{-1} \cdot E[N(t)]$
Non-repairable case	$h_X(x)$, truncated at time of failure	$f_X(x)$	$x^{-1} \cdot (1 - R_X(x))$
Repairable case	A sequence of truncated FOMs (defined in local time)	ROCOF or $v(t)$	Average ROCOF, i.e. AROCOF

In Table 2.1, a concise summary (adapted from Hokstad (1997)) of the failure intensity concepts discussed in this section is presented. Note that local time, denoted by x, is used as time scale for the non-repairable case, consistent with the terminology introduced in Section 1.2.1.

*$N(t)$ has right continuous sample paths.

University of Pretoria etd – Vlok, P-J (2006)

2.3 Literature survey on advanced failure intensity models

There are countless publications on advanced failure intensity models attempting to represent the intensity concepts outlined in Table 2.1 as part of practical statistical failure analysis exercises. Most of these publications consider variations on a small number of fundamentally different approaches. The fundamentally different approaches are referred to as model classes and are listed below:

(i) Multiplicative intensity models
(ii) Additive intensity models
(iii) Models with mixed or modified time scales
(iv) Marginal regression analysis
(v) Competing risks
(vi) Frailty or mixture models

These model classes are discussed in Section 2.3. Publications that consider combinations of two or more model classes are discussed as part of the model class where it makes the most significant contribution. At the end of this section, noteworthy extensions of the listed model classes are also discussed.

2.3.1 Multiplicative Intensity Models

Multiplicative intensity models represent the intensity of a failure process as the product of a baseline intensity, that is a function of time only, and a functional term, that may be a function of both time and covariates. Covariates are allowed to be time-independent or time-dependent.

2.3.1.1 Proportional Hazards Model (PHM)

Survival data analysis underwent a revolution with the introduction of the PHM by Cox (1972). The model was originally intended for biomedical applications but was soon applied in reliability engineering. As the name implies, this model represents the FOM, i.e. the failure intensity of non-repairable items, as a proportion of different FOMs.

The PHM is constructed as the product of a totally arbitrary and unspecified baseline FOM, $h_0(x)$, and a functional term $\lambda(x, \boldsymbol{z})$, where $\boldsymbol{z}$'s dependence on time is not important, i.e. †

$$h(x, \boldsymbol{z}) = h_0(x) \cdot \lambda(x, \boldsymbol{z}(x)) \tag{2.7}$$

†The subscript x is dropped here for notational convenience.

University of Pretoria etd – Vlok, P-J (2006)

There are several possible forms for the functional term. Some are: the exponential form, $\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}(x))$; the logarithmic form, $\log(1+\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}(x)))$; the inverse linear form, $1/(1+\boldsymbol{\gamma}\cdot\boldsymbol{z}(x))$; or the linear form, $1+\boldsymbol{\gamma}\cdot\boldsymbol{z}(x)$, where $\boldsymbol{\gamma}$ is a vector of regression coefficients associated with a particular data set. The exponential form of the functional term is used most often in reliability applications and results in the following PHM:

$$h(x,\boldsymbol{z}) = h_0(x)\cdot\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}(x)) \tag{2.8}$$

The model assumes the following:

(i) Event times are IID.
(ii) All influential variables are included in the model.
(iii) The ratio of any two hazard rates as determined by any two sets of time-independent covariates $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ associated with a particular item has to be constant with respect to time, i.e. $h_X(x,\boldsymbol{z}_1) \propto h_X(x,\boldsymbol{z}_2)$. (This assumption is not valid for time-dependent covariates).

The biggest advantage of the PHM, as defined in (2.8) in its semi-parametric form, is that no assumption needs to be made about the baseline FOM when fitting the model. This is a result of partial likelihood theory developed by Cox (1975). Kalbfleisch and Prentice (1980) explain partial likelihood in detail. Partial likelihood only yields relative risks but can be very useful in gross analyses.

Suppose m items are under observation and n events have occurred up to time x. Let $\mathbb{F}(x_i)$ be a risk set of the events up to time x_i and let l be the number of events yet to occur. The partial likelihood of (2.8) is then given by

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{n}\frac{\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}_i)}{\sum\limits_{l\in\mathbb{F}(x_i)}\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}_l)} \tag{2.9}$$

In the case where relatively few ties, d_i, are present, the following relation holds:

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{n}\frac{\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}_i)}{\left[\sum\limits_{l\in\mathbb{F}(x_i)}\exp(\boldsymbol{\gamma}\cdot\boldsymbol{z}_l)\right]^{d_i}} \tag{2.10}$$

It is also possible to stratify the PHM into different strata, i.e.

$$h(x,\boldsymbol{z}) = h_{0_j}(x)\cdot\exp(\boldsymbol{\gamma}_j\cdot\boldsymbol{z}(x)) \tag{2.11}$$

with partial likelihood given by,

$$L(\boldsymbol{\gamma}) = \prod_{j=1}^{r}\prod_{i=1}^{k_j}\frac{\exp(\boldsymbol{\gamma}_j\cdot\boldsymbol{z}_{ij})}{\sum\limits_{l,j\in\mathbb{F}(x_{ij})}\exp(\boldsymbol{\gamma}_j\cdot\boldsymbol{z}_{lj})} \tag{2.12}$$

University of Pretoria etd – Vlok, P-J (2006)

where r denotes the number of strata and k_j is the number of events in the j^{th} stratum. Ascher, Kobbacy, and Percy (1997) applied the stratified PHM successfully.

If absolute risks are required, a fully parameterized PHM is required. A distribution often used to perform the parameterization is the Weibull distribution because of its flexibility. Substitution of the Weibull distribution in (2.8) yields

$$h(x, \boldsymbol{z}) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \cdot \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \tag{2.13}$$

where β and η are the shape and scale parameters of the Weibull distribution respectively.

The parameters PHM in (2.13) can be calculated by constructing the full likelihood as,

$$\begin{aligned} L(\beta, \eta, \boldsymbol{\gamma}, \boldsymbol{z}) &= \prod_{i=1}^{n} h(x_i, \boldsymbol{z}) \cdot \exp\left(-\int_0^{x_i} h(x, \boldsymbol{z}) \mathrm{d}x\right) \\ &= \prod_{i=1}^{n} \frac{\beta}{\eta} \cdot \left(\frac{x_i}{\eta}\right)^{\beta-1} \cdot e^{\boldsymbol{\gamma} \cdot \boldsymbol{z}_i} \cdot \exp\left(-\int_0^{x_i} \frac{\beta}{\eta} \cdot \left(\frac{x}{\eta}\right)^{\beta-1} \cdot e^{\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)} \, \mathrm{d}x\right) \end{aligned} \tag{2.14}$$

The solution of (2.14) is complex if $\boldsymbol{z}$ is dependent on time. Press et al. (1993) discuss some numerical techniques with which an economic solution can be obtained.

It is also possible to stratify the fully parametric PHM. Usually, either the baseline FOM or the regression coefficients are stratified, not both. This is done to limit the number of parameters in the model and to obtain synergy amongst different strata.

A useful extension of the PHM is Aalen's Regression Model as discussed in Aalen (1980) and Aalen (1989). This model can be used to test time dependence of covariates in the PHM and adds significant value to PHM analysis. In this model, the vector $\boldsymbol{h}(x; \boldsymbol{z})$ of the FOMs $h_j(x; \boldsymbol{z})$ for $j = 1, 2, .., n$, is given by:

$$\boldsymbol{h}(x; \boldsymbol{z}) = \boldsymbol{Y}(x) \cdot \boldsymbol{\alpha}(x) \tag{2.15}$$

Here $\boldsymbol{Y}(x)$ is an $n \times (q+1)$ matrix whose rows at time x_j consist of those vectors,

$$\boldsymbol{z}^j = \left[1, z_1^j(x), ..., z_q^j(x)\right] \tag{2.16}$$

where $z_i^j(x)$, $i = 1, 2, ..., q$ are covariate values, corresponding to those failure times that have not occurred up to time x_j. In the vector,

$$\boldsymbol{\alpha}(x) = [\alpha_0(x), \alpha_1(x), ..., \alpha_q(x)] \tag{2.17}$$

$\alpha_0(x)$ is the baseline parameter function, while $\alpha_i(x)$, $i = 1, 2, ..., q$ are called regression functions, defining the effects of covariates. The effect of a covariate is represented by the cumulative regression function, $\boldsymbol{A}(x)$, defined as:

$$\boldsymbol{A}_i(x) = \int_0^x \alpha_i(s) \mathrm{d}s \tag{2.18}$$

University of Pretoria etd – Vlok, P-J (2006)

for $i = 0, 1, ..., q$. To study the time-varying effect of the i^{th} covariate, an estimate of the i^{th} cumulative regression function should be plotted against the failure times. There are 4 possible outcomes:

(i) Straight line with an incline m. The effect is independent of time.
(ii) Constant line at value y. Indicates no effect at all.
(iii) Increasing at a decreasing rate. Indicates a decreasing effect over time.
(iv) Increasing at an increasing rate. Indicates an increasing effect over time.

Aalen's approach is particularly useful in analyzing condition monitoring data since condition monitoring data is almost always time-dependent.

2.3.1.2 Proportional Mean Intensity Models (PMIM)

Proportional Mean Intensity Models or Proportional ROCOF Models are constructed by the product of a baseline ROCOF multiplied with a functional term, dependent on time and covariates. PMIMs are very similar to PHMs as far as construction and estimation is concerned but they are based on fundamentally different representations of the intensity of failure processes. In the literature, the terminology for these concepts are often inconsistent, e.g. Kumar (1996) investigated the use of "Proportional Hazards Modeling" on repairable systems while he was actually using PMIMs.

Suppose the PMIM is constructed as the product of a baseline ROCOF, $\iota_{u0}(t)$, and a functional term $\lambda(t, \boldsymbol{z}(t))$, where $\boldsymbol{z}$ may or may not depend on time, i.e.

$$\iota_u(t, \boldsymbol{z}) = \iota_{u0}(t) \cdot \lambda(t, \boldsymbol{z}(t)) \tag{2.19}$$

As before, it is possible to estimate the semi-parametric model in (2.19) without making any assumptions about $\iota_{u0}(t)$ by using partial likelihood theory. Let m denote the number of items under observation and let n represent the total number of failures that have occurred. Let $\mathbb{F}(t_i)$ be the risk set of the failure events and let l represent the number of events yet to occur at time. The partial likelihood is then given by

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{n} \frac{\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i)}{\sum\limits_{l \in \mathbb{F}(t_i)} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_l)} \tag{2.20}$$

If the number of ties, d_i, in the data set is small, the following relation holds

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{n} \frac{\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i)}{\left[\sum\limits_{l \in \mathbb{F}(t_i)} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_l)\right]^{d_i}} \tag{2.21}$$

If the PMIM is stratified into r strata, i.e. $\iota_{u0j}(t, \boldsymbol{z}) = \iota_{u0j}(t)\exp(\boldsymbol{\gamma}_j \cdot \boldsymbol{z}(t))$, the partial likelihood becomes,

$$L(\boldsymbol{\gamma}) = \prod_{j=1}^{r}\prod_{i=1}^{k_j} \frac{\exp(\boldsymbol{\gamma}_j \cdot \boldsymbol{z}_{ij})}{\sum_{l,j \in F(t_{ij})} \exp(\boldsymbol{\gamma}_j \cdot \boldsymbol{z}_{lj})} \tag{2.22}$$

where r is the number of strata and k_j is the number of of events in the j^{th} stratum.

If an absolute mean intensity is required, the PMIM can be parameterized. The log-linear representation of a NHPP is often used to perform the parameterization, i.e. $\iota_{u_0}(t) = \exp(\alpha_0 + \alpha_1 \cdot t)$. The full likelihood becomes,

$$\begin{aligned} L(\alpha_0, \alpha_1, \boldsymbol{\gamma}, \boldsymbol{z}) &= \prod_{i=1}^{n} \iota_{u_0}(t) \cdot \exp\left(-\int_0^{Tn} \iota_u(t, \boldsymbol{z}(t))\mathrm{d}t\right) \\ &= \prod_{i=1}^{n} \left(e^{\alpha_0+\alpha_1 \cdot T_i} \cdot e^{\boldsymbol{\gamma}\cdot \boldsymbol{z}(T_i)}\right) \cdot \exp\left(-\int_0^{Tn} e^{\alpha_0+\alpha_1 \cdot t} \cdot e^{\boldsymbol{\gamma}\cdot\boldsymbol{z}(t)}\mathrm{d}t\right) \end{aligned} \tag{2.23}$$

As in the case of the parametric PHM, it is difficult to maximize (2.23) if the covariates are time-dependent.

If the fully parametric PMIM is stratified, usually, either the baseline ROCOF or the regression coefficients are stratified, not both. This is done to limit the number of parameters in the model and to obtain synergy amongst different strata.

2.3.1.3 Proportional Odds Model (POM)

The proportional odds model originated from epidemiological studies and was introduced by Bennet (1983) for use in biomedicine. This model is structurally similar to the PHM, but not a direct extension. It models the odds of an event occurring and unlike the PHM, the effect of covariates in the POM model diminishes as time approaches infinity. This diminishing property of the covariates means that the model is suitable for situations where an item adjusts to factors imposed on it or the factors only operate in early stages.

For this model the *odds* of a failure occurring is defined in terms of the survivor function as,

$$\frac{F_X(x)}{R_X(x)} = \frac{1 - R_X(x)}{R_X(x)} \tag{2.24}$$

This definition of odds is used to introduce the POM:

$$\frac{1 - R(x, \boldsymbol{z})}{R(x, \boldsymbol{z})} = \psi \cdot \frac{1 - R_X(x)}{R_X(x)} \tag{2.25}$$

University of Pretoria etd – Vlok, P-J (2006)

Equation (2.25) states that the odds for a failure to occur under the influence of covariates are ψ times higher than the odds of a failure without the effects of covariates. If ψ increases, so does the probability of a shorter life time. Differentiation of (2.25) with respect to time leads to,

$$\frac{h(x,\boldsymbol{z})}{R(x,\boldsymbol{z})} = \psi \cdot \frac{h_X(x)}{R_X(x)} \tag{2.26}$$

after using the coefficient rule. By rearranging the terms in (2.26) and re-using (2.25), a FOM ratio can be obtained:

$$\frac{h(x,\boldsymbol{z})}{h_X(x)} = \psi \cdot \frac{R(x,\boldsymbol{z})}{R_X(x)} = \frac{1-R(x,\boldsymbol{z})}{1-R_X(x)} \tag{2.27}$$

Inspection shows that $\psi|_{x=0} = \psi$ and $\psi|_{x=\infty} = 1$, from there the diminishing effect of the covariates.

Bennet (1983) derives the full likelihood for the model in his original paper to estimate the model parameters. Research done by Shen (1998) provides more efficient estimation methods and methods to enable the model to handle suspended observations.

A special case of the POM arise when it assumed that event times are distributed according to a log-logistic distribution. Kalbfleisch and Prentice (1980) describe this special case in detail. The FOM of an item with event times following a log-logistic distribution is given by:

$$h(x;\boldsymbol{z}) = \frac{\delta}{x\cdot(1+x^{-\delta}\cdot\exp(-\boldsymbol{\gamma}\cdot\boldsymbol{z}(x)))} \tag{2.28}$$

where δ is a measure of precision. The FOM is assumed to be increasing first and then decreasing with a change at time

$$x = \{(1-\delta)\exp(-\boldsymbol{\gamma}\cdot\boldsymbol{z}(x))\}^{1/\delta} \tag{2.29}$$

If $x\to\infty$, $x^{-\delta}\cdot\exp(-\boldsymbol{\gamma}\cdot\boldsymbol{z}(x))\to 0$ (see (B.4)) and subsequently covariates will influence the FOM less and less as the item ages.

2.3.2 Additive Intensity Models (AIMs)

Additive Intensity Models represent the intensity of a failure process as the sum of a baseline intensity and a functional term containing covariates. Pijnenburg (1991) deals with AIMs in completely general terms. Newby (1993) compare this type of model, for the case where the FOM is used as intensity, to various other regression models. Authors often refer to AIMs incorrectly as Additive *Hazard* Models (AHM) in reliability modeling literature. This section describes AIMs in general terms.

University of Pretoria etd – Vlok, P-J (2006)

Suppose two items are in series, S_1 and S_2. Suppose S_1 represents a repairable system and S_2 is an item representing the influence of covariates. Let T_i be be the time at which the i^{th} system failure occurs and X_i the system's i^{th} interarrival time, i.e. $X_i = T_i - T_{i-1}$. The system is supposed to have a survival time $X_{1,1}$ and FOM $h_1(\cdot)$ and the item representing the covariates has a survival time of $X_{2,1}$ with a FOM, $\lambda(t_0, \boldsymbol{z}_0)$. For the moment $\lambda(t_0, \boldsymbol{z}_0)$ is defined as constant in-between interarrival times, i.e. constant covariates, but variable over successive lifetimes, i.e. dependent on time.

After the first failure at system level at time, T_1, i.e. $T_1 = X_1 = \min(X_{1,1}, X_{2,1})$ both components are replaced, such that:

(i) S_1 is renewed by an identical component, also called S_1, with lifetime $X_{1,2}$ and FOM $h_1(\cdot)$.
(ii) S_2 is replaced by a component with lifetime $X_{2,2}$ and FOM $\lambda(t_1, \boldsymbol{z}_1)$.

In general terms it means that after the i^{th} failure on system level at time T_i, i.e. $T_i = T_{i-1} + \min(X_{1,i}, X_{2,i})$:

(i) S_1 is replaced by an identical new component with lifetime $X_{1,i+1}$ and FOM $h_1(\cdot)$. The lifetimes $X_{1,k}$ are assumed to be IID.
(ii) S_2 is replaced by a component with a lifetime $X_{2,i+1}$ and FOM $\lambda(t_i, \boldsymbol{z}_i)$. The lifetimes $X_{2,k}$ are assumed to be statistically independent.

It is also assumed that the survival times $X_{1,k}$ and $X_{2,k}$ are mutually independent.

For the various FOMs, $\lambda(t_i, \boldsymbol{z}_i)$, it is assumed that the covariates, $\boldsymbol{z}_i$, are constant in $[T_i, T_{i+1})$ but may change for different lifetimes. Higher covariate values, generally represent more severe environmental stresses and $x_{2,i} = \min(x_{1,i}, x_{2,i})$ should be interpreted as a system failure due to these higher environmental stresses. Pijnenburg (1991) suggests a few forms for $\lambda(t_i, \boldsymbol{z}_i)$. The simplest form for $\lambda(t_i, \boldsymbol{z}_i)$ is a linear function,

$$\lambda(t, \boldsymbol{z}) = \boldsymbol{\gamma} \cdot \boldsymbol{z} = \sum_{i=1}^{p} \gamma_i \cdot z_i \tag{2.30}$$

for p covariates. A linear term can be included in (2.30) by simply specifying $z_1 = 1$. In the case where higher order terms are present in the polynomial, $\lambda(t_i, \boldsymbol{z}_i)$ can be specified as,

$$\lambda(t, \boldsymbol{z}) = \sum_{i=1}^{p} \sum_{j=0}^{m} \gamma_{ij} \cdot z_i^j \tag{2.31}$$

If covariates appear to interact, $\lambda(t_i, \boldsymbol{z}_i)$ can be chosen as,

$$\lambda(t, \boldsymbol{z}) = \boldsymbol{\gamma}_0 + \sum_{i=1}^{p} \sum_{j=1}^{i-1} \gamma_{ij} \cdot z_i \cdot z_j \tag{2.32}$$

University of Pretoria etd – Vlok, P-J (2006)

A suitable form to handle both higher order terms and interaction can be,

$$\lambda(t, \boldsymbol{z}) = \sum_{i=1}^{p} \sum_{j=1}^{i} \sum_{k=0}^{r} \sum_{l=0}^{r} \gamma_{ijkl} \cdot z_i^k \cdot z_j^l \tag{2.33}$$

Data limitations often cause that only (2.30) is practical.

Following the argument above, the AIM is completely generalized by allowing covariates to dependent on time, i.e.

$$\iota(t, \boldsymbol{z}) = \iota_0(t) + \lambda(t, \boldsymbol{z}(t)) \tag{2.34}$$

The AIM can be stratified and parameterized in the same manner as the PHM and PMIM. Pijnenburg (1991) also allows the AIM to have modified time scales. These extensions are discussed in sections to follow. Crowder, Kimber, Smith, and Sweeting (1991) and Newby (1993) derive the full likelihood to fit AIMs‡.

2.3.3 Models with mixed or modified time scales

Modified or mixed time scales in intensity models can be interpreted as an additional covariate in data sets to provide more flexibility. Modified time scales increase or decrease the modeled intensity of a failure process by either accelerating or decelerating the actual age of an item. Mixed time scale models incorporate local and global time in the same model to utilize the advantages of both long and short term history. Newby (1993) refers to the result of these concepts as the *virtual age* of an item since the actual survival time differs from the survival time used in models.

2.3.3.1 The Prentice Williams Peterson (PWP) model

Prentice, Williams, and Peterson (1981) published the so-called PWP model after research done by Williams (1981). This model is generally considered as the most significant extension of the PHM by Cox (1972) according to Ascher and Feingold (1984). Two versions of the PWP model were proposed, both of the stratified Proportional Hazards type, which means this discussion would also be applicable in Section 2.3.1 where multiplicative models were considered but it is believed that this model made a more significant contribution to models with modified or mixed time scales.

The PWP model is specifically directed towards the analysis of situations where only a small number of observations is available on an item but where a large number of items is studied.

‡Partial likelihood can not be used because of the summation of terms in the model and relative risks are thus not possible

University of Pretoria etd – Vlok, P-J (2006)

Specific items are also allowed to experience multiple failures. This makes the PWP very attractive in reliability modeling where data sets are often limited in size.

The model is constructed as follows. Let $\boldsymbol{z} = [z_1(t), ..., z_p(t)]$ denote a vector of covariates of a specific item, part of the covariate process, $Z(t)$. Also, let $N(t)$, denote the counting process of the the number of failures, $n(t)$, on an item up to time t. The counting process, $N(t)$, is equivalent to the random failure times $T_1 < .... < T_{n(t)}$ in $[0, t)$. Prentice, Williams and Peterson then define the intensity of a failure process as,

$$\iota(t|N(s), Z(s), s \leq t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T_{n(t)+1} \leq t + \Delta t | N(s), Z(s), s \leq t]}{\Delta t} \tag{2.35}$$

Some special cases of (2.35), in the absence of covariates, are:

(i) $\iota(t|N(t)\} = \iota_u(t)$ for some $\iota_u(\cdot) \geq 0$ is the unconditional intensity function of a NHPP.

(ii) $\iota(t|N(t)) = n(t)\iota_u(t)$ specifies a nonhomogeneous pure birth process.

(iii) $\iota(t|N(t)) = \iota_k(t - t_k)$ for an arbitrary $\iota_k(\cdot) \geq 0$ $(k = n(t) + 1 = 1, 2, ...)$ gives a semi-Markov process.

(iv) A further restriction on the semi-Markov process that $\iota_k(\cdot) = \iota_0(\cdot)$ for all k gives an ordinary renewal process.

Prentice, Williams and Peterson suggest two models based on (2.35), both of a stratified Proportional Hazards type,

$$\text{PWP Model 1}: \ \iota(t|N(t), Z(t)) = \iota_{0_s}(t) \cdot \exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}(t)) \tag{2.36}$$

$$\text{PWP Model 2}: \ \iota(t|N(t), Z(t)) = \iota_{0_s}(t - t_{n(t)}) \cdot \exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}(t)) \tag{2.37}$$

The stratification variable $s = s\{N(t), Z(t), t\}$ may change as a function of time for a given item, e.g. $s = n(t) + 1$ and the subject moves to stratum k immediately following its $(k-1)^{\text{th}}$ failure and remains there until the k^{th} failure. More refined stratum conditions can easily be constructed. For Model 1, it is possible to define $s = 2 \cdot n(t) + \Lambda\{N(t)\}$, where $\Lambda\{N(t)\} = 1$ if the time since the last failure, $t - t_{n(t)}$, is less than some specified value and $\Lambda\{N(t)\} = 2$, otherwise. In the case of Model 2, it is possible to define $s = 2 \cdot n(t) + \Delta(t)$, where $\Delta(t) = 1$ if t is less than some value and $\Delta(t) = 2$, otherwise.

The PWP formulations differ from Andersen (1985) in two aspects: (a) the risk sets of the $(k+1)^{\text{th}}$ recurrences are restricted to the individuals who have experienced the first k recurrences; and (b) the underlying intensity functions and regression parameters are allowed to vary amongst distinct recurrences. Gail, Santner, and Brown (1980) published a two-sample special case of Model 2 with strata defined at least as finely as $s = n(t) + 1$. Clifton and Crowley (1978) considered a special case of Model 1 without covariates and with $s = 1$ if $n(t) = 0$ and $s = 2$ if $n(t) \geq 1$.

University of Pretoria etd – Vlok, P-J (2006)

Partial likelihood can be used to estimate relative risks with both Models 1 and 2. For Model 1 the partial likelihood is,

$$L(\gamma) = \prod_{s\geq 1}\prod_{i=1}^{d_s} \frac{\exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}_{si}(t_{si}))}{\sum\limits_{l\in\mathbb{R}(t_{si},s)} \exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}_l(t_{si}))} \tag{2.38}$$

where t_{si} denotes the failure time of item i in stratum s, $\boldsymbol{z}_{si}(t_{si})$ refers to the covariate vector of item i at time t_{si} and d_s denotes the total number of events in stratum s.

The partial likelihood for Model 2 is,

$$L(\gamma) = \prod_{s\geq 1}\prod_{i=1}^{d_s} \frac{\exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}_{si}(t_{si}))}{\sum\limits_{l\in R(u_{si},s)} \exp(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}_l(\zeta_l + u_{si}))} \tag{2.39}$$

where u_{si} are the interarrival times of the different items in various stratums and ζ_l is the last failure time on item l prior to entry into stratum s ($\zeta_l = 0$ if no prior failure on the item).

Prentice et al. also extend the PWP model to multivariate failure time applications where there will be more than one type of failure. Let $J \in \{1, 2, ..., m\}$ denote m mutually exclusive failure type classes. Analogous to (2.35) it is possible to define type-specific intensity functions at time t by,

$$\iota_j(t|N(t), Z(t)) = \lim_{\Delta t\to 0} \frac{\Pr[t \leq T_{n(t)+1} \leq t + \Delta t, J = j|N(t), Z(t)]}{\Delta t} \tag{2.40}$$

where, in this case, $N(t) = \{N_1(t), ..., N_m(t)\}$ is the counting process for each of the m types of failure and $n(t) = n_1(t) + ... + n_m(t)$. This leads to the following extensions of (2.36) and (2.37):

$$\iota_j(t|N(t), Z(t)) = \iota_{0_{sj}}(t) \cdot \exp(\boldsymbol{\gamma}_{sj} \cdot \boldsymbol{z}(t)) \tag{2.41}$$

$$\iota_j(t|N(t), Z(t)\} = \iota_{0_{sj}}(t - t_{n(t)}) \cdot \exp(\boldsymbol{\gamma}_{sj} \cdot \boldsymbol{z}(t)) \tag{2.42}$$

Prentice et al. applied their models on a data set from Atkinson et al. (1979) and generally achieved better results than with an ordinary PHM. In another example, Ascher (1983) used PWP Model 2 on marine gas turbine failure data by using indicative covariates, i.e. 0's and 1's, with good results. Ascher believes the "custom tailoring" allowed by the PWP models, is essential in failure data analysis.

2.3.3.2 Accelerated Failure Time Models (AFTM)

Pike (1966) introduced the AFTM and it is often a useful alternative to the PHM in many reliability modeling situations, according to Newby (1988). This model incorporates the effect of covariates by allowing for changes in the time scale of, for example, the reliability function.

University of Pretoria etd – Vlok, P-J (2006)

Let the probabilistic reliability function be given by $R_X(x)$ and the accelerated reliability function be denoted by $\widehat{R_X(x)}$, due to environmental stresses, i.e.

$$\widehat{R_X(x)} = R[(x-c)/b] \tag{2.43}$$

where c is a location parameter and b is a scale parameter. The model is similar to regression models which assume that $(x-c)/b$ is distributed according to a known parametric form. Some density functions often used, are,

Weibull:

$$\widehat{f_X(x)} = \frac{k}{b}u^{k-1}\exp(-u^k), \quad u = (x-c)/b \tag{2.44}$$

Gamma:

$$\widehat{f_X(x)} = \frac{1}{b\Gamma(k)}u^{k-1}\exp(-u), \quad u = x/b \tag{2.45}$$

Log-normal:

$$\widehat{f_X(x)} = \frac{1}{buk\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\frac{\ln(u)}{k}\right]^2\right\}, \quad u = (x-c)/b \tag{2.46}$$

Inverse Gaussian:

$$\widehat{f_X(x)} = \frac{1}{b}\sqrt{\frac{k}{2\pi}}u^{-3/2}\exp\left\{-\frac{-k(u-1)^2}{2u}\right\}, \quad u = x/b \tag{2.47}$$

where k is a shape parameter.

A logical variation on the AFTM is where the underlying distribution is only a function of x and k, $f_X(x,k)$, but the accelerated distribution is $f_X\widehat{(x/b_i,k)}$, with $b_i = b(\boldsymbol{z}_i;q)$ where $\boldsymbol{z}_i$ is a vector of covariates and q is a flexibility parameter. Commonly used models for b are:

(i) Constant

(ii) Linear as a function of stress, e.g. $b = \boldsymbol{z}_i \cdot a$

(iii) Exponential, e.g. $b_i = \exp(\boldsymbol{z}_i \cdot a)$

(iv) Inversely exponential, e.g. $b_i = \exp(-1/(\boldsymbol{z}_i \cdot a))$ (Arrhenius model)

Constructing the likelihood for the mentioned models is similar to simple two-parameter likelihood construction. Many authors have discussed the likelihood construction, including Smith and Naylor (1987), Cheng and Amin (1983) and Cheng and Isles (1987).

Solomon (1984) has shown that, in the absence of censoring, the relative effect of covariates are identical in the AFTM and the PHM. Great care should thus be taken in such cases that either one of the two models is not misspecified.

University of Pretoria etd – Vlok, P-J (2006)

A very popular application of the AFTM is fatigue crack growth in the field of structural mechanics. The acceleration-property of the model is used to estimate fatigue crack growth rates. Many examples of this kind can be found in the literature, including Crowder, Kimber, Smith, and Sweeting (1991) and Newby (1988).

Ciampi and Etezadi-Amoli (1985) and Etezadi-Amoli and Ciampi (1987) combined the PHM and the AFTM in the so-called Extended Hazard Regression Model (EHRM), i.e.

$$h(x; \boldsymbol{z}) = h_0\left(x \cdot \psi_1(\boldsymbol{z}(x) \cdot \alpha)\right) \psi(\boldsymbol{z}(x) \cdot \beta) \tag{2.48}$$

where $\psi_1(\boldsymbol{z}(x) \cdot \alpha)$ and $\psi(\boldsymbol{z}(x) \cdot \beta)$ are positive functions equal to 1 when all covariate values are equal to 0. When $\alpha = 0$, the model in (2.48) becomes an ordinary PHM and when $\alpha = \beta$, the corresponding model is an AFTM.

2.3.3.3 Proportional Age Setback (PAS)

In this approach, introduced by Martorell, Munoz, and Serradell (1996), each maintenance action is assumed to shift the origin of time from where the age of the component is evaluated. Let every maintenance action reduce the age of a component, just before maintenance, proportionally by a factor ε, where ε lies in $[0, 1]$. If $\varepsilon = 0$, the PAS produces the BAO situation and if $\varepsilon = 1$, the GAN situation results. Thus, the virtual age, τ, of an item after it has undergone its first maintenance action[§] is given by:

$$\tau_1^+ = (1 - \varepsilon_1) \cdot \lambda(\boldsymbol{z}_1) \cdot T_1 \tag{2.49}$$

In (2.49), $\lambda(\cdot)$ is a functional term containing covariates in the vector $\boldsymbol{z}_1$. Covariates could be time-dependent or time-independent. The superscript "+" indicates that the virtual age is applicable shortly after the event at T_1. After the second maintenance action the virtual age is,

$$\tau_2^+ = (1 - \varepsilon_2) \cdot [\tau_1 + \lambda(\boldsymbol{z}_2) \cdot (T_2 - T_1)] \tag{2.50}$$

Substitution of the above yields:

$$\tau_2^+ = (1 - \varepsilon_2) \cdot [(1 - \varepsilon_1) \cdot \lambda(\boldsymbol{z}_1) \cdot T_1 + \lambda(\boldsymbol{z}_2) \cdot (T_2 - T_1)] \tag{2.51}$$

If m denotes the maintenance number, the virtual age of a component is generally given by,

$$\tau_m^+ = \sum_{k=0}^{m-1} \lambda(\boldsymbol{z}_{m-k}) \cdot \prod_{r=0}^{k} (1 - \varepsilon_{m-r}) \cdot (T_{m-k} - T_{m-k-1}) \tag{2.52}$$

Martorell, Sanchez, and Serradell (1999) simplify the virtual age model in (2.52) by assuming that,

[§] *Maintenance action* could be interpreted here as renewal, minimal repair or imperfect repair.

University of Pretoria etd – Vlok, P-J (2006)

(i) the effectiveness of each maintenance action is equal to some constant value ε, i.e. $\varepsilon_k = \varepsilon$.

(ii) constant operating conditions apply, i.e. $\boldsymbol{z}_k = \boldsymbol{z}$.

This leads to a simplification of (2.52), i.e.

$$\begin{aligned} \tau_m^+ &= \lambda(\boldsymbol{z}) \cdot \left[\sum_{k=0}^{m-1} (1-\varepsilon)^{k+1} \cdot (T_{m-k} - T_{m-k-1}) \right] \\ &= \lambda(\boldsymbol{z}) \cdot (t_m - \Delta\tau_m) \end{aligned} \tag{2.53}$$

where

$$\Delta\tau_m = \sum_{k=0}^{m-1} (1-\varepsilon)^k \cdot k \cdot T_{m-k} \tag{2.54}$$

This simplification is useful in cases where information in data sets are limited or where a first approximation suffices.

Atwood (1992) used an approach similar to the PAS, specifically on the ROCOF of items, i.e. $\rho(t) = \rho_0 \cdot g(t;\beta)$. Here, ρ_0 is a constant multiplier and $g(t;\beta)$ is the portion of the expression that determines the shape of $\rho(t)$. Three models for the ROCOF are proposed (exponential, linear and power law):

$$\rho(t) = \begin{cases} \rho_0 \exp[\beta(t-t_0)] \\ \rho_0[1+\beta(t-t_0)] \\ \rho_0(t/t_0)^{\beta}] \end{cases} \tag{2.55}$$

The value of t_0 can be selected for convenience. In the first two cases, if t_0 is set to zero, $t - t_0$ is the time measured from the system's installation. In the third case, t_0 normalizes the scale in which time is measured. In all three models, ρ_0 has units of 1/time. If $\beta > 0$, $\rho(t)$ is increasing, $\beta = 0$, $\rho(t)$ is constant or $\beta < 0$, $\rho(t)$ is decreasing. The value of ρ_0 is the value of $\rho(t)$ at time $t = t_0$. Atwood uses a Bayesian approach to fit the models, i.e. a preliminary analysis is done first based on the conditional likelihood of the models given in (2.55) whereafter the full likelihood is constructed and the values of parameters are estimated.

2.3.3.4 Proportional Age Reduction (PAR)

Malik (1979) introduced the PAR model, where the virtual age is based on the survival time of the most recent lifetime. This differs from the PAS approach where the virtual age is based on the entire history.

Let ε be the efficiency factor, as before, that lies within $[0,1]$. The virtual age of an item, after it has undergone its first maintenance action, in the PAR model is given by:

$$\tau_1^+ = (1-\varepsilon_1) \cdot \lambda(\boldsymbol{z}_1) \cdot T_1 \tag{2.56}$$

University of Pretoria etd – Vlok, P-J (2006)

The functional term, $\lambda(\cdot)$, incorporates covariates and the superscript "+" denotes applicability of the τ_1 shortly after event T_1 occurred. After the second maintenance action the virtual age is,

$$\tau_2^+ = \tau_1 + (1-\varepsilon_2)\cdot\lambda(\boldsymbol{z}_2)\cdot(T_2 - T_1) \tag{2.57}$$

Immediately after maintenance action m, the virtual age is given by:

$$w_m^+ = \sum_{k=m}^{m}(1-\varepsilon_k)\cdot\lambda(\boldsymbol{z}_k)\cdot(T_m - T_{m-1}) \tag{2.58}$$

If ε and $\boldsymbol{z}$ are fixed, (2.58) simplifies to,

$$\tau_m^+ = (1-\varepsilon)\cdot\lambda(\boldsymbol{z})\cdot T_m \tag{2.59}$$

This simplified estimation of the PAR was applied by, amongst others, Malik (1979) and Shin, Lim, and Lie (1996). Shin, Lim, and Lie, for example, implemented the PAR concept on two models namely the power-law intensity function (Weibull) and log-linear intensity function. From here the PAR model is defined as

$$\iota_{k+1}(t) = \iota_k(t - \zeta\cdot\tau_k),\ t > \tau_k \tag{2.60}$$

where ζ is an improvement factor or factor of rejuvenation and $0 \leq \zeta \leq 1$. This particular model was only used on a single item under observation but it can be extended to handle counts of multiple system copies.

2.3.4 Marginal regression analysis

Marginal regression analysis has been used with success in the field of biomedicine to represent multiple-event time data. See for example Pepe and Cai (1993) and Wei, Lin, and Weissfeld (1989). The approach of marginal regression analysis is similar to the stratified PHM approach with vaguely defined strata. This approach has the attractive attribute that no explicit model needs to be formulated for the probabilistic association between failures of the same individual. Wei et al. also allow for k different failure types, censoring and missing observations, which could be very useful in reliability modeling.

Let X_{ki} be the failure time of the i^{th} subject ($i = 1,...n$) that experiences the k^{th} type of failure ($k = 1,...,K$). In some instances a bivariate vector, $\widetilde{\boldsymbol{X}}_{ki}$, is observed consisting of $[X_{ki}, \Delta_{ki}]$, where $X_{ki} = \min(\widetilde{\boldsymbol{X}}_{ki}, C_{ki})$ and C_{ki} is the censoring time. Let $\Delta_{ki} = 1$ if $X_{ki} = \widetilde{\boldsymbol{X}}_{ki}$ and $\Delta_{ki} = 0$ otherwise. If $\widetilde{\boldsymbol{X}}_{ki}$ is missing, $C_{ki} = 0$, which implies that $X_{kk} = 0$ and $\Delta_{ki} = 0$, since $X_{ki} = \widetilde{\boldsymbol{X}}_{ki}$ is positive. Let $\boldsymbol{z}_{ki} = [z_{1ki}, ..., z_{1ki}]$ be a vector of p covariates for the i^{th} subject with respect to the k^{th} type of failure. Conditional on $\boldsymbol{z}_{ki}$, the failure vector $\widetilde{\boldsymbol{X}}_i = [\widetilde{X}_{1i}, ..., \widetilde{X}_{Ki}]$ and the censoring vector $\boldsymbol{C}_1 = [C_{1i}, ..., C_{Ki}]$ ($i = 1, ..., n$) are assumed to

University of Pretoria etd – Vlok, P-J (2006)

be independent. For the k^{th} type of failure of the i^{th} subject, the FOM $h_{ki}(x)$ is assumed to take the form,

$$h_{ki}(x) = h_{k0}(x) \cdot \exp[\boldsymbol{\gamma}_k \cdot \boldsymbol{z}_{ki}(x)] \tag{2.61}$$

where $h_{k0}(x)$ is an unspecified baseline FOM and $\boldsymbol{\gamma}_k$ is a vector of failure-specific regression coefficients. If $\mathbb{F}_k(x) = \{l : X_{kl} \geq x\}$ is defined as the set of subjects at risk just prior to time x with the respect to the k^{th} type of failure, the k^{th} failure-specific partial likelihood is,

$$L_k(\boldsymbol{\gamma}) = \prod_{i=1}^{n} \left[\frac{\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_{ki}(X_{ki}))}{\sum\limits_{l \in \mathbb{F}_k(X_{ki})} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_{kl}(X_{ki}))} \right]^{\Delta_{ki}} \tag{2.62}$$

Pepe and Cai (1993) considered a simplification of the approach above by defining a FOM $h^F(x)$ for individuals at risk at time x but not previously infected and a FOM $h^R(x)$ for individuals at risk and previously infected. It is possible to decompose $h^R(x)$ into further components, i.e. $\{h^{2/1}, h^{3/2}, ...\}$, where $h^{k/(k-1)}(x)$ is the FOM of individuals with the k^{th} infection amongst those who have already experienced $k-1$ infections. Ascher, Kobbacy, and Percy (1997) proposed a similar approach by specifying a different FOM for items following either corrective maintenance or preventive maintenance. Every model has its own baseline and regression coefficients.

The approach of marginal regression analysis as outlined above requires large data sets - something that is not common in reliability. The failure type specific regression coefficients is an important attribute however, since machines rarely fail repeatedly because of the same type of failure.

2.3.5 Competing risks

Crowder (1991) believes the principle of competing risks is best explained by an example from the field of biomedicine. Suppose the time to recurrence of a specific type of cancer in a group of patients is modeled. Patients not only run the risk of the recurrence of cancer but also, for example, of dying before recurrence or developing a different disease before recurrence. This problem is defined as competing risks in data and is common in reliability problems.

Competing risk models have two interpretations: (1) it describes the lifetime of a system subject to several potential causes of failure; and (2) it describes the lifetime of a system consisting of a series of components which fails as soon as one of the components fail. The occurrences of potential failures can be regarded as a vector of random variables $\boldsymbol{X} = [X_1, ..., X_n]$, so that the actual stopping time is at the smallest element of $\boldsymbol{X}$, say X_i. If the random variables in

University of Pretoria etd – Vlok, P-J (2006)

$\boldsymbol{X}$ are independent, the system reliability is given by,

$$R_{sys}(x) = \prod_{i=1}^{n} R_i(x) \tag{2.63}$$

with FOM

$$h_{sys}(x) = \sum_{i=1}^{n} h_i(x) \tag{2.64}$$

Competing risks situations arise naturally in reliability problems, particularly where series systems are considered. Equations (2.63) and (2.64) are, for example, directly applicable in the "weakest-link" argument of Blanchard and Fabrycky (1990). Lewis (1987) considered an approach similar to that of competing risks, called the "β-factor" method. This method analyzes a system as a series of (i)subsystems of independent components; and (ii) common-cause components. Crowder (1991) derives the likelihood for competing risks models in general terms.

2.3.6 Frailty or Mixture Models

The concept of frailty or mixture is used in two ways in reliability models: (1) as a way of introducing an idea of heterogeneity into the construction of a model; and (2) as an object of interest in itself. Mixture models are more applicable in reliability than frailty models.

Frailty in this context is an unobservable random effect shared by subjects in a group. It is defined by Vaupel, Manton, and Stallard (1979) rather like the PHM, but differs in that the relative risk factor is a random variable in this case. The frailty, ξ, is defined in terms of the FOMs of individuals in a population, i.e.

$$h(x|\xi) = \xi \cdot h_0(x) \tag{2.65}$$

is the FOM of an individual with frailty ξ and baseline FOM, h_0. If the frailty at time x has a density $f_x(\cdot)$, the average FOM at time x is,

$$\begin{aligned} \overline{h}(x) &= \int_0^\infty h(x|\xi) \cdot f_x(\xi) \mathrm{d}\xi \\ &= h_0(x) \int_0^\infty \xi \cdot f_x(\xi) \mathrm{d}\xi \\ &= \overline{\xi} \cdot h_0(x) \end{aligned} \tag{2.66}$$

If the frailty decreases with time, so will $\overline{\xi}$ (since the weakest die young if no fatal external influences are present). This leads to a situation where the average FOM is declining more rapidly than the FOM for individuals.

University of Pretoria etd – Vlok, P-J (2006)

Many authors have also used frailties in regression models. See for example Klein and Moeschberger (1990). The inclusion of frailties overcome the limiting assumption of most regression models that survival times of distinct subjects are independent of each other. This assumption is for instance not valid for a study on litter mates that share the same genetic makeup or married couples that share the same, unmeasured environment.

Klein and Moeschberger (1990) present two types of frailty models (both based on Cox's PHM). For the first it is assumed that the FOM of the j^{th} subject in the i^{th} group, given the frailty, to be

$$h_{ij}(x) = h_0(x) \cdot \exp(\sigma w_i + \boldsymbol{\gamma} \cdot \boldsymbol{z}_{ij}) \tag{2.67}$$

where $w_1, ..., w_G$ are frailties. It is assumed that the w's are an independent sample from some distribution with mean 0 and variance 1. The second model is given by

$$h_{ij}(x) = h_0(x) \cdot u_i \cdot \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_{ij}) \tag{2.68}$$

where the u_i's are an independent and identically distributed sample from a distribution with mean 1 and some unknown variance. Common models proposed in the literature for the random effect are the one-parameter gamma distribution, the inverse Gaussian distribution and the log normal distribution.

Lawless (1987) introduced frailties in a Poisson process model by including a variable α_i which accounts for unobservable random effects for each subject, i.e. $\rho(t, \boldsymbol{z}) = \alpha_i \rho_0(t) \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i)$. The α_i's are independent and identically distributed random variables, independent of the z_i's with some distribution $G(\alpha)$. The likelihood for subject i's event history over $(0, T_i]$ is

$$L_i(\boldsymbol{\theta}) = \int_0^{\infty} \prod_{j=1}^{n_i} \alpha_i \rho(t_{ij}) \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i) \exp\left\{-\int_0^{T_i} \alpha_i \rho(t) \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i) \mathrm{d}t\right\} \mathrm{d}G(\alpha_i) \tag{2.69}$$

Mixture models arise naturally in reliability according to, amongst others, Lancaster (1990) and Littlewood and Verrall (1973). These models are expressed as a conditional FOM, $h(x|\boldsymbol{z}, \sigma)$, where σ is a random variable with density ω. The conditional density and survivor functions for x are,

$$\begin{aligned} f(x|\boldsymbol{z}) &= \int f(x|\boldsymbol{z}, \sigma)\omega(\sigma)\mathrm{d}\sigma \\ &= \int h(x|\boldsymbol{z}, \sigma) \exp[-H(x|\boldsymbol{z}, \sigma)]\mathrm{d}\sigma \end{aligned} \tag{2.70}$$

and

$$\begin{aligned} R(x|\boldsymbol{z}) &= \int R(x|\boldsymbol{z}, \sigma)\omega(\sigma)\mathrm{d}\sigma \\ &= \int \exp[-H(x|\boldsymbol{z}, \sigma)]\mathrm{d}\sigma \end{aligned} \tag{2.71}$$

University of Pretoria etd – Vlok, P-J (2006)

Note that,

$$h(x|\boldsymbol{z}) = \frac{f(x|\boldsymbol{z})}{R(x|\boldsymbol{z})} \neq \int h(x|\boldsymbol{z},\sigma)w(\sigma)\mathrm{d}\sigma = \overline{h}(x|\boldsymbol{z}) \tag{2.72}$$

since the FOM defined in terms of frailty is not the FOM of the unconditional distribution.

Mixture models are also often interpreted as Bayesian models with prior $\omega(\sigma)$ for a parameter σ. See for example Lancaster (1990) and Ridder (1990).

2.3.7 Noteworthy extensions of intensity concepts

On a few occasions authors published extensions of the failure intensity concepts described above that are beneficial to this study. These extensions are mostly integrations of different approaches to suit particular applications.

2.3.7.1 A point-process model incorporating renewals and time trends, with application to repairable systems

Lawless and Thiagarajah (1996) presented a family of models that incorporates both Poisson¶ and renewal behavior although multiple system copies are not considered. The authors studied models of the form,

$$\iota(t, \boldsymbol{z}) = e^{\boldsymbol{\gamma}\cdot\boldsymbol{z}(t)} \tag{2.73}$$

where $\boldsymbol{z}(t) = [z_1(t), ..., z_p(t)]$ and $\boldsymbol{\gamma} = [\gamma_1, ..., \gamma_p]$. This model is a special case of that considered by Berman and Turner (1992). Two important Poisson processes can be modeled by (2.73) by specifying the covariates intelligently:

(i) $\rho_1(t) = \exp(\alpha + \beta t)$ by letting $\boldsymbol{z}(t) = [1, t]$ and $\boldsymbol{\gamma} = [\alpha, \beta]$

(ii) $\rho_2(t) = \alpha t^{\beta}$ by letting $\boldsymbol{z}(t) = [1, \log t]$ and $\boldsymbol{\gamma} = [\log \alpha, \beta]$

Renewal processes are obtained by taking $\boldsymbol{z}(t)$ as a function of the backward recurrent time, $B(t)$, as defined in Section 1.2.1. For example, $\boldsymbol{z}(t) = [1, \log B(t)]$ and $\boldsymbol{\gamma} = [\log \alpha, \beta]$ produce a renewal process with a Weibull distribution and FOM $h_X(x) = \alpha x^{\beta}$. Models with $\boldsymbol{z}(t) = [1, g_1(t), g_2(B(t))]$, where g_1 and g_2 are specified functions, incorporate both renewal and time trend behavior.

¶See Section A.3.3 for details on Poisson processes.

University of Pretoria etd – Vlok, P-J (2006)

Anderson, Borgan, Gill, and Keiding (1993) and Berman and Turner (1992) have shown that the maximum likelihood of (2.73) is given by,

$$L(\gamma) = \prod_{i=1}^{n} \rho(t_i) \cdot \exp\left\{-\int_0^T \rho(t)\mathrm{d}t\right\} \tag{2.74}$$

for Poisson processes and for renewal processes by,

$$L(\gamma) = \prod_{i=1}^{n} f(x_i) \cdot \overline{F}(x^*_{n+1}) \tag{2.75}$$

where x_i is defined as before, x^*_{n+1} is a suspension time and $F_X(\cdot)$ is the survivor function.

In an example where Prochan's[‖] "famous" airplane air-conditioning data is modeled, the usefulness of this approach is illustrated. The general model with $g_1(t)$ and $g_2(B(t))$ as covariates was fitted on the data. After evaluation of the significance of the covariates by means of the Wald test statistic it was clear that only $g_1(t)$ was significant, i.e. the the data was more suitable for repairable systems theory because an underlying trend was present in the data. Laplace's trend test (see De Laplace (1773)) and the test by Cox and Lewis (1966) confirmed this result.

Calabria and Pulcini (2000) presented a special case of the model by Lawless and Thiagarajah (1996) where the model determines the characteristics of the failure process during fitting procedures. The two most popular NHPPs (Power-Law Process (PLP) and Log-Linear Process (LLP)) are considered in terms of (2.73), together with the Weibull Renewal Process (WRP). The proposed models are as follows:

(i) The Power-Law Weibull Renewal process with an intensity of,

$$\iota(t|H_t) = \frac{\beta+\delta-1}{\theta^{\beta+\delta-1}} t^{\beta-1}[u(t)]^{\delta-1} \tag{2.76}$$

where $\theta > 0$, $\beta + \delta > 1$ for $0 < t \leq T_1$ and $u(t) = t - t_{N(t)}$. The intensity up to the first failure time T_1 is $\iota(t) = \gamma t^{\beta+\delta-2}$ where $\gamma = (\beta+\delta-1)/\theta^{\beta+\delta-1}$. This is a power law function which does not depend on H_t or the maintenance policy. If minimal repair was done after each failure, the failure process should evolve on the basis of the intensity. Thus, the ratio of (2.76) and the intensity up to the first failure gives a measure of the improvement or worsening introduced by the actual maintenance policy (minimal repair), i.e. $[u(t)/t]^{\delta-1}$. An indication of the departure from minimal repair is thus given by δ. For example, if $\delta > 1$, the ratio is less than 1 for any $t > T_1$ and, at a given distance $B(t)$ from the most recent failure, it becomes smaller and smaller as the number of occurred failures increases, thus indicating a repeated beneficial effect of maintenance actions on the equipment reliability.

[‖] See Prochan (1963).

University of Pretoria etd – Vlok, P-J (2006)

The parameter β in (2.76) measures the departure from perfect maintenance. If $\beta = 1$, then (2.76) reduces to Weibull renewal. When $\beta > 1$, reliability degradation is experienced and if $1 - \delta < \beta < 1$, reliability improvement is experienced.

(ii) The Log-Linear Weibull Renewal process with intensity function,

$$\lambda(t|H_t) = \delta \exp(\theta + \beta t)[u(t)]^{\delta - 1} \tag{2.77}$$

with $-\infty < \theta, \beta < \infty$ and $\delta > 0$ for $0 < t \leq T_1$. Up to the first failure, $\iota(t|H_t) = \exp(\gamma + \beta t)t^{\delta-1}$, where $\gamma = \theta + \ln \delta$, which is exactly the intensity of an ordinary Poisson process. When $\beta = 0$, (2.77) does not depend on global age but only on $B(t)$ which implies perfect maintenance, i.e. Weibull renewal. If $\delta = 1$ the process intensity does not depend on local time and reduces to a log-linear process.

The value of δ has the same physical meaning here as in the Power-Law Weibull Renewal model. But, for the value of β, if $\beta > 0$ a reliability deterioration of the equipment with the operating time is described. The more β differs from 0, the bigger the time trend. Finally, if $\beta = 0$ and $\delta = 1$, the Log-Linear Weibull Renewal process reduces to the HPP.

Likelihood construction for the above models is trivial and will not be discussed here.

2.3.7.2 Simple and robust methods for the analysis of recurrent events in repairable systems

Lawless and Nadeau (1995) considered some robust methods to estimate the behavior of point process data based on the Poisson model. This is an extension of the techniques described by Nelson (1982) for IID data.

Suppose k systems are observed and system i is under consideration over a time period $[0; T_i]$. Let $N_i(t)$ be the number of events up to time t. It follows that the Cumulative Mean Function (CMF) is $M_i(t) = E[N_i(t)]$. In the continuous sense $m_i(t) = M_i^{'}(t)$, which is the ROCOF.

To estimate the common CMF in a discrete sense, let $n_i(t) > 0$ be the number of events that occur to system i at time t. This means $m(t) = E[n_i(t)]$ and hence $M(t) = \sum_{s=1}^{t} m(s)$. System i is observed over $[0; T_i]$ and we define $\delta_i(t) = 1$ if $t \leq T_i$ and $\delta_i(t) = 0$ if $t > T_i$ to indicate whether i is observed at t. The total number of events is given by $n.(t) = \sum_{i=1}^{k} \delta_i n_i(t)$ and the total number of systems observed at t is $\delta.(t) = \sum_{i=1}^{k} \delta_i(t)$. Further, assume the k systems under observation are mutually independent. Then, if the $n_i(t)$'s are independent Poisson random variables with means $m(t)$, the MLE's of $m(t)$ is given by:

$$\hat{m}(t) = \frac{n.(t)}{\delta.(t)} \tag{2.78}$$

University of Pretoria etd – Vlok, P-J (2006)

Similarly for $M(t)$ we have,

$$\hat{M}(t) = \sum_{s=0}^{t} \frac{n.(t)}{\delta.(t)} \tag{2.79}$$

The authors present a valve seat replacement example as well as an automobile warranty claim example to illustrate the above concepts.

This publication is very useful for this study since it is robust and simple. It is ideal for a preliminary analysis of data. The assumption that the end observation times τ_i are independent of the event process may be somewhat unrealistic in reliability problems. For example, if system failures are studied and systems with many failures are withdrawn from service earlier, the estimates of $M(t)$ or the regression coefficients could be badly biased.

2.4 Conclusion

Chapter 2 covered the majority of advanced failure intensity model classes in the literature as well as a few noteworthy extensions of model classes. Kumar and Westberg (1996b) compiled a summary of advanced failure intensity models for non-repairable systems and how these models are interrelated. This summary was broadened and generalized for both non-repairable and repairable systems and is shown in Figure 2.1 on the next page.

The theory, models and concepts discussed in this chapter are used in chapters to follow to achieve the objectives set in Section 1.6.

University of Pretoria etd – Vlok, P-J (2006)

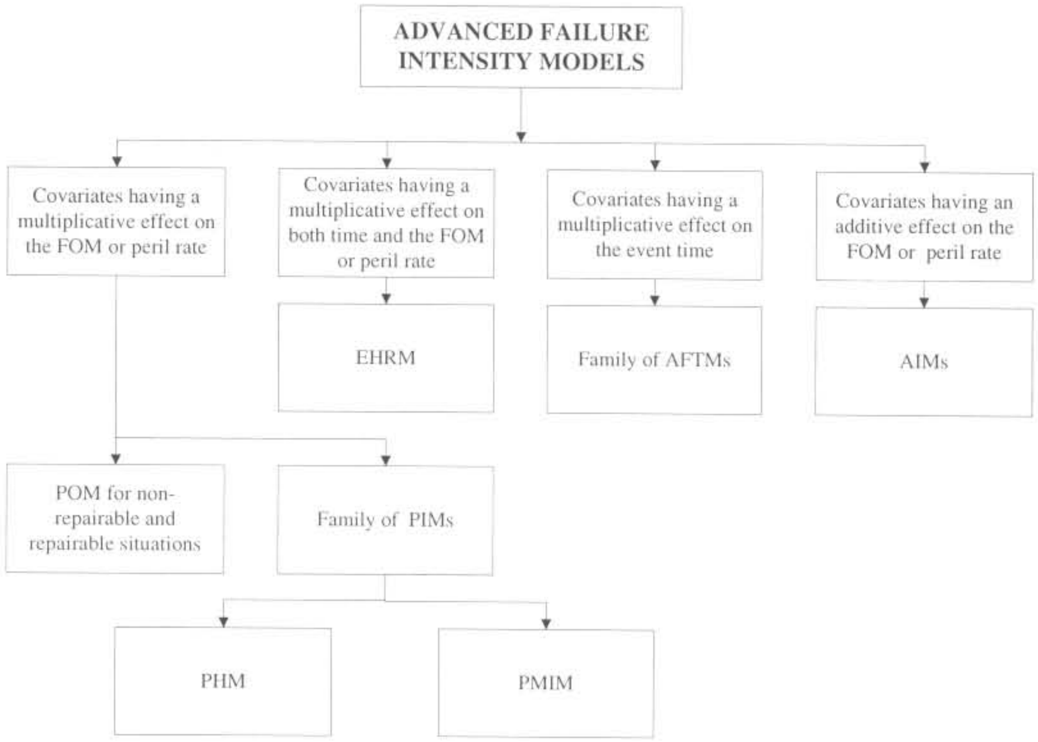

Figure 2.1: Summary of different advanced failure intensity models

University of Pretoria etd – Vlok, P-J (2006)

Chapter 3

Combined Advanced Failure Intensity Models

3.1 Introduction

In Chapter 2, advanced failure intensity models found in the literature were discussed and categorized in different model classes. The theoretical foundation, implementability and practical applicability of each model were evaluated. From this literature survey and evaluation, the following conclusions were drawn:

(i) The distinction between models applicable for non-repairable systems and repairable systems are not clear enough and are rarely emphasized in the literature. According to Ascher (1999) this contributes to the confusion between the two approaches.

(ii) Most models only consider relative risks. Relative risks are attractive because no assumption needs to be made about the underlying baseline, but it does not provide any information with regards to absolute probabilities. Absolute risks are required to utilize the techniques described in Section 1.4.2 and also to estimate residual life. For this study, models need to be fully parametric to be able to calculate absolute risks and hence residual life, as was stated in the problem statement in Section 1.6.

(iii) With the exception of the Extended Hazard Regression Model (EHRM) (see Section 2.3.3.2), models focus on one particular enhancement (such as addition, multiplication, frailty, mixed time-scales, etc.) rather than combining different enhancements. Models will be more practical if more than one enhancement is allowed in the same model.

(iv) Data sets are often modeled with only one type of model and the results of this model are accepted without comparing it to other models. Part of the reason for this practice is because it is such a laborious task to manipulate data, estimate coefficients and refine algorithms for any particular model.

University of Pretoria etd – Vlok, P-J (2006)

Subsequent to these conclusions, a methodology was established to improve the shortcomings outlined. The methodology is illustrated in Figure 3.1 and elucidated below.

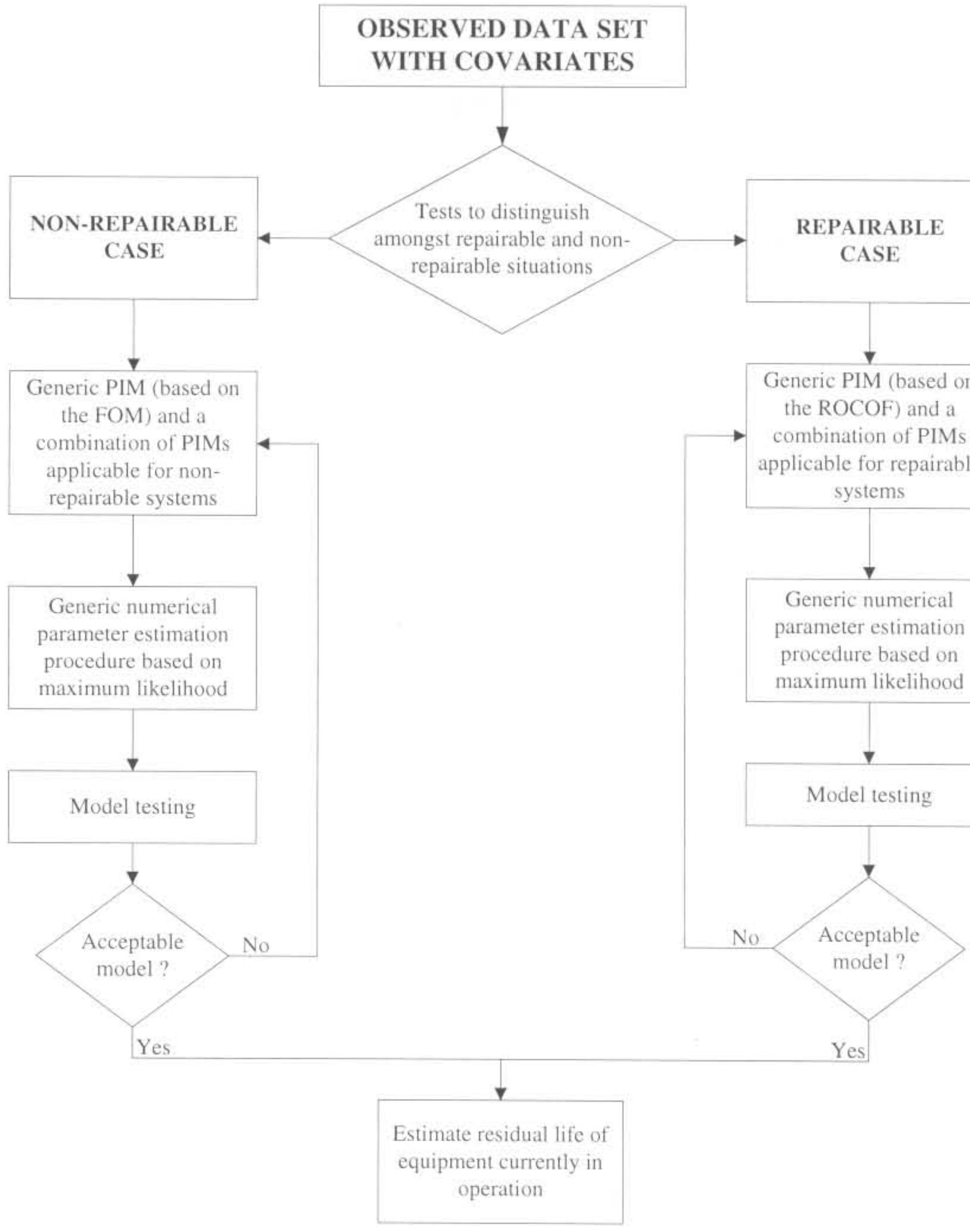

Figure 3.1: Modeling methodology

(i) Fundamentally different models will be constructed for non-repairable and repairable situations and it is therefore important to distinguish between these cases. Techniques

University of Pretoria etd – Vlok, P-J (2006)

such as those of De Laplace (1773), Bates (1955), Bartholomew (1956a), Bartholomew (1956b), Boswell (1966), Cox and Lewis (1966), Boswell and Brunk (1969), Lorden and Eisenberger (1973), Saw (1975), Bain, Engelhardt, and Wright (1985), Lawless and Thiagarajah (1996), Martz and Kvam (1996) and Vaurio (1999), as mentioned in Section 1.2.2, will be used on data sets to determine whether non-repairable or repairable systems theory is more appropriate. These techniques are applied in Chapter 5.

(ii) For both the non-repairable and repairable case, generic fully parametric PIMs will be developed that are able to simplify to the majority of models (or combination of models) described in Section 2.3. For non-repairable cases, the full intensity or conditional intensity is used, i.e. FOM, and for repairable cases the mean intensity or unconditional intensity is used, i.e. ROCOF. (See Table 2.1). Such generic models have the advantage that data can be modeled with the aid of more than one of the conventional enhancements.

(iii) Numerical parameter estimation techniques and algorithms will be developed for the generic PIMs based on maximum likelihood techniques. These algorithms will also be able to estimate the parameters of any simplification of the generic models. This simplifies the modeling processes because different models can be tested without having to develop an estimation algorithm for every special case of the generic PIMs.

(iv) Statistical techniques similar to those described by Kay (1984), Anderson (1982) and Moreau, O'Quigly, and Mesbah (1985) will be used as part of the testing of models' quality. Model quality will also be evaluated by "forecasting" observed events, following case studies by Vlok (1999) and Vlok, Coetzee, Banjevic, Jardine, and Makis (2001) that have shown that models with relatively poor statistical performance can provide very useful practical results. The motivation for this approach is discussed as part of the residual life estimation procedure in Chapter 4.

The methodology above addresses the shortcomings outlined earlier in this section. In the remainder of Chapter 3, the generic PIMs for both the non-repairable and repairable cases are developed with likelihood construction for parameter estimation. Several assumptions are made while developing the theory. These assumptions are motivated in Section 3.4 at the end of this chapter where the practical implementation of the combined advanced failure intensity models is discussed.

3.2 The non-repairable case

A single model that incorporates all the conventional model enhancements related to non-repairable systems (discussed in Chapter 2) is required. In this section such a model is developed. The following assumptions are made:

University of Pretoria etd – Vlok, P-J (2006)

(i) Multiple system copies are nominally similar and are operating in similar conditions.

(ii) All items considered are behaving according to renewal processes.

(iii) The Weibull distribution is used to parameterize models, except for the case of the POM.

(iv) Covariates are assumed to be positive.

The validity and practical implications of these assumptions are discussed in Section 3.4.

3.2.1 Model development

Suppose $k = 1, 2, .., w$ nominally similar single-part system copies are studied and the event times on each system are recorded. This scenario is illustrated in Figure 3.2.

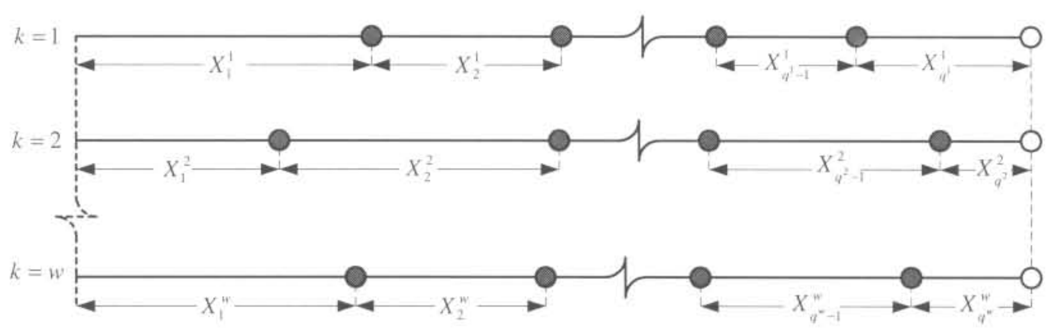

Figure 3.2: w nominally similar single-part system copies (renewed / replaced after each failure) with m time-dependent covariate measurements on each copy (Dots denote failures, circles denote suspensions)

It is assumed that tests revealed that all the particular systems can be modeled with non-repairable systems theory. Event data for each system is recorded in a $q^k \times 2$ matrix, where each row contains X_i^k, the time to event and C_i^k, the event type indicator, i.e. $C_i^k = 0$ denotes suspension and $C_i^k = 1$ denotes failure. This is similar to the approach of Wei, Lin, and Weissfeld (1989). The total observation period for any system k is $\sum X_i^k$ (for $i = 1, 2, ..., q^k$). On each system m time-dependent covariates are measured, i.e. $\boldsymbol{z}_i^k = [z_{i_1}^k(x)\ z_{i_2}^k(x)\ \ldots\ z_{i_m}^k(x)]$*, where x refers to the local time of system k during lifetime i. It is also assumed that the data is categorized in $s = 1, 2, ..., r$ different strata, where r is the highest stratum of item k. A general model that represents the FOM of any of the observed system copies is given by,

$$h(x, \boldsymbol{\theta}) = \zeta_s^k \left(g_s^k(x, \tau_s^k, \psi_s^k) \cdot \lambda(\boldsymbol{\gamma}_s^k \cdot \boldsymbol{z}_s^k) + \nu(\boldsymbol{\alpha}_s^k \cdot \boldsymbol{z}_s^k) \right) \quad (3.1)$$

*For notational convenience, the indication of the time dependence of covariates, "(x)", is suppressed in expressions to follow.

University of Pretoria etd – Vlok, P-J (2006)

where $\boldsymbol{\theta}^{\dagger}$ consists of

- k: the system copy indicator
- s: the current stratum indicator
- ζ_s^k: a random variable that acts as a frailty in the model that could be system copy- and stratum-specific
- g_s^k: a fully parametric baseline function that could be system copy- and stratum-specific
- τ_s^k: a factor that acts additively on x in g_s^k to represent a time jump or time setback that could be system copy- and stratum-specific
- ψ_s^k: a factor that acts multiplicatively on x in g_s^k to result in an acceleration or deceleration of time that could be system copy- and stratum-specific
- λ_s^k: a multiplicative functional term that is determined by $\boldsymbol{z}_s^k$ and that acts on g_s^k
- ν_s^k: an additive functional term determined by $\boldsymbol{z}_s^k$
- $\boldsymbol{z}_s^k$: a vector of time-dependent covariates

Before any comment is made on (3.1), the model will be fully parameterized first. The Weibull distribution is used throughout this thesis for a parametric baseline function because of its versatility, except for the special case of the Proportional Odds Model (see Section B.1.2). (β and η denote the Weibull shape and scale parameters respectively). Both the multiplicative and additive terms are assumed to be exponential. Every element in $\boldsymbol{\theta}$ has potentially unique values for every k and every s, i.e. $\boldsymbol{\gamma}_s^k = [\gamma_{s1}^k\ \gamma_{s2}^k\ \cdots\ \gamma_{sm}^k]$, $\zeta_s^k \in \{\zeta_1^k, \zeta_2^k, ..., \zeta_r^k\}$, $\psi_s^k \in \{\psi_1^k, \psi_2^k, ..., \psi_r^k\}$, $\tau_s^k \in \{\tau_1^k, \tau_2^k, ..., \tau_r^k\}$, $\beta_s^k \in \{\beta_1^k, \beta_2^k, ..., \beta_r^k\}$ and $\eta_s^k \in \{\eta_1^k, \eta_2^k, ..., \eta_r^k\}$. For any given value of s, k and x, the baseline function g_s^k is,

$$g_s^k(x,\boldsymbol{\theta}) = \frac{\beta_s^k}{\eta_s^k}\left(\frac{\psi_s^k\left(x-\tau_s^k\right)}{\eta_s^k}\right)^{\beta_s^k-1} \tag{3.2}$$

Similarly, for the functional terms,

$$\lambda_s^k(x,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m}\gamma_{s_j}^k \cdot z_{i_j}^k\right) \tag{3.3}$$

$$\nu_s^k(x,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m}\alpha_{s_j}^k \cdot z_{i_j}^k\right) \tag{3.4}$$

---

† $\boldsymbol{\theta}$ also include any additional parameters used in the parametric baseline.

University of Pretoria etd – Vlok, P-J (2006)

The FOM of any item, k, under consideration at any point in time, x, and for any stratum, s, is thus given by,

$$h(x,\boldsymbol{\theta}) = \zeta_s^k \left(\frac{\beta_s^k}{\eta_s^k} \left(\frac{\psi_s^k \left(x - \tau_s^k\right)}{\eta_s^k} \right)^{\beta_s^k - 1} \cdot e^{\sum_{j=1}^{m} \gamma_{s_j}^k \cdot z_{i_j}^k} + e^{\sum_{j=1}^{m} \alpha_{s_j}^k \cdot z_{i_j}^k} \right) \tag{3.5}$$

The model in (3.5) is probably unrealistic from a reliability modeling point of view because of the large number of parameters that need to be estimated. Huge data sets would be required to fit such a model. The objective with this model however, is not to use it in its complete form as presented above, but to simplify it to conventional enhanced models or to combinations of models with different enhancements by applying restrictions on some of the parameters. For example, to obtain a conventional Weibull-parameterized PHM from w system copies, the restrictions summarized in Table 3.2 is applied on (3.5).

Table 3.2: Parameter restrictions for equation (3.5) to obtain a conventional Weibull-parameterized PHM from w system copies

Parameter	Restriction		
k:	k	$=$	$1, 2, ..., w$
s:	s	$=$	1, for all values of i^k
ζ_s^k:	ζ_s^k	$=$	1, for all values of s and k
ψ_s^k:	ψ_s^k	$=$	1, for all values of s and k
τ_s^k:	τ_s^k	$=$	0, for all values of s and k
α_s^k:	$\alpha_{s_j}^k$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s and k
γ_s^k:	$\gamma_{s_j}^k$	$=$	γ, for $j = 1, 2, ..., m$ and all values of s and k
β_s^k:	β_s^k	$=$	β, for all values of s and k
η_s^k:	η_s^k	$=$	η, for all values of s and k

The restrictions in Table 3.2 applied to (3.5) gives,

$$h(x,\boldsymbol{\theta}) = \frac{\beta}{\eta} \left(\frac{x}{\eta} \right)^{\beta - 1} \cdot e^{\sum_{j=1}^{m} \gamma_j \cdot z_j} \tag{3.6}$$

To further illustrate the usefulness of (3.5), a special case of the Weibull-parameterized PWP Model 2 (similar to (2.37)) is constructed. Suppose the following requirements for the model are set:

(i) No frailty.

(ii) No accelerative or decelerative component.

University of Pretoria etd – Vlok, P-J (2006)

(iii) No PAR or PAS component, i.e. no time jump or setback.

(iv) No additive component.

(v) Strata are defined as $s = i^k$, i.e. $s = 1$ for $x \leq X_1^k$, $s = 2$ for $X_1^k < x \leq X_2^k$, etc.

(vi) Regression coefficients in the multiplicative functional term are stratum-specific but not system copy specific.

(vii) The Weibull shape parameter is neither stratum nor system copy specific.

(viii) The Weibull scale parameter is system copy specific but not stratum-specific.

To obtain the desired model, certain restrictions are applied on (3.5). These restrictions are summarized in Table 3.3.

Table 3.3: Parameter restrictions for equation (3.5) to obtain a special case of the Weibull-parameterized PWP Model 2 from w system copies

Parameter	Restriction		
k:	k	$=$	$1, 2, ..., w$
s:	s^k	$=$	i^k, for all values of i^k
ζ_s^k:	ζ_s^k	$=$	1, for all values of s and k
ψ_s^k:	ψ_s^k	$=$	1, for all values of s and k
τ_s^k:	τ_s^k	$=$	0, for all values of s and k
α_s^k:	$\alpha_{s_j}^k$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s and k
γ_s^k:	γ_s^k	$=$	γ_s, for $j = 1, 2, ..., m$ and all values of s and k
β_s^k:	β_s^k	$=$	β, for all values of s and k
η_s^k:	η_s^k	$=$	η^k, for all values of s and k

Applying the restrictions in Table 3.3 on (3.5), results in:

$$h(x, \boldsymbol{\theta}) = \frac{\beta}{\eta^k} \left(\frac{x}{\eta^k} \right)^{\beta - 1} \cdot e^{\sum_{j=1}^{m} \gamma_{s_j} \cdot z_{i_j}^k} \tag{3.7}$$

Equation (3.5) can be generalized even further by allowing for system copies that consist of multiple parts in series, where the total system success is dependent on the success of each individual part. On failure of any part, the total system is renewed or replaced. A situation of competing risks arise in such a case.

Reconsider the configuration in Figure 3.2, but suppose that each system copy now consists of $l = 1, 2, ..., n$ parts in series (see Figure 3.3). The success of the entire system is therefor dependent on the success of each individual part. It is assumed that tests confirmed the validity of non-repairable systems theory on all l parts of each of the k systems. On failure

University of Pretoria etd – Vlok, P-J (2006)

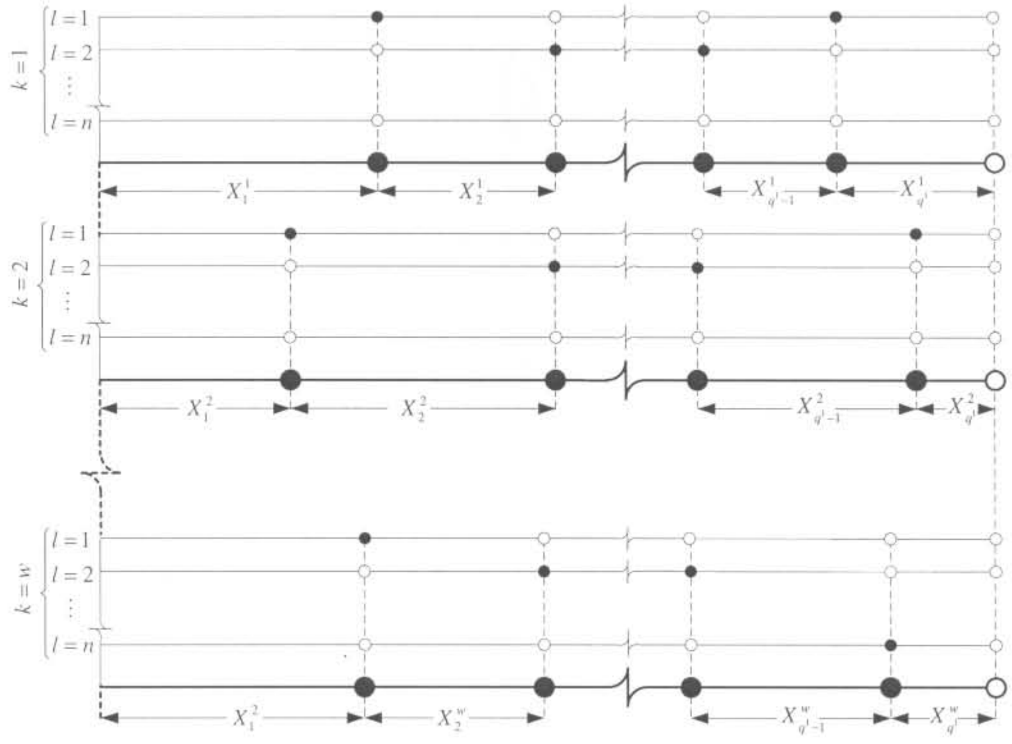

Figure 3.3: w nominally similar system copies containing n parts each with m_l time-dependent covariate measurements on each copy (Dots denote failures, circles denote suspensions)

of any of the n parts, all the parts are renewed or replaced before the system is put back into service. The event history of every part in every system is recorded in a $q^{k_l} \times 2$ matrix, consisting of event times, $X_i^{k_l}$ and event type indicators, $C_i^{k_l}$. Every system has a similar event history matrix that can be deduced from the part histories, i.e. $X_i^k = \min\{X_i^{k_l}\}$ for $l = 1, 2, ..., n$ and C_i^k corresponds to the event type of $\min\{X_i^{k_l}\}$. On each part in each system, m_l time-dependent covariates are measured, i.e. $\boldsymbol{z}_i^{k_l} = [z_{i_1}^{k_l}\ z_{i_2}^{k_l}\ ...\ z_{i_{ml}}^{k_l}]^{\ddagger}$, where x denotes the local time during lifetime i of system k. Event data is categorized in $s = 1, 2, ..., r^l$ strata, where r^l is the highest stratum of any part l. A general model for such a situation (analogous to 3.1) is given by,

$$h(x, \boldsymbol{\theta}) = \sum_{l=1}^{n} \varsigma_s^{k_l} \left(g_s^{k_l}(x, \tau_s^{k_l}, \psi_s^{k_l}) \cdot \lambda(\boldsymbol{\gamma}_s^{k_l} \cdot \boldsymbol{z}_s^{k_l}) + \nu(\boldsymbol{\alpha}_s^{k_l} \cdot \boldsymbol{z}_s^{k_l}) \right) \tag{3.8}$$

where the baseline function g for any item l, associated with system k in stratum s becomes,

$$g_s^{k_l}(x, \boldsymbol{\theta}) = \frac{\beta_s^{k_l}}{\eta_s^{k_l}} \left(\frac{\psi_s^{k_l}\left(x - \tau_s^{k_l}\right)}{\eta_s^{k_l}} \right)^{\beta_s^{k_l} - 1} \tag{3.9}$$

‡As before, "(x)" is suppressed.

University of Pretoria etd – Vlok, P-J (2006)

The functional terms become,

$$\lambda_s^{k_l}(x,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{i_j}^{k_l}\right) \quad (3.10)$$

$$\nu_s^{k_l}(x,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{i_j}^{k_l}\right) \quad (3.11)$$

The FOM for any part l in system copy k at any time in stratum s is given by,

$$h(x,\boldsymbol{\theta}) = \zeta_s^{k_l}\left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}}\left(\frac{\psi_s^{k_l}\left(x-\tau_s^{k_l}\right)}{\eta_s^{k_l}}\right)^{\beta_s^{k_l}-1} \cdot e^{\sum_{j=1}^{m_l}\gamma_{s_j}^{k_l}\cdot z_{i_j}^{k_l}} + e^{\sum_{j=1}^{m_l}\alpha_{s_j}^{k_l}\cdot z_{i_j}^{k_l}}\right) \quad (3.12)$$

while the FOM of the entire system is represented by,

$$h(x,\boldsymbol{\theta}) = \sum_{l=1}^{n}\zeta_s^{k_l}\left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}}\left(\frac{\psi_s^{k_l}\left(x-\tau_s^{k_l}\right)}{\eta_s^{k_l}}\right)^{\beta_s^{k_l}-1} \cdot e^{\sum_{j=1}^{m_l}\gamma_{s_j}^{k_l}\cdot z_{i_j}^{k_l}} + e^{\sum_{j=1}^{m_l}\alpha_{s_j}^{k_l}\cdot z_{i_j}^{k_l}}\right) \quad (3.13)$$

Equation (3.13) is more general than (3.5) and can be reduced to (3.5) by letting $n = 1$ for all values of k. Thus, the model constructed in (3.7), can also be achieved by (3.13). An advantage of the model in (3.13) is that different failure modes in systems are accommodated even if the system's condition is monitored by only one set of covariates, as is often the case in practice. For such a situation, $\boldsymbol{z}_s^{k_l} = \boldsymbol{z}_s^k$, $\boldsymbol{\gamma}_s^{k_l} = \boldsymbol{\gamma}_s^k$ and $\boldsymbol{\alpha}_s^{k_l} = \boldsymbol{\alpha}_s^k$, for all l.

In Appendix B it is shown that (3.8), and hence (3.13), can be reduced to the majority of models discussed in Section 2.3.

3.2.2 Likelihood construction

The general approach of Anderson, Borgan, Gill, and Keiding (1993) combined with the method used by Prentice, Williams, and Peterson (1981) is used to construct the likelihood for equation (3.13). Suspensions are accommodated in the likelihood, but should preferably only be used for calendar suspensions and for cases where the system was withdrawn from service preventively, according to Ascher (1999). Even if a system was withdrawn from service before failure, it is usually done for good reason, i.e. it is believed that the system is near to the end of its lifetime. In such a case it is more meaningful to include the observation as a failure as apposed to a suspension in the intensity model.

To simplify the calculation of the likelihood, data should be structured in the following way. Suppose d_s events are observed in stratum s. An auxiliary $d_s \times 4$ matrix is introduced, consisting of the chronologically ordered event times of stratum s in column 1, the corresponding

University of Pretoria etd – Vlok, P-J (2006)

event type indicators in column 2, the system on which the event occurred in column 3 and the part on the system which caused the event in column 4. In this matrix, the events are denoted $\underset{\sim}{X}_b^s$, the event indicators $\underset{\sim}{C}_b^s$, the system identifiers in $\underset{\sim}{k}_b^s$ and the part identifiers in $\underset{\sim}{l}_b^s$, for $b = 1, 2, ..., d_s$. The general form of the likelihood for (3.13) is given by,

$$L(x, \boldsymbol{\theta}) = \sum_{l=1}^{n} \left[\prod_{s=1}^{r^l} \prod_{b=1}^{d_s} h(\underset{\sim}{X}_b^s, \boldsymbol{\theta})^{\underset{\sim}{C}_b^s} \cdot \prod_{s=1}^{r^l} \prod_{b=1}^{d_s} e^{-\int_0^{\underset{\sim}{X}_b^s} h(x,\boldsymbol{\theta})\mathrm{d}x} \right] \tag{3.14}$$

For numerical convenience, the natural logarithm of the likelihood, i.e. $\ln L(x, \boldsymbol{\theta})$, is maximized because,

$$\underset{x,\boldsymbol{\theta}}{\operatorname{argmax}}\, L(x, \boldsymbol{\theta}) = \underset{x,\boldsymbol{\theta}}{\operatorname{argmax}} \ln L(x, \boldsymbol{\theta}) \tag{3.15}$$

which leads to,

$$\ln L(x, \boldsymbol{\theta}) = \prod_{l=1}^{n} \left[\underbrace{\sum_{s=1}^{r^l} \sum_{b=1}^{d_s} \ln h(\underset{\sim}{X}_b^s, \boldsymbol{\theta})^{\underset{\sim}{C}_b^s}}_{\text{Term 1}} - \underbrace{\sum_{s=1}^{r^l} \sum_{b=1}^{d_s} \int_0^{\underset{\sim}{X}_b^s} h(x, \boldsymbol{\theta})\mathrm{d}x}_{\text{Term 2}} \right] \tag{3.16}$$

Term 1 in equation (3.16) is,

$$\sum_{s=1}^{r^l} \sum_{b=1}^{d_s} \ln \left[\zeta_s^{k_l} \left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}} \left(\frac{\psi_s^{k_l} \left(\underset{\sim}{X}_b^s - \tau_s^{k_l} \right)}{\eta_s^{k_l}} \right)^{\beta_s^{k_l}-1} \cdot e^{\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{s_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{s_j}^{k_l}} \right) \right]^{\underset{\sim}{C}_b^s} \tag{3.17}$$

where $k = \underset{\sim}{k}_b^s$ and $l = \underset{\sim}{l}_b^s$. Term 2 is,

$$\sum_{s=1}^{r^l} \sum_{b=1}^{d_s} \int_0^{\underset{\sim}{X}_b^s} \zeta_s^{k_l} \left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}} \left(\frac{\psi_s^{k_l} \left(x - \tau_s^{k_l} \right)}{\eta_s^{k_l}} \right)^{\beta_s^{k_l}-1} \cdot e^{\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{s_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{s_j}^{k_l}} \right) \mathrm{d}x \tag{3.18}$$

also with $k = \underset{\sim}{k}_b^s$ and $l = \underset{\sim}{l}_b^s$.

The maximum value of equation (3.16) is found where,

$$\frac{\partial \ln L(x, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0 \tag{3.19}$$

for all values of $\boldsymbol{\theta}$. Numerical optimization techniques with which (3.19) can be obtained are described in Appendix C.

University of Pretoria etd – Vlok, P-J (2006)

3.3 The repairable case

A single model that incorporates all the conventional model enhancements related to repairable systems (discussed in Chapter 2) is required. In this section such a model is developed. The following assumptions are made:

(i) Multiple system copies are nominally similar and are operating in similar conditions.
(ii) All items considered are by the definition of Section 1.2.1, repairable systems.
(iii) NHPPs of log-linear and power-law forms are used to parameterize models.
(iv) Covariates are assumed to be positive.

The validity and practical implications of these assumptions are discussed in Section 3.4.

3.3.1 Model development

Suppose $k = 1, 2, .., w$ nominally similar single-part system copies are studied (see Figure 3.4) on which q^k observations were recorded. Assume that tests revealed that repairable systems theory is suitable to model the reliability of each of these systems. Event data of each system is saved in a $q^k \times 2$ matrix, where each row contains T_i^k, the arrival time a particular event and C_i^k, the event type indicator, i.e. $C_i^k = 0$ in case of suspension and $C_i^k = 1$ otherwise. On each system m time-dependent covariates are measured, i.e. $\boldsymbol{z}^k = [z_1^k(t)\ z_2^k(t)\ ...\ z_m^k(t)]$[§] for $i = 1, 2, ..., q^k$, where t refers to the global time of system k. It is also assumed that the data of each system is categorized in $s = 1, 2, ..., r$ different strata, where r is the highest stratum of system k.

A general model that represents the mean intensity, i.e. ROCOF, of any of the observed system copies is given by,

$$v(t, \boldsymbol{\theta}) = \zeta_s^k \left(g_s^k(t, \tau_s^k, \psi_s^k) \cdot \lambda(\boldsymbol{\gamma}_s^k \cdot \boldsymbol{z}_i^k) + \nu(\boldsymbol{\alpha}_s^k \cdot \boldsymbol{z}_i^k) \right) \tag{3.20}$$

where $\boldsymbol{\theta}$[¶] consists of

k:	the system copy indicator
s:	the current stratum indicator
ζ_s^k:	a random variable that acts as a frailty in the model that could be system copy- and stratum-specific
g_s^k:	a fully parametric baseline function that could be system copy- and stratum-specific
τ_s^k:	a factor that acts additively on t in g_s^k to represent a time

[§]For notational convenience, the indication of the time dependence of covariates, "(t)", is suppressed in expressions to follow.

[¶]$\boldsymbol{\theta}$ also include any additional parameters used in the parametric baseline.

University of Pretoria etd – Vlok, P-J (2006)

jump or time setback that could be system copy- and stratum-specific

ψ_s^k: a factor that acts multiplicatively on t in g_s^k to result in an acceleration or deceleration of time that could be system copy- and stratum-specific

λ_s^k: a multiplicative functional term that is determined by $\boldsymbol{z}_s^k$ and that acts on g_s^k

ν_s^k: an additive functional term determined by $\boldsymbol{z}_s^k$

$\boldsymbol{z}_s^k$: a vector of time-dependent covariates

The same symbols that were used to present the general combined model for the FOM of a number of system copies in (3.5), are used in (3.20) to introduce the general combined model for the ROCOF of a number of system copies. It is assumed that the different variables will be interpreted in context, i.e. as FOMs in the non-repairable case and as ROCOFs in the repairable case.

NHPP models have gained general acceptance for non-repairable situations as described in Section 1.2.2. In this thesis two types of NHPP models are used namely the log-linear process where $\rho_1 = \exp(\Gamma + \Upsilon t)$ and power-law process where $\rho_2 = \kappa\beta t^{\beta-1}$. Theory will only be developed for the log-linear process but the same principles apply for the power-law process. Both the multiplicative and additive terms are assumed to be exponential. Every element in $\boldsymbol{\theta}$ has potentially unique values for every k and every s, i.e. $\boldsymbol{\gamma}_s^k = [\gamma_{s_1}^k \; \gamma_{s_2}^k \; ... \; \gamma_{s_m}^k]$, $\zeta_s^k \in \{\zeta_1^k, \zeta_2^k, ..., \zeta_{r_k}\}$, $\psi_s^k \in \{\psi_1^k, \psi_2^k, ..., \psi_{r_k}\}$, $\tau_s^k \in \{\tau_1^k, \tau_2^k, ..., \tau_{r_k}\}$, $\Gamma_s^k \in \{\Gamma_1^k, \Gamma_2^k, ..., \Gamma_{r_k}\}$ and $\Upsilon_s^k \in \{\Upsilon_1^k, \Upsilon_2^k, ..., \Upsilon_{r_k}\}$. For any given value of s, k and t, the baseline function g_s^k is,

$$g_s^k(t, \boldsymbol{\theta}) = \exp(\Gamma_s^k + \psi_s^k \Upsilon_s^k (t - \tau_s^k)) \tag{3.21}$$

Similarly, for the functional terms,

$$\lambda_s^k(t, \boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m} \gamma_{s_j}^k \cdot z_{i_j}^k\right) \tag{3.22}$$

$$\nu_s^k(t, \boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m} \alpha_{s_j}^k \cdot z_{i_j}^k\right) \tag{3.23}$$

The peril rate of any item, k, under consideration at any point in time, t, and for any stratum, s, is thus given by,

$$\rho_1(t, \boldsymbol{\theta}) = \zeta_s^k \left(e^{\Gamma_s^k + \psi_s^k \Upsilon_s^k (t - \tau_s^k) + \sum_{j=1}^{m} \gamma_{s_j}^k \cdot z_{i_j}^k} + e^{\sum_{j=1}^{m} \alpha_{s_j}^k \cdot z_{i_j}^k} \right) \tag{3.24}$$

University of Pretoria etd – Vlok, P-J (2006)

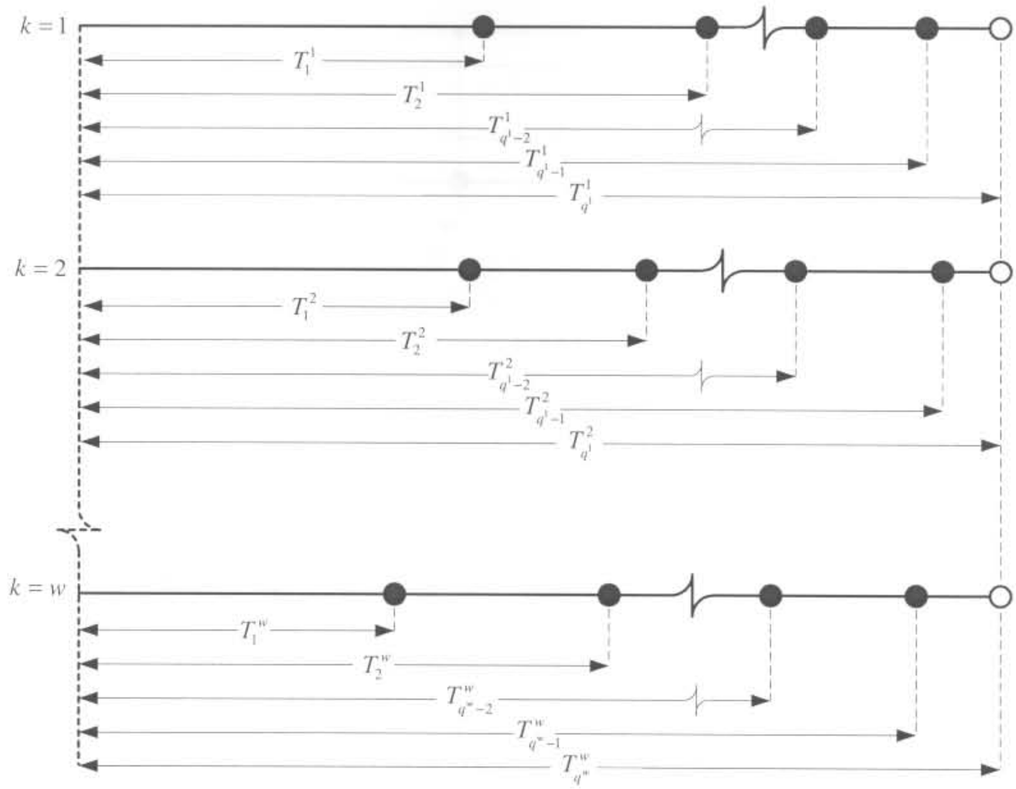

Figure 3.4: w nominally similar single-part system copies (repaired after each failure) with m time-dependent covariate measurements on each copy (Dots denote failures, circles denote suspensions)

As in the case of (3.5), the model in (3.24) is probably unrealistic from a reliability modeling point of view because of the data requirements. The objective with this model is also not to use it in its complete form as presented above, but to simplify it to conventional enhanced models or to combinations of models with different enhancements by applying restrictions on some of the parameters.

Equation (3.24) can be generalized further by allowing for system copies that consist of multiple parts in series, where the total system success is dependent on the success of each individual part. On failure of any part, only the particular part is repaired and the system is put back into service. A situation of competing risks arise in such a case.

Suppose a system is considered with parts $l = 1, 2, ..., n$ in series (see Figure 3.5) where the success of the system is dependent on the success of each individual part. Event data from each part in each system is recorded in $q^{k_l} \times 2$ matrices, consisting of event times, $T_i^{k_l}$ and event type indicators, $C_i^{k_l}$. Every system has a similar event history matrix that can be deduced from the part histories, i.e. $T_i^k = \min\{T_i^{k_l}\}$ for $l = 1, 2, ..., n$ and $C_i^{k_l}$ corresponds to

University of Pretoria etd – Vlok, P-J (2006)

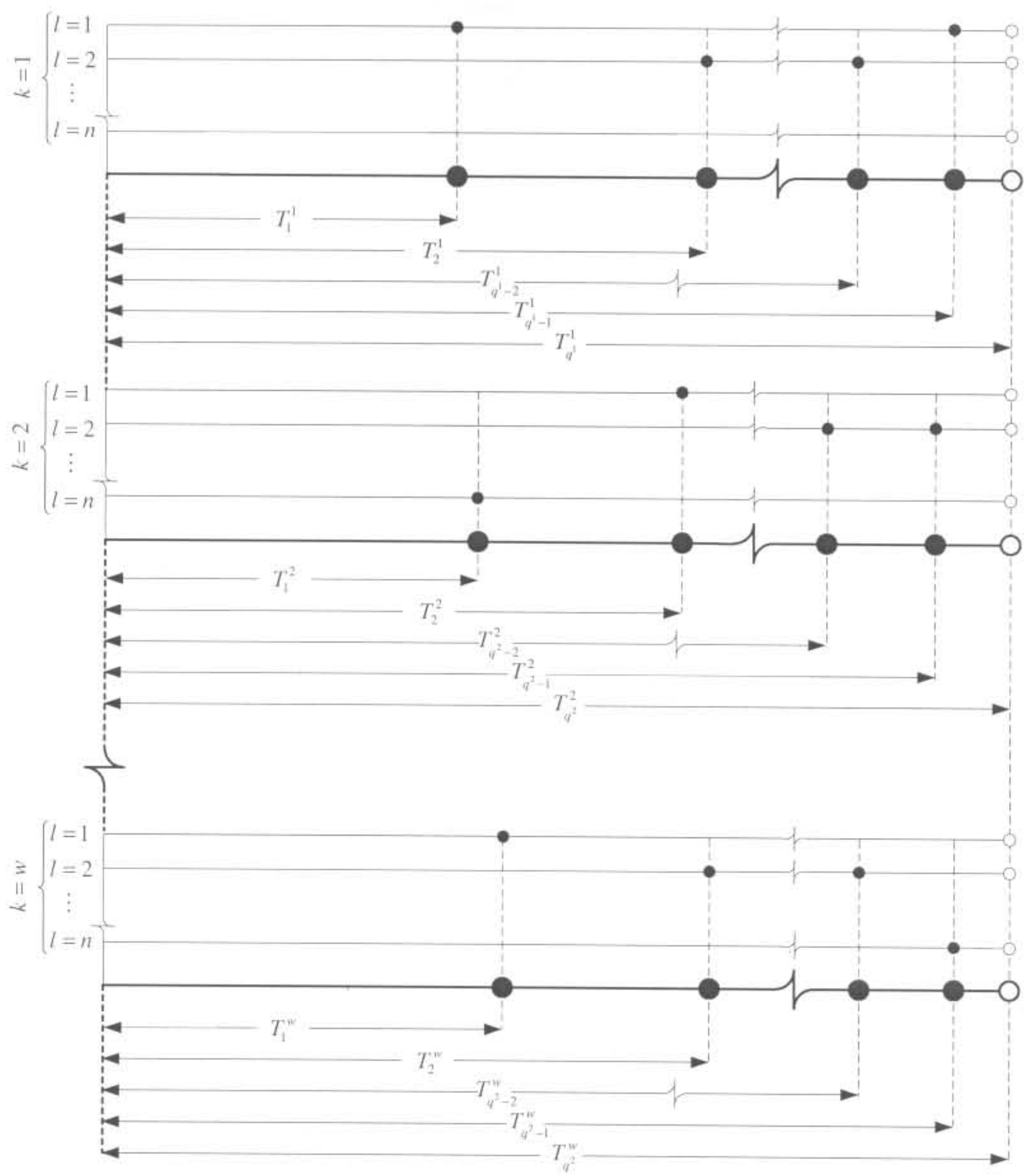

Figure 3.5: w nominally similar repairable system copies containing n parts each with m_l time-dependent covariate measurements on each copy (Dots denote failures, circles denote suspensions)

the event type of $\min\{T_i^{k_l}\}$. On each part m_l time-dependent covariates are measured, i.e. $\mathbf{z}^{k_l} = [z_1^{k_l}\ z_2^{k_l}\ ...\ z_{m_l}^{k_l}]$[||], where t refers to the global time of system k. Event data is categorized in $s = 1, 2, ..., r^l$ different strata, where r^l is the highest stratum of any part l. The general

[||]As before, "(t)" is suppressed.

University of Pretoria etd – Vlok, P-J (2006)

model for such a situation (analogous to 3.20) is given by,

$$v(t,\boldsymbol{\theta}) = \sum_{l=1}^{n} \zeta_s^{k_l} \left(g_s^{k_l}(t, \tau_s^{k_l}, \psi_s^{k_l}) \cdot \lambda(\boldsymbol{\gamma}_s^{k_l} \cdot \boldsymbol{z}_i^{k_l}) + \nu(\boldsymbol{\alpha}_s^{k_l} \cdot \boldsymbol{z}_i^{k_l}) \right) \tag{3.25}$$

where the baseline function g for any item l, associated with system k in stratum s becomes,

$$g_s^{k_l}(t,\boldsymbol{\theta}) = \exp(\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l}(t - \tau_s^{k_l})) \tag{3.26}$$

The coefficients $\beta_s^{k_l}$, $\eta_s^{k_l}$, $\psi_s^{k_l}$ and $\tau_s^{k_l}$ can not be represented as matrices because different systems could be in different strata. The functional terms become,

$$\lambda_s^{k_l}(t,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_j^{k_l} \right) \tag{3.27}$$

$$\nu_s^{k_l}(t,\boldsymbol{\theta}) = \exp\left(\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_j^{k_l} \right) \tag{3.28}$$

The peril rate for any part l in a system copy k at any time t in stratum s is given by,

$$\rho_1(t,\boldsymbol{\theta}) = \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l}(t-\tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_j^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_j^{k_l}} \right) \tag{3.29}$$

while the peril rate of an entire system is represented by,

$$\rho_1(t,\boldsymbol{\theta}) = \sum_{l=1}^{n} \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l}(t-\tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_j^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_j^{k_l}} \right) \tag{3.30}$$

Equation (3.30) is more general than (3.24) and can be reduced to (3.24) by letting $n = 1$ for all values of k. The biggest advantage of the model in (3.30) is that different failure modes in systems are accommodated even if the system's condition is monitored by only one set of covariates, as is often the case in practice. For such a situation, $\boldsymbol{z}_s^{k_l} = \boldsymbol{z}_s^k$, $\boldsymbol{\gamma}_s^{k_l} = \boldsymbol{\gamma}_s^k$ and $\boldsymbol{\alpha}_s^{k_l} = \boldsymbol{\alpha}_s^k$, for all l.

In Appendix B it is shown that (3.30) can be reduced to the majority of models discussed in Section 2.3.

3.3.2 Likelihood construction

The general approach of Anderson, Borgan, Gill, and Keiding (1993) combined with the method used by Prentice, Williams, and Peterson (1981) is used to construct the likelihood for equation (3.30). Suspensions are accommodated in the likelihood, but should preferably

University of Pretoria etd – Vlok, P-J (2006)

only be used for calendar suspensions and for cases where the system was withdrawn from service preventively, according to Ascher (1999). Even if a system was withdrawn from service before failure, it is usually done for good reason, i.e. it is believed that the system is near to the end of its lifetime. In such a case it is more meaningful to include the observation as a failure as apposed to a suspension in the intensity model.

To simplify the calculation of the likelihood, data should be structured in the following way. Suppose d_s events are observed in stratum s. An auxiliary $d_s \times 4$ matrix is introduced, consisting of the chronologically ordered event times of stratum s in column 1, the corresponding event type indicators in column 2, the system on which the event occurred in column 3 and the part on the system that caused the event in column 4. In this matrix, the events are denoted $\underset{\sim}{T}_b^s$, the event indicators $\underset{\sim}{C}_b^s$, the system identifiers as $\underset{\sim}{k}_b^s$ and the part identifiers as $\underset{\sim}{l}_b^s$ for $b = 1, 2, ..., d_s$.

The general form of the likelihood for (3.30) is given by,

$$L(t,\boldsymbol{\theta}) = \sum_{l=1}^{n} \left[\prod_{s=1}^{r^l} \left[\prod_{b=1}^{d_s} \rho_1(\underset{\sim}{T}_b^s, \boldsymbol{\theta})^{\underset{\sim}{C}_b^s} \cdot \left(e^{-\int_0^{\underset{\sim}{T}_b^s} \rho_1(t,\boldsymbol{\theta})\mathrm{d}t} \right)^{\underset{\sim}{C}_b^s} \right] \cdot e^{-\int_0^{\underset{\sim}{T}_{d_s}^s} \rho_1(t,\boldsymbol{\theta})\mathrm{d}t} \right] \tag{3.31}$$

For numerical convenience, the natural logarithm of the likelihood, i.e. $\ln L(t,\boldsymbol{\theta})$, is maximized because,

$$\underset{t,\boldsymbol{\theta}}{\operatorname{argmax}}\, L(t,\boldsymbol{\theta}) = \underset{t,\boldsymbol{\theta}}{\operatorname{argmax}} \ln L(t,\boldsymbol{\theta}) \tag{3.32}$$

which leads to,

$$\ln L(t,\boldsymbol{\theta}) = \prod_{l=1}^{n} \left[\sum_{s=1}^{r^l} \left[\left[\underbrace{\sum_{b=1}^{d_s} \underset{\sim}{C}_b^s \ln \rho_1(\underset{\sim}{T}_b^s, \boldsymbol{\theta})}_{\text{Term 1}} - \underbrace{\sum_{b=1}^{d_s} \underset{\sim}{C}_b^s \int_0^{\underset{\sim}{T}_b^s} \rho_1(t,\boldsymbol{\theta})\mathrm{d}t}_{\text{Term 2}} \right] - \underbrace{\int_0^{\underset{\sim}{T}_{d_s}^s} \rho_1(t,\boldsymbol{\theta})\mathrm{d}t}_{\text{Term 3}} \right] \right] \tag{3.33}$$

In equation (3.33) Term 1 is,

$$\sum_{b=1}^{d_s} \underset{\sim}{C}_b^s \ln \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l} (\underset{\sim}{T}_b^s - \tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} \right) \tag{3.34}$$

where $k = \underset{\sim}{k}_b^s$ and $l = \underset{\sim}{l}_b^s$. Term 2 is,

$$\sum_{b=1}^{d_s} \underset{\sim}{C}_b^s \int_0^{\underset{\sim}{T}_b^s} \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l} (t - \tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} \right) \mathrm{d}t \tag{3.35}$$

University of Pretoria etd – Vlok, P-J (2006)

with $k = \underset{\sim}{k}_b^s$ and $l = \underset{\sim}{l}_b^s$ and Term 3 is,

$$\int_0^{\underset{\sim}{T}_b^s} \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l} (t - \tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{b_j}^{k_l}} \right) \mathrm{d}t \tag{3.36}$$

with $k = \underset{\sim}{k}_b^s$ and $l = \underset{\sim}{l}_b^s$.

The maximum value of equation (3.33) is found where,

$$\frac{\partial \ln L(x, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0 \tag{3.37}$$

for all values of $\boldsymbol{\theta}$. Numerical optimization techniques with which (3.37) can be obtained are described in Appendix C.

3.4 Practical implementation of the combined models

In this section some comments are made with regards to the practical implementation of the combined models. As part of this, the validity of the assumptions for the models are also considered.

3.4.1 Comments on the assumption that covariates are always positive

Covariates were restricted to be positive during the model development in order to simplify the specification of restrictions. For example, in the non-repairable case, to restrict λ to 0 it is simply required to fix all elements of $\boldsymbol{\gamma}$ to $-\infty$, i.e. $\lambda(x, \boldsymbol{\theta}) = \exp(\sum -\infty \cdot z_j^{k_l}) = 0$, for all valid values of j. The assumption of positive covariates has no other influence on the combined models. Positive and negative covariates and also decreasing covariates do play a role in the estimation of residual life. This is discussed in Chapter 4.

3.4.2 Different modeling scenarios

Four different modeling scenarios are identified and numbered in Figure 3.6.

The different scenarios correspond to the following equations: (1) Equation (3.5); (2) Equation (3.13); (3) Equation (3.24); and (4) Equation (3.30). While developing these equations, it was assumed for Scenarios (1) and (2) that all parts of all system copies form part of renewal processes and that all parts of all system copies form part of NHPPs for Scenarios (3) and (4). In practice this will probably seldom be true but by separating renewal processes from NHPPs in "mixed" data sets, this can be overcome. In Table 3.5 below, scenarios

University of Pretoria etd – Vlok, P-J (2006)

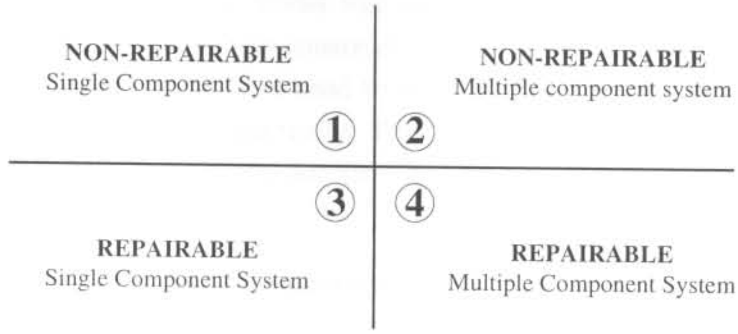

Figure 3.6: Modeling scenarios

other than the four in Figure 3.6 are sketched, i.e. mixed scenarios, with proposed modeling methodologies.

Table 3.5: Methodologies to model mixed scenarios

Scenario	Proposed modeling approach
Model w single part system copies of which y forms part of renewal processes and $w-y$ forms part of NHPPs.	Model the y renewal systems with Scenario (1) of Figure 3.6 and the $w-y$ remaining system with Scenario (3) of Figure 3.6. To estimate the next event time of a renewal system, use the model calculated for the y systems and to estimate the next event time for a repairable system, use the model for the $w-y$ systems.
Model w system copies consisting of n parts each. Of part l, y^l copies follow a renewal process while $n-y^l$ behave according to NHPP processes.	Model the $\sum y^l$ renewal copies with Scenario (2) of Figure 3.6 and the $\sum(n-y^l)$ remaining copies with Scenario (3) of Figure 3.6. To estimate the next event time of a renewal system, use the model calculated for the $\sum y^l$ copies and to estimate the next event time for a repairable system, use the model for the $\sum(n-y^l)$ systems.

Table 3.5 emphasizes the importance of separating renewal processes and repairable systems as was discussed in Section 1.2.2.

3.5 Conclusion

The models developed in this chapter primarily arose from a need to include more than one conventional enhancement in the same model. Generic models were developed to address this need with a clear distinction between the non-repairable and repairable cases.

University of Pretoria etd – Vlok, P-J (2006)

For non-repairable cases, a generic model was constructed to estimate the FOM while for repairable cases, a generic model was constructed to estimate the peril rate. The Weibull distribution and log-linear NHPP were used to parameterize the generic model for the non-repairable and repairable cases, respectively. This was done to be able to calculate absolute risks and eventually estimate residual life (in Chapter 4). A summary of these models is presented in Table 3.6.

Table 3.6: Summary of generic models

Non-repairable Case[**]
w single part system copies (all forming part of a renewal process):
$$h(x,\boldsymbol{\theta}) = \zeta_s^k \left(\frac{\beta_s^k}{\eta_s^k} \left(\frac{\psi_s^k \left(x - \tau_s^k\right)}{\eta_s^k} \right)^{\beta_s^k - 1} \cdot e^{\sum_{j=1}^{m} \gamma_{s_j}^k \cdot z_{i_j}^k} + e^{\sum_{j=1}^{m} \alpha_{s_j}^k \cdot z_{i_j}^k} \right)$$
w system copies consisting of n parts in series each, where every part forms part of a renewal process:
$$h(x,\boldsymbol{\theta}) = \sum_{l=1}^{n} \zeta_s^{k_l} \left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}} \left(\frac{\psi_s^{k_l} \left(x - \tau_s^{k_l}\right)}{\eta_s^{k_l}} \right)^{\beta_s^{k_l} - 1} \cdot e^{\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{i_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{i_j}^{k_l}} \right)$$
Repairable Case[††]
w single part system copies (all forming part of a NHPP):
$$\rho_1(t,\boldsymbol{\theta}) = \zeta_s^k \left(e^{\Gamma_s^k + \psi_s^k \Upsilon_s^k (t - \tau_s^k) + \sum_{j=1}^{m} \gamma_{s_j}^k \cdot z_{i_j}^k} + e^{\sum_{j=1}^{m} \alpha_{s_j}^k \cdot z_{i_j}^k} \right)$$
w system copies consisting of n parts in series each, where every part forms part of a NHPP:
$$\rho_1(t,\boldsymbol{\theta}) = \sum_{l=1}^{n} \zeta_s^{k_l} \left(e^{\Gamma_s^{k_l} + \psi_s^{k_l} \Upsilon_s^{k_l} (t - \tau_s^{k_l}) + \sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_j^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_j^{k_l}} \right)$$

The models in Table 3.6 are generic and it was proved in Appendix B that, in most cases, they can be reduced to the models considered in the literature survey of Section 2.3. Data constraints encountered in practice make these generic models unrealistic in their complete form but provide a basis from which simpler models (with more than one enhancement) can be derived. This concludes point (i) of Section 1.6 - the problem statement.

[**] Variables for the models corresponding to the non-repairable case are declared and described in Section 3.2

[††] Variables for the models corresponding to the repairable case are declared and described in Section 3.3

University of Pretoria etd – Vlok, P-J (2006)

Chapter 4

Estimating Residual Life based on Failure Intensities

4.1 Introduction

Chapter 4 deals with Residual Life Estimation (RLE) of items based on observed FOMs (in the non-repairable case) and peril rates (in the repairable case). The intention was not to develop new theory for RLE but to apply existing theory to the combined models that were developed in Chapter 3. Three steps in the process of estimating residual life are identified:

(i) Prediction of covariate behaviour because covariates are assumed to be time-dependent.

(ii) Calculating residual life based on observed FOMs or peril rates with the assistance of covariate behaviour predictions.

(iii) Presentation of results in a comprehensible manner.

These steps are discussed thoroughly in the remainder of the chapter.

In Step (i) covariates are assumed to be time-dependent for more generality but also because of practicality. This assumption requires prediction of future covariate behaviour to be able to estimate residual life. Various approaches can be followed to perform the covariate behaviour prediction and it is discussed in the first part of this chapter.

Residual life calculations are based on the models for items' FOMs or peril rates developed in Chapter 3, i.e. equations (3.13) and (3.30). A detailed literature survey was done on this subject and subsequent procedures were based on the literature survey. Step (ii) is the ultimate objective of this thesis.

To make this study more useful, a method to present the results of calculations to maintenance practitioners in a user-friendly manner, is proposed at the end of this chapter to conclude

University of Pretoria etd – Vlok, P-J (2006)

Step (iii). This was identified as one of the main research objectives in Section 1.6 and should be achieved to make a contribution to practical reliability modeling.

4.2 Covariate characteristics and behaviour prediction

Section 4.2 covers covariate characteristics and the prediction of covariate behaviour. Before techniques for covariate behaviour prediction can be considered, three covariate characteristics need to be discussed. Covariates can be either,

(i) time-dependent or time-independent;
(ii) internal or external; or
(iii) stochastic or non-stochastic.

These characteristics are discussed in the next three subsections after which techniques are discussed to predict stochastic and non-stochastic covariate behaviour. The section ends with formal assumptions on covariate characteristics (in Section 4.2.6) that is applied in the remainder of the thesis.

4.2.1 Time-dependent vs. time-independent covariates

In this thesis it is consistently assumed that covariates are time-dependent, not only for generality but for more practicality. If covariates were time-independent, the section on prediction of future covariate values would be unnecessary since covariate values would be known and would remain constant. PIMs and residual life calculations with time-independent covariates are special cases of models that allow for time-dependent covariates.

4.2.2 Internal vs. external covariates

Covariates can be either internal or external. This subject is discussed in detail by Fahrmeir and Tutz (1994). External covariates can be measured on a system regardless of whether an event has occurred on the system and the value of an external covariate is not changed materially by the occurrence of an event. An example of an external covariate is the ambient temperature close to a system.

Internal covariates, on the contrary, can only be meaningfully measured on a system before an event has occurred. The value of an internal covariate generally changes dramatically after the event occurs and is generally uninterpretable after the event. An example from the

University of Pretoria etd – Vlok, P-J (2006)

biomedicine field is a living organism's heartbeat, which is by definition zero after the event of death.

The type of covariate, i.e. internal or external, does not play a mathematical role in predicting covariate behaviour but it is an important aspect to consider when constructing a model for covariate behaviour. In the case of internal covariates for example, it is important to study the effect of event-type on covariates before it is attempted to model the covariate behaviour.

4.2.3 Stochastic vs. non-stochastic covariates

Covariates can be either stochastic or non-stochastic. For the stochastic case, covariates can only be predicted within certain confidence bounds and an exact prediction is not possible. At least three techniques exist to model stochastic covariates, i.e. time series analysis, state space models and Markov chains. Both time series analysis and state space models requires large quantities of data (observations in this context) to produce reasonable models. Markov chains are less dependent on large data sets and because the case study of Chapter 5 deals with a fairly small data set, only Markov chains are considered in Section 4.2.4. For more detail on the theory and application of state space models, see Cmiel and Gurgul (2000), Christer, Wang, and Sharp (1997) and Wang, Wang, and Mao (1999). Harvey (1981) and Chatfield (1980) provide an introduction to time series analysis.

Non-stochastic covariates can, by definition, be predicted with reasonable accuracy. Two cases of non-stochastic covariates exist however. In the first case, covariates are known from the origin of time and observations are unnecessary in intensity models, e.g. the complete peril rate of a system would simply be a complex parametric formulation of a NHPP as a function of time. In the second case, information about a covariate's behaviour up to a point x or t is required to be able to predict covariate behaviour beyond x or t. These cases are considered in Section 4.2.5.

4.2.4 Predicting stochastic covariate behaviour

Markov chains have been used by, amongst other, Makis and Jardine (1992), Makis and Jardine (1991), Vlok (1999) and Vlok, Coetzee, Banjevic, Jardine, and Makis (2001) to predict covariate behaviour. Other authors that have applied Markov chains in reliability include, Lagakos, Sommer, and Zelen (1978), Ng (1999), Billard and Meshkani (1995), Collins (1973) and Zhang and Love (2000). Christer and Wang (1995) oppose the use of Markov chains to model covariate behaviour because present covariate levels are in practice more often than not dependent on immediately preceding levels. Ross (1990) and Hines and Montgomery (1980) discuss Markov chains in detail. In this section an overview of the theory required to

University of Pretoria etd – Vlok, P-J (2006)

predict covariate behaviour with Markov chains is presented.

Covariate states have to be defined for the covariates before it can be modeled with Markov chains. For this reason, every range of covariate values is divided into appropriate intervals or bands and every covariate band is defined as a covariate state. Covariate bands are then used as boundaries for the transition probabilities in the Transition Probability Matrix (TPMX). For numerical convenience, 4 or 5 bands are usually selected between upper and lower bands except for the last band which does not have an upper bound.

Following the covariate states, suppose that $\{X_0, X_1, X_2, ...\}$ is a multidimensional Markov process which makes up an item's event history such that $X_k = (z_{k1}(x), z_{k2}(x), ..., z_{km}(x)) \in \mathbb{R}^m$, where m is the number of covariates, and $z_{ki}(x)$ is the k^{th} observation of variable i before an event, performed at time $x = k\Delta$ where Δ is a fixed inspection interval. A stochastic process $\{X_0, X_1, X_2, ...\}$ is assumed to be Markovian if, for every $k \geqslant 0$,

$$P\{X_{k+1} = j | X_k = i, X_{k-1} = i_{k-1}, X_{k-2} = i_{k-2}, ..., X_0 = i_0\} = P\{X_{k+1} = j | X_k = i\} \tag{4.1}$$

where $j, i, i_0, i_1, ..., i_{k-1}$ are defined states of the process, in this case the covariate bands.

The transition probability for any covariate in state i to undergo a transition to state j for a given inspection interval Δ is,

$$P_{ij}(k) = P_{ij}(k, \Delta) = P(X_{k+1} = j | X > (k+1)\Delta, X_k = i) \tag{4.2}$$

where X denotes time to event as before and i and j denote any two possible states.

Suppose a sample $X_0, X_1, X_2, ...$ is observed and let $n_{ij}(k)$ denote the number of transitions from state i to j at k throughout the sample, where the sample may contain several histories, i.e.

$$n_{ij}(k) = \#\{X_k = i, X_{k+1} = j\} \tag{4.3}$$

Similarly, the number of transitions from i at time $k\Delta$ to any other state can be calculated by,

$$n_i(k) = \#\{X_k = i\} = \sum_j n_{ij}(k) \tag{4.4}$$

It is hence possible to estimate the probability of a transition from state i to state j at time $k\Delta$ with the following relationship derived with the maximum likelihood method,

$$\hat{P}_{ij}(k) = \frac{n_{ij}(k)}{n_i(k)} \tag{4.5}$$

University of Pretoria etd – Vlok, P-J (2006)

If it is assumed that the Markov chain is homogeneous within the interval $a \leqslant k \leqslant b$, i.e. $P_{ij}(k) = P_{ij}(a)$, the transition probability can be estimated by,

$$\hat{P}_{ij}(k) = \frac{\sum\limits_{a \leqslant k \leqslant b} n_{ij}(k)}{\sum\limits_{a \leqslant k \leqslant b} n_i(k)} \tag{4.6}$$

It would also be possible to assume that the entire Markov chain is homogeneous, then $P_{ij} = P_{ij}(k)$, for $k = 0, 1, 2, ...$ and hence the transition probabilities are estimated by,

$$\hat{P}_{ij} = \frac{n_{ij}}{n_i}, \text{ where } n_{ij} = \sum_{k \geqslant 0} n_{ij}(k), \; n_i = \sum_j n_{ij} \tag{4.7}$$

As mentioned before, covariates are assumed to be time-dependent by default. For this reason continuous time is divided into u intervals, $(a_1, a_2], ..., (a_u, \infty)$, in which the transition probabilities are considered to be homogeneous. This manipulation simplifies the calculation of the TPMX considerably without loosing much accuracy.

The estimations of the TPMX above assumed that the inspection interval Δ was constant. In practice, this is rarely the case. This would mean that recorded data with inspection intervals different from Δ have to be omitted from TPMX calculations, thereby loosing valuable information about the covariates' behavior. To overcome this problem a technique utilizing transition densities (or rates) is used. Assume that the Markov chain is homogeneous for a short interval of time. The probability of transition from $i|_{x=0} \rightarrow j|_{x=x}$ is $P_{ij}(x) = P(X(x) = j|X(0) = i)$ and the rate at which the transition will take place is $D_x[P_{ij}(x)] = \lambda_{ij}$ $(i \neq j)$. For the case where $i = j$ the transition rate can be derived with the following argument. Suppose the system is in state $i|_{x=0}$ and state $j|_{x=x}$ with r possible states. If the sum over all probabilities over x is taken,

$$\begin{aligned} P_{i0}(x) + P_{i1}(x) + P_{i2}(x) + ... + P_{ir}(x) = 1 \\ \sum_j P(X(x) = j|X(0) = i) = 1 \\ \text{or } \sum_j P_{ij}(x) = 1 \end{aligned} \tag{4.8}$$

If the time derivative is taken,

$$\begin{aligned} \sum_j \frac{\partial}{\partial x} [P_{ij}(x)] = 0 \\ \therefore \;\; \lambda_{i0} + \lambda_{i1} + \lambda_{ii} + ... + \lambda_{ir} = 0 \\ \lambda_{ii} = -\sum_{i \neq j} \lambda_{ij} \end{aligned} \tag{4.9}$$

University of Pretoria etd – Vlok, P-J (2006)

The value of any λ_{ij} $(i \neq j)$ can be estimated by,

$$\hat{\lambda}_{ij} = \frac{n_{ij}}{\Omega_i}, \quad n_{ij} = \sum_k n_{ij}(k) \tag{4.10}$$

where, k runs over the given interval of time and Ω_i is the total length of time that a state is occupied in the sample. The calculation of the transition rates can be generalized for the system from any state i to j at any time x with,

$$P'_{ij}(x) = \sum_l P_{il}(x)\lambda_{lj} \tag{4.11}$$

Equation (4.11) provides a system of differential equations that has to be solved to obtain the transition probability matrix. A solution to the system of differential equations solution is,

$$P(x) = \exp(A \cdot x) \tag{4.12}$$

where $P(x) = (P_{ij}(x))$ and $A = (\lambda_{ij})^*$. This can be calculated by the series,

$$P(x) = \sum_{n=0}^{\infty} A^n \frac{x^n}{n!} \tag{4.13}$$

which is fast and accurate. Statistical tests (such as χ^2) can be used to confirm the validity of the homogeneity assumption over the given time intervals.

4.2.5 Predicting non-stochastic covariate behaviour

When selecting a parametric function to predict future covariate behaviour, the first option should be to select a function that has a physical relationship to the observed phenomenon. In vibration analysis, for example, it is expected that the spectral component related to unbalance would increase quadratically with increasing rotational velocity according to Rao (1995). If this spectral component is used as a covariate, its future behaviour should typically be modeled by a parametric function of some parabolic type. Selecting an appropriate parametric function should as far as possible not be a curve fitting exercise but rather a physical interpretation of the actual situation.

Rao (1980) formulated a few basic parametric functions that could be used to predict covariate behaviour[†]. These functions with solutions to their parameters are summarized in Table 4.1.

[*] Brackets denote matrices.

[†] These parametric functions are general and were not intended to predict covariate behaviour.

University of Pretoria etd – Vlok, P-J (2006)

Table 4.1: Parametric functions suitable to predict covariate behaviour

Linear curve	
Form	$y = ad + b$
Solution	$a = \frac{(d_1 - \overline{d})(y_1 - \overline{y}) + (d_2 - \overline{d})(y_2 - \overline{y}) + \ldots + (d_n - \overline{d})(y_n - \overline{y})}{(d_1 - \overline{d})^2 + (d_2 - \overline{d})^2 + \ldots + (d_n - \overline{d})^2}$ $b = \overline{y} - a\overline{d}$
Quadratic curve	
Form	$y = ad^2 + bd + c$
Solution	$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} d_i^2 + b\sum_{i=1}^{n} d_i + cn$ $\sum_{i=1}^{n} d_i y_i = a\sum_{i=1}^{n} d_i^3 + b\sum_{i=1}^{n} d_i^2 + c\sum_{i=1}^{n} d_i$ $\sum_{i=1}^{n} d_i^2 y_i = a\sum_{i=1}^{n} d_i^4 + b\sum_{i=1}^{n} d_i^3 + c\sum_{i=1}^{n} d_i^2$
Hyperbolic curve	
Form	$y = a/d + b$
Solution	$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} \frac{1}{d_i} + bn$ $\sum_{i=1}^{n} \frac{y_i}{d_i} = a\sum_{i=1}^{n} \frac{1}{d_i^2} + b\sum_{i=1}^{n} \frac{1}{d_i}$
Exponential curve	
Form	$y = ab^d$
Solution	$\sum_{i=1}^{n} \log y_i = n\log a + \log b\sum_{i=1}^{n} d_i$ $\sum_{i=1}^{n} d_i \log y_i = \log a\sum_{i=1}^{n} d_i + \log b\sum_{i=1}^{n} d_i^2$
Geometric curve	
Form	$y = ad^b$
Solution	$\sum_{i=1}^{n} \log y_i = n\log a + \log b\sum_{i=1}^{n} \log d_i$ $\sum_{i=1}^{n} d_i \log y_i = \log a\sum_{i=1}^{n} d_i + \log b\sum_{i=1}^{n} \log d_i^2$

A simple technique to test the goodness of fit of a straight line to a particular data set, is to

University of Pretoria etd – Vlok, P-J (2006)

calculate the correlation coefficient, R, by

$$R = \frac{(d_1 - \overline{d})(y_1 - \overline{y}) + (d_2 - \overline{d})(y_2 - \overline{y}) + ... + (d_n - \overline{d})(y_n - \overline{y})}{\sqrt{[(d_1 - \overline{d})^2 + ... + (d_n - \overline{d})^2]\,[(y_1 - \overline{y})^2 + ... + (y_n - \overline{y})^2]}} \tag{4.14}$$

To be able to calculate R for the non-linear functions in Table 4.1, data should be linearized first. A summary of the linearization procedure is shown in Table 4.2.

Table 4.2: Linearization of data to calculate correlation coefficient

Curve	d_i	y_i
Quadratic	d_i	$y_{i+1} - y_i$
Hyperbolic	$1/d_i$	y_i
Exponential	d_i	$\log y_i$
Geometric	$\log d_i$	$\log y_i$

Rao (1980) suggests that for an acceptable fit, $0.88 < R < 1$ for 5 samples, $0.28 < R < 1$ for 50 samples and $0.20 < R < 1$ for 100 samples.

It would also be possible to fit a n^{th} order polynomial through the covariate history and then predict future covariate values. This should preferably only be done for fixed inspection intervals. A n^{th} order polynomic approximation for a particular future covariate value, y_i at instant d_i is given by,

$$y_i = a_n d_i^n + a_{n-1} d_i^{n-1} + ... + a_1 d_i + a_0 \tag{4.15}$$

Ellis and Gulick (1990) describe numerical methods to solve for the coefficients of non-linear systems such as (4.15).

4.2.6 Assumptions on covariate characteristics

According to Banjevic (2001), assumptions on covariate characteristics should be based primarily on the specific situation that is considered since there are too many different possible scenarios to generalize these assumptions. The assumptions on covariate behaviour made in this thesis are however generalized to a certain extent, since it is valid for the majority of situations where condition monitoring measurements are used as covariates. A summary of the assumptions is presented in Table 4.3.

University of Pretoria etd – Vlok, P-J (2006)

Table 4.3: Summary of assumptions on covariate behaviour

Covariate Characteristic	Assumption
Time dependency	Time-dependent covariates are assumed. The majority of diagnostic parameters measured on equipment in industry are functions of time. Constant covariates are special cases of time varying covariates.
Internal vs. external	No assumption needs to be made with regards to this characteristic but it will be considered when postulating the method of covariate behaviour prediction.
Stochasticity	Non-stochastic covariates are assumed for two reasons. Firstly, it is believed that covariate behaviour in a condition monitoring environment can be predicted with reasonable accuracy provided information about the covariate is available up to a certain time x or t. Secondly, it is shown in Section 4.5 that this approach has more appeal to maintenance practitioners.

4.3 Residual life estimation based on an observed FOM

Section 4.3 discusses RLE of a system based on an observed FOM. Relevant literature is presented in Subsection 4.3.1 after which the most applicable approach is applied on equation (3.13) in Subsection 4.3.2.

4.3.1 Literature survey

Residual life is defined as the time from some current point in time, x, until the following event. This concept is not unique to reliability modeling. In reliability modeling the event corresponds to failure or suspension, in queuing theory it could correspond to the time from a customer arrives until he/she is served (see Gross and Harris (1985)) or in inventory management theory it could be the time to the reorder point (see Tijms (1976)).

The vast majority of literature found on RLE[‡] based on observed FOMs, deals with the situation where covariates are not observed or recorded. A possible reason for this is the complexity of estimating the conditional survival function of a system where the system is a function of time-dependent covariates. Percy, Kobbacy, and Ascher (1998) confirm this statement by describing the procedure as "tricky" and mention it as a possible subject for

[‡]RLE in a reliability modeling context is considered in the remainder of this section.

University of Pretoria etd – Vlok, P-J (2006)

future research.

Many authors studied the relationship between the FOM and RLE, for example, Ghai and Mi (1999), Ruiz and Navarro (1994), Guess and Prochan (1988) and Ebrahimi (1996). The univariate residual life, μ, of a system at time x is defined as a conditional expectation, i.e.

$$\mu(x) \equiv E[X - x|X \geqslant x] = \frac{\int_x^\infty R_X(x)\mathrm{d}x}{R_X(x)} \tag{4.16}$$

$R_X(x)$ is related to $h_X(x)$ by,

$$R_X(x) = \mathrm{e}^{-\int_0^x h_X(s)\mathrm{d}s} \tag{4.17}$$

Tang, Lu, and Chew (1999) describe the discrete relationship between FOM and residual life and Baganha, Geraldo, and Pyke (1999) propose a simple algorithm with which the conditional expectation can be solved for discrete relationships.

Rausand and Reinertsen (1996) believe that the probability distribution selected for RLE should primarily be based on knowledge of the underlying failure mechanism in the system. Event data should only be used to quantify parameters. Lee, Chung, Kim, Ford, and Andersen (1999) follow this school of thinking with some examples from the nuclear power generation industry. Huang, Miller, and Okogbaa (1995) describe these approaches as proof that data in industry is too limited to estimate residual life.

Guess and Park (1988), Guess and Park (1991) and Mi (1995) address the possibility that the residual life of a system is not constant. This is not done by allowing for the effects of covariates but by assuming a bathtub-curved FOM and then base RLEs on this assumption. It is also proved in these publications that by minimizing the FOM, the residual life is not necessarily maximized. Lim and Park (1995) considered a similar scenario. They propose a procedure for testing constant residual life against increasing or decreasing residual life, assuming that the proportion of the population that fails at or before the change point of the residual life function, is known.

In Pulkkinen (1991), the residual life is calculated as the time until the degree of wear of a system reaches a certain threshold level. The estimate is formed by updating the distribution of a stochastic process by describing the degree of wear. The updating procedure is based on successive application of Bayes' formula. Pulkkinen concludes that even though analytical calculation is difficult, the approach is promising. Shimizu, Ando, Morioka, and Okuzumi (1991) used a similar approach but based estimates on a threshold reliability level.

Karpinski (1988) developed a general method to determine the distribution of residual life (conditional expectation) of a system after some partial failures. The RLE starts after the first partial failure and the approach is based on the knowledge of a special distribution of component lives and system life. This method was applied with success on systems operating in nuclear power plants.

University of Pretoria etd – Vlok, P-J (2006)

Nair and Nair (1989) and Kulkarni and Rattihalli (1996) extended the common univariate residual life concept to the bivariate case where two random variables are observed at each event. This extension could be useful in cases where survival times of different parts on the same system are dependent because of common environmental influences. Zahedi (1985) and Arnold and Zahedi (1988) generalized the bivariate approached further by introducing multivariate residual life estimation. No practical examples of the bivariate or multivariate approach were found in the literature.

An approach fundamentally different to the conventional conditional expectation was proposed by Zahedi (1991). Zahedi constructed a proportional mean remaining life model, analogous to the PHM proposed by Cox (1972), i.e.

$$\mu(x, \boldsymbol{z}) = \mu_0(x) \cdot \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}) \tag{4.18}$$

where $\mu_0(x)$ is a baseline residual life function dependent on time which is influenced by a functional term containing covariates. Regression parameters are determined in a similar manner as with the model of Cox. Zahedi mentioned in this publication that practical tests were done on the model and that results would be published shortly. No further publication on this approach could be located.

4.3.2 Application of residual life theory on the combined model for non-repairable systems

In the previous section it became evident that a conditional expectation approach is most suitable for this application. This approach is a natural extension of the FOM (which is the conditional probability of failure) and has been proven in many reliability applications.

Following the assumptions made in Section 4.2.6, it is required to substitute equation (3.13) into equation (4.16). This yields,

$$\mu(x, \boldsymbol{\theta}) = \frac{\int_x^\infty R(\phi, \boldsymbol{\theta}) \mathrm{d}\phi}{R(x, \boldsymbol{\theta})} \tag{4.19}$$

where,

$$R(x, \boldsymbol{\theta}) = \exp\left[-\int_0^x \sum_{l=1}^{n} \zeta_s^{k_l} \left(\frac{\beta_s^{k_l}}{\eta_s^{k_l}} \left(\frac{\psi_s^{k_l}\left(\phi - \tau_s^{k_l}\right)}{\eta_s^{k_l}} \right)^{\beta_s^{k_l}-1} \cdot e^{\sum_{j=1}^{m_l} \gamma_{s_j}^{k_l} \cdot z_{i_j}^{k_l}} + e^{\sum_{j=1}^{m_l} \alpha_{s_j}^{k_l} \cdot z_{i_j}^{k_l}} \right) \mathrm{d}\phi \right] \tag{4.20}$$

In (4.19) and (4.20), x corresponds to the time of the last observation in the history of a system currently in operation. To be able to integrate these functions to infinity, the

University of Pretoria etd – Vlok, P-J (2006)

covariate behaviour should be extrapolated with suitable techniques up to the point where $R(x, \boldsymbol{\theta}) = 0$.

A 95% confidence bound can be constructed around the mean residual life of (4.20) to quantify the certainty of an estimate. First it is assumed that the density corresponding to the FOM in (3.13) is given by,

$$f(x, \boldsymbol{\theta}) = \mathrm{D}_x \left[1 - \exp\left(-\int_0^x h(x, \boldsymbol{\theta})\mathrm{d}x\right)\right] \tag{4.21}$$

The lower limit for the residual life is $\underset{\sim}{\mu}(x, \boldsymbol{\theta}) = \underset{\sim}{x} - x$, where $\underset{\sim}{x}$ is calculated from,

$$\int_x^{\underset{\sim}{x}} \frac{f(x, \boldsymbol{\theta})}{\int_x^\infty f(x, \boldsymbol{\theta})\mathrm{d}x} \mathrm{d}x = 0.025 \tag{4.22}$$

Similarly, the upper limit $\tilde{\mu}(x, \boldsymbol{\theta}) = \tilde{x} - x$ is found where,

$$\int_x^{\tilde{x}} \frac{f(x, \boldsymbol{\theta})}{\int_x^\infty f(x, \boldsymbol{\theta})\mathrm{d}x} \mathrm{d}x = 0.975 \tag{4.23}$$

Both equations (4.22) and (4.23) are solved numerically.

4.4 Residual life estimation based on an observed peril rate

Section 4.4 discusses RLE of a system based on an observed peril rate. Relevant literature is presented in Subsection 4.4.1 after which the most applicable approach is applied on equation (3.30) in Subsection 4.4.2.

4.4.1 Literature survey

Publications on RLE of repairable systems are not nearly as common as for non-repairable systems. Reinertsen (1996) supports this statement. All the publications found on this particular subject follow a conditional expectation approach, similar to what was described in Section 4.3.1.

Bhattacharjee (1994) investigated RLE for repairable systems and concluded that the time to first failure cannot adequately reflect its degradation over time because the aging property is influenced by maintenance. He developed a framework that attempts to formulate appropriate ratios of aging under repair and the corresponding implications. This formulation is based on conditional expectation of the next time to failure.

Calabria, Guida, and Pulcini (1990) propose a point estimation procedure for the m^{th} future failure of a repairable system based on the observation of n preceding failures. The procedure

University of Pretoria etd – Vlok, P-J (2006)

is based on the conditional expectation of the next failure calculated by maximizing the likelihood. Monte Carlo simulations done to evaluate the approach produced good results.

4.4.2 Application of residual life theory on combined model for repairable systems

Following the approaches in the literature study, a conditional expectation approach will be used to estimate residual life of a system based on an observed peril rate. According to Banjevic (2001) this is in general complicated but for the NHPP it is simple because of the definition of the NHPP (see Section A.3.3).

Meeker and Escobar (1998) describe RLE of repairable systems modeled by NHPPs in detail. It is required to calculate the mean of the distribution of failure times of a repairable system that experienced its most recent failure at time T_i and is currently operating at time t where $t > T_i$. The residual life μ for a repairable system is $\mu = T_{i+1} - t$ and for a NHPP with covariates it is expected to be,

$$\mu(t,\boldsymbol{\theta}) = \frac{\int_t^{\infty} \vartheta \cdot \mathrm{D}_{\vartheta}\left[1 - \exp\left(-\int_t^{\infty}(\rho(s,\boldsymbol{\theta}) - \rho(t,\boldsymbol{\theta}))\,\mathrm{d}s\right)\right]\mathrm{d}\vartheta}{1 - \exp\left(-\int_t^{\infty}(\rho(s,\boldsymbol{\theta}) - \rho(t,\boldsymbol{\theta}))\,\mathrm{d}s\right)} - t \tag{4.24}$$

A 95% confidence bound can be constructed around the mean residual life of (4.24) to quantify the certainty of an estimate. First it is assumed that the density corresponding to the relevant portion of the peril rate is given by,

$$f(t,\boldsymbol{\theta}) = \mathrm{D}_t\left[1 - \exp\left(-\int_{T_i}^{t}(\rho(s,\boldsymbol{\theta}) - \rho(t,\boldsymbol{\theta}))\,\mathrm{d}s\right)\right] \tag{4.25}$$

The lower limit for the residual life is $\underset{\sim}{\mu}(t,\boldsymbol{\theta}) = \underset{\sim}{t} - t$, where $\underset{\sim}{t}$ is calculated from,

$$\int_t^{\underset{\sim}{t}} \frac{f(t,\boldsymbol{\theta})}{\int_t^{\infty} f(t,\boldsymbol{\theta})\mathrm{d}t}\mathrm{d}t = 0.025 \tag{4.26}$$

Similarly, the upper limit $\tilde{\mu}(t,\boldsymbol{\theta}) = \tilde{t} - t$ is found where,

$$\int_t^{\tilde{t}} \frac{f(t,\boldsymbol{\theta})}{\int_t^{\infty} f(t,\boldsymbol{\theta})\mathrm{d}t}\mathrm{d}t = 0.975 \tag{4.27}$$

Both equations (4.26) and (4.27) are solved numerically.

4.5 Presentation of results to maintenance practitioners

Up to this point, objectives (i) and (ii) of the problem statement (see Section 1.6) were addressed. But, if these results are not presented in a user-friendly and comprehensible

University of Pretoria etd – Vlok, P-J (2006)

manner, the contribution of the thesis to practical reliability modeling will be small. It is necessary to "sell" the concept of RLE to maintenance practitioners on two levels. The first level of people is the middle and upper level maintenance managers. This would typically include maintenance supervisors, general maintenance managers and engineering managers. The second level of people will be referred to as end-users of the RLE methodology and could include maintenance planners and highly skilled maintenance technicians responsible for a limited scope of equipment.

Different approaches should be followed to promote the RLE methodology amongst the two identified levels of maintenance practitioners because each level will evaluate the methodology differently. Middle and upper level maintenance practitioners will be interested in the process-flow of the concept and how the concept will be integrated with conventional practices and processes.

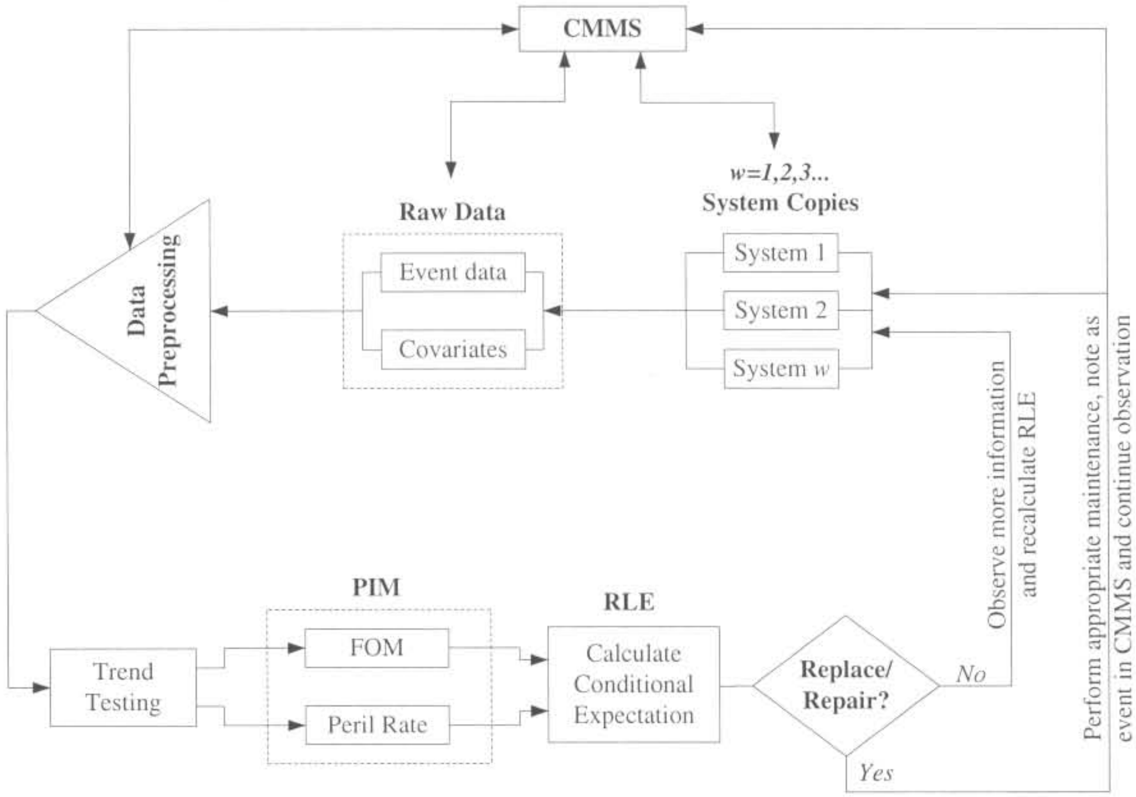

Figure 4.1: Diagrammatic overview of the process-flow of the RLE concept

Figure 4.1 shows a diagrammatic overview of the process-flow of the RLE methodology that can be used as a basis to introduce the concept to the first level of maintenance practitioners. The involvement of an organization's Computerized Maintenance Management System (CMMS) is emphasized in Figure 4.1 because data required for RLE would typically be re-

University of Pretoria etd – Vlok, P-J (2006)

trieved and stored in the CMMS. Note that even though the illustration in Figure 4.1 is very conceptual, the difference between the non-repairable and repairable cases is stressed. The process is sketched as a loop to show that as soon as new information is available (a new CM inspection was done) or a maintenance action was taken, the process is repeated.

End-users of the RLE concept will not be interested in the process-flow of the methodology but rather in the outputs of the concept. For this reason, the methodology should be introduced and results should be presented in a practical (graphical) manner. End-users should understand that RLE algorithms developed in this thesis are only decision support tools and that the final maintenance decision is still up to the individual. Guidelines for decision making are covered in Section 4.6.

Two different illustrations of the output of the RLE methodology should be presented to end-users to promote the concept: one for a non-repairable system and another for a repairable system. Figure 4.2 shows the output of the RLE methodology on a non-repairable system.

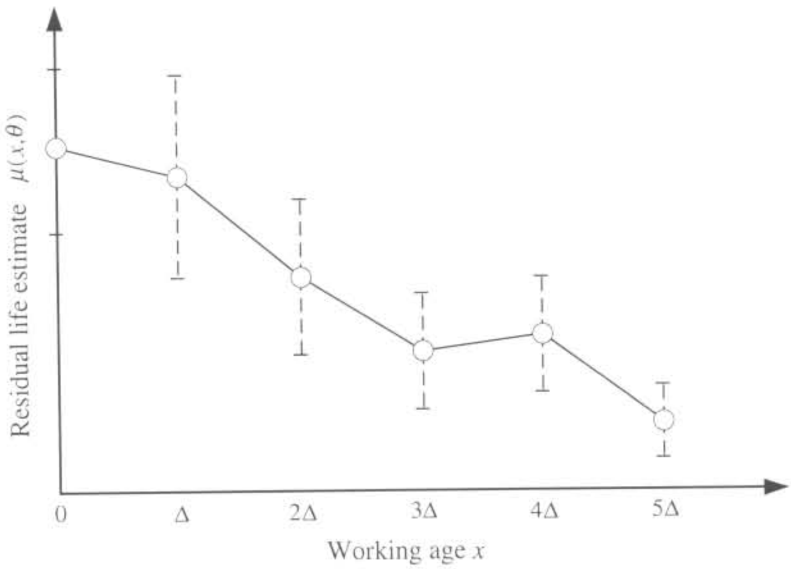

Figure 4.2: Presentation of RLE results of non-repairable systems to end-users

Suppose a fixed CM inspection interval of Δ is used to monitor a particular non-repairable system. At time $x = 0$ when the system was installed, an estimate of the residual life is made within statistical bounds (indicated as a dotted vertical lines). After Δ time units new information (covariates) are observed and the residual life is recalculated. If the system's degradation is linear and the model is accurate, the residual life estimate should reduce by Δ. It is expected that the $\mu(x, \boldsymbol{\theta})$ would decrease monotonically with wear but minor maintenance to the system could elongate the system's life and $\mu(x, \boldsymbol{\theta})$ could increase (see $x = 4\Delta$ on Figure 4.2).

University of Pretoria etd – Vlok, P-J (2006)

The output of the RLE methodology on a repairable system is virtually the same as for the non-repairable system except that the time scale is different. See Figure 4.3.

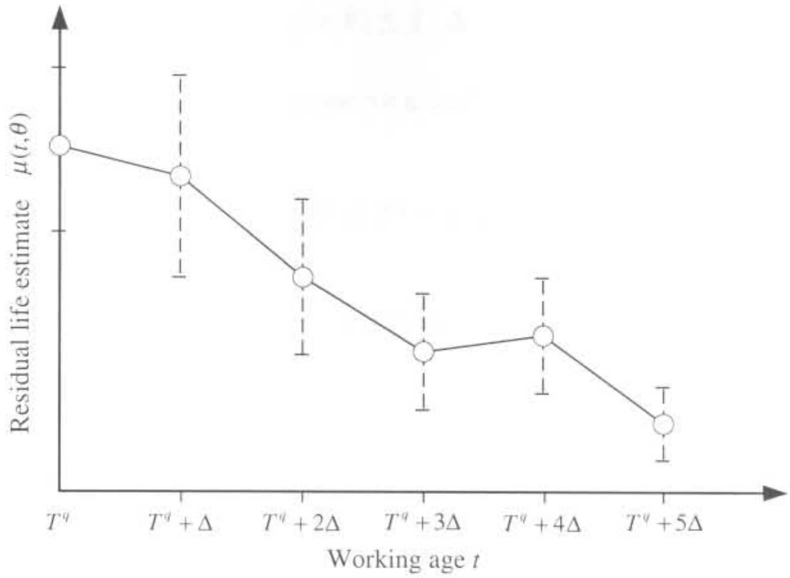

Figure 4.3: Presentation of RLE results of repairable systems to end-users

Suppose a fixed CM inspection interval of Δ is used to monitor a particular repairable system. At time T^q when the system was re-installed, an estimate of the residual life is made within statistical bounds (indicated as dotted vertical lines). After Δ time units new information (covariates) are observed and the residual life is recalculated. If the system's degradation is linear and the model is accurate, the residual life estimate should reduce by Δ. It is expected that the $\mu(t,\boldsymbol{\theta})$ would decrease monotonically with wear but minor maintenance to the system could elongate the system's life and $\mu(t,\boldsymbol{\theta})$ could increase (see $t = T^q + 4\Delta$ on Figure 4.3).

4.6 Decision making with the assistance of dynamic residual life estimates

The ideal decision rule for the RLE approach would be to take action as soon as the lower limit of the residual life estimate equals zero. This is not practical however because of two reasons. Firstly, because the lower limit of the residual life estimate is calculated from the conditional expectation of failure it implies that this estimate will only approach zero when time (local or global) approaches infinity. Secondly, inspections are most often done discretely which means that the lower limit of the residual life estimate could become zero in-between inspections and the failure could occur before the next inspection. For these reasons a simple but practical decision rule is proposed: take preventive maintenance action as soon as the lower residual

University of Pretoria etd – Vlok, P-J (2006)

life estimate is less than the time to the next inspection. Thus, for the non-repairable case it is when,

$$\underset{\sim}{\mu}(x,\boldsymbol{\theta}) \leq j \cdot \Delta - x \tag{4.28}$$

where the j^{th} inspection is the next inspection and Δ is the inspection interval. Similarly, for the repairable case it is when,

$$\underset{\sim}{\mu}(t,\boldsymbol{\theta}) \leq T^q + j \cdot \Delta - t \tag{4.29}$$

It is important to realize that these decision rules are influenced by the following:

(i) Quality of data
(ii) Quantity of data
(iii) Selection of the PIM
(iv) Selection of covariates
(v) Accuracy of covariate behavior prediction
(vi) Accuracy of covariate measurements

It is difficult to quantify the effect of these influences on residual life estimates and it would differ for each situation. The proposed decision rule should thus only be used as a maintenance decision support tool and the final decision should still be made by the maintenance practitioner.

4.7 Conclusion

In this chapter, the RLE process was divided into three steps: (i) prediction of future covariate behaviour; (ii) calculating residual life based on observed FOMs or peril rates with the assistance of covariate behaviour predictions; and (iii) the presentation of results in a comprehensible manner. These steps are similar to the 2-phase approach of Christer and Wang (1995) that constructed a model for the prediction of covariate behaviour first and then used this prediction as a covariate in a PIM to make maintenance decisions. Christer and Wang restricted their approach to the renewal case modeled by a simple Weibull FOM and used the results to minimize long term cost, downtime or risk (similar to Makis and Jardine (1991)).

Covariates in this thesis are assumed to be time-dependent, internal and non-stochastic, provided that the covariates have been observed up to a certain point x or t. Several possible parametric functions with their solutions that could be used for covariate behaviour prediction are summarized in Table 4.1.

University of Pretoria etd – Vlok, P-J (2006)

A detailed literature study was done on RLE based on observed FOMs or peril rates and it was found that very few attempts have been made to calculate RLE based on an observed peril rate. Many publications were found on RLE based on FOMs. In both cases, the vast majority of authors used a conditional expectation approach to calculate residual life. Following this, it was decided to use the conditional expectation approach to calculate $\mu(x, \boldsymbol{\theta})$ and $\mu(t, \boldsymbol{\theta})$. The conditional expectation approach applied to the theory of Chapter 3 is summarized in Table 4.4.

Table 4.4: Summary of RLE calculations based on a FOM or peril rate

Non-repairable Case[§]
$\mu(x, \boldsymbol{\theta}) = \frac{\int_x^\infty R(\phi, \boldsymbol{\theta}) \mathrm{d}\phi}{R(x, \boldsymbol{\theta})}$
Repairable Case[¶]
$\mu(t, \boldsymbol{\theta}) = t - \frac{\int_t^\infty \vartheta \cdot \mathrm{D}_\vartheta \left[1 - \exp\left(-\int_t^\infty \left(\rho(s, \boldsymbol{\theta}) - \rho(t, \boldsymbol{\theta})\right) \mathrm{d}s\right)\right] \mathrm{d}\vartheta}{1 - \exp\left(-\int_t^\infty \left(\rho(s, \boldsymbol{\theta}) - \rho(t, \boldsymbol{\theta})\right) \mathrm{d}s\right)}$

The RLE process used in this thesis is of a fairly complex nature and would probably not find acceptance amongst maintenance practitioners in its mathematical form. For this reason, simplified graphical representations of the approach were developed for middle to upper level management and for end-users in Section 4.5.

In this chapter, Steps (ii) and (iii) of the objectives set out in Section 1.6 were achieved. It is hence required to apply the theory and methodologies developed on an actual data set. This is done in Chapter 5.

[§]Variables for the models corresponding to the non-repairable case are declared and described in Section 4.3

[¶]Variables for the models corresponding to the repairable case are declared and described in Section 4.4

University of Pretoria etd – Vlok, P-J (2006)

Chapter 5

Case study

5.1 Introduction

It was stated earlier in this thesis that a useful contribution to the field of reliability modeling can only be claimed if the developed theory is implemented successfully on an actual data set. Data can be obtained in two ways: (i) it can be generated in a laboratory under controlled conditions; or (ii) it can be collected in a typical industrial situation. Successful application of the theory on data obtained from a laboratory may be doubtful because of the controlled conditions in laboratories that do no exist in practice. It was hence decided to use data from a real industrial situation.

Data satisfying the requirements of proportional intensity modeling was found at SASOL Coal's Twistdraai plant at Secunda*. The Twistdraai plant was started up in September 1996 as a coal washing plant that washes coal to a certain cleanliness before it enters the petrochemical process. In the plant, eight Warman® axial in, radial out pumps are used to circulate a water and magnetite solution which is used in the washing process. A condition monitoring maintenance strategy through vibration monitoring was applied on the pumps from the startup date. Despite this strategy, to date several failures have occurred on these pumps. The events produced by the pumps together with the vibration information are used in the PIM theory of Chapter 3.

The theory that was developed in this thesis not only needs to be applied to real data but also to be benchmarked against existing approaches. The only existing approach that the present research can be compared against is the decision technique of Makis and Jardine mentioned in Section 1.4.2. This approach was applied to the above-mentioned data set by Vlok (1999) and the results are briefly repeated in this chapter and compared to the RLE approach.

*SASOL is a major petrochemical company in South Africa

University of Pretoria etd – Vlok, P-J (2006)

Chapter 5 starts off with a description of the SASOL data set after which the approach of Makis and Jardine is applied and discussed. Because this approach was only briefly introduced in Chapter 1, more details on the theory are also presented for completeness. The second part of this chapter consists of the application of the RLE theory developed in this thesis on the same data set. Chapter 5 ends with a detailed comparison of the two approaches.

5.2 Description of SASOL data

The data set under discussion has many shortcomings, including missing observations and irregular inspection intervals, but was the best data set found after an extensive search for suitable data in the South African industry. The Twistdraai plant was started up in September 1996 and is thus still relatively new. Data was collected from September 1, 1996 to November 1, 1998 which gives an analysis time horizon of 791 days. A second data set was collected from November 1, 1998 to February 28, 1999 to further evaluate the combined PIM's performance.

5.2.1 Background

A total of eight identical axial in, radial out, Warman pumps are used in a specific section of the plant to circulate a water and magnetite solution. These pumps are very important in the washing process and significant production losses are suffered when one of the pumps breaks down. All eight pumps work under nominally similar conditions. Figure 5.1 shows the configuration of the eight pumps while Figure 5.2 shows a close-up of one of the pumps.

All the elements visible in Figure 5.2 are implied when referring to a *pump* except for the 220 kW electrical motors used to drive these pumps. A pump consists principally of an impeller housing, impeller, bearing housing, 2 × SKF 938 932 bearings, drive shaft, V-belt pulley and seals.

Because of the aggressive nature of the fluid being circulated and the robust environment of the pumps, destructive failures are encountered frequently. These destructive failures often occur abruptly, i.e. a pump's state literally change overnight from being in an acceptable condition to being completely destructed. Functional failures are usually caused by one (or a combination) of the following:

(i) Complete bearing seizure
(ii) Broken or defective impeller
(iii) Damaged or severely eroded pump housing

University of Pretoria etd – Vlok, P-J (2006)

Figure 5.1: Pump layout

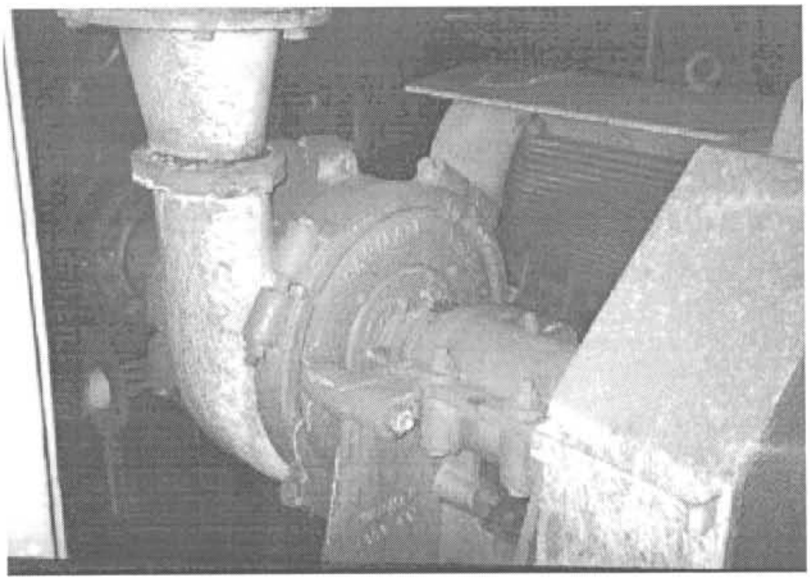

Figure 5.2: Warman Pump

(iv) Broken drive shaft

When a pump has failed due to one of the reasons above, it is overhauled completely regardless of the amount of work that needs to be done. This may include replacement of bearings, repair or renewal of impeller, repair or renewal of impeller housing or replacement of the main shaft. No complete spare pumps are stocked at the plant but only spare parts, since some parts tend to fail more often than others.

University of Pretoria etd – Vlok, P-J (2006)

During the analysis time horizon, the plant's management prescribed a condition based preventive maintenance strategy based on vibration monitoring results. No fixed inspection interval was used and vibration levels were only measured sporadically or when a notable deterioration in a pump's condition became evident, whereafter more regular inspections were done. This strategy lead to several unexpected failures.

Vibration levels of the pumps were measured on the shaft bearings in two directions, horizontally and vertically, to assess a pump's condition. Figure 5.3 shows the horizontal measuring positions.

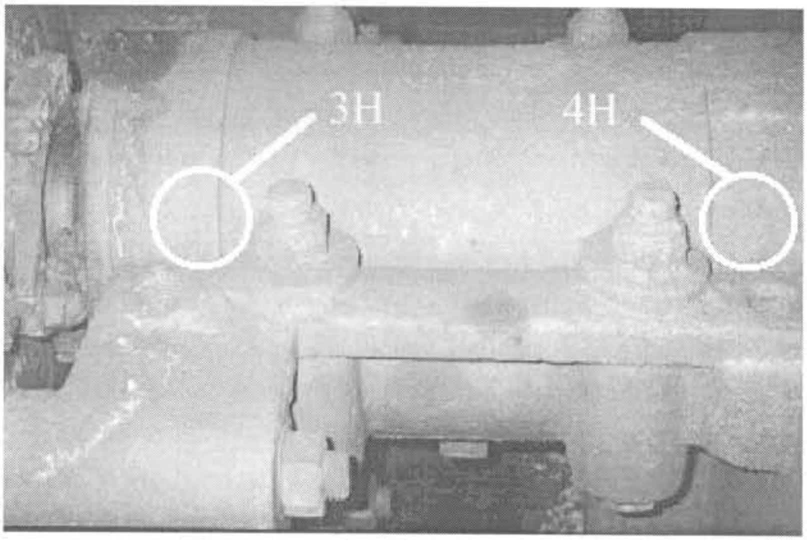

Figure 5.3: Monitoring spots on pumps

The "wet-end" bearing (the bearing closest to the impeller) is referred to as Bearing 3 while the "dry-end" bearing is labeled Bearing 4. Measuring positions 3H and 4H are thus the horizontal measurements on Bearing 3 and 4, respectively. Only the horizontal measurements were used in the analyses - reasons are presented later.

As in most typical vibration monitoring programs, the maintenance decisions on the pumps were based on spectral vibration analysis. Several important frequencies are enveloped with alarm levels and the required maintenance is performed as soon as two or three of the alarm levels are exceeded. Alarm levels were determined by a combination of technician experience and OEM specifications.

Vibration data loggers were used to capture vibration data on the pumps, from where the information was downloaded onto a dedicated computerized vibration measurement database. Data used in this research was retrieved from this database. Frequency spectrums of all

University of Pretoria etd – Vlok, P-J (2006)

measurements are stored in the database and the chosen covariate levels (discussed later) could be retrieved accurately.

The vibration measurement database does not contain information regarding events during a pump's life, nor does the plant's CMMS. This is not considered to be a serious shortcoming for this research since the only event or action performed on a pump during its life time is additional lubrication. It is assumed that additional lubrication does not effect covariate levels severely.

Root cause failure analysis records obtained from the CMMS provided insight on the state of a pump when maintenance was performed, i.e. whether it was in the failed state or was preventively withdrawn from service.

5.2.2 Covariates

Covariate selection was largely based on the experience of vibration technicians involved with the pumps at the plant. These technicians are of the opinion that the horizontal vibration measurements on the bearings alone is a sufficient indication of a pump's condition and that not much additional information is obtained from the vertical measurements. According to the theory of vibration analysis this viewpoint is not necessarily correct, but it was nevertheless decided to use only the horizontal vibration measurements to show that the combined PIMs can improve decision making even if covariates have certain shortcomings.

As mentioned earlier, the vibration monitoring program that was used on the pumps was based on spectral analysis. A number of important frequency bands (selected on theory and experience) are monitored and a pump is maintained as soon as two or three of these frequency bands' amplitudes exceed certain alarm levels. It was decided to use all of these frequency bands as covariates in the combined PIMs, thereby incorporating vibration theory and technician experience in the models. Table 5.1 summarizes the 12 selected covariates (6 on each bearing).

Table 5.1: Summary of covariates

	Covariate	Description
1.	RF043H	0.4× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of a bearing defect.
2.	RF13H	1× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of unbalance in the pump.
3.	RF23H	2× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of misalignment in the pump.

University of Pretoria etd – Vlok, P-J (2006)

4.	RF53H	5× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of cavitation in the pump.
5.	HFD3H	High frequency domain components between 1200-2400 Hz, measured on Bearing 3, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.
6.	LNF3H	Lifted noise floor in 600-1200 Hz range, measured on Bearing 3, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.
7.	RF044H	0.4× rotational frequency amplitude, measured on horizontally on Bearing 4 in mm/s, indicative of a bearing defect.
8.	RF14H	1× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of unbalance in the pump.
9.	RF24H	2× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of misalignment in the pump.
10.	RF54H	5× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of cavitation in the pump.
11.	HFD4H	High frequency domain components between 1200-2400 Hz, measured on Bearing 4, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.
12.	LNF4H	Lifted noise floor in 600-1200 Hz range, measured on Bearing 4, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.

The biggest challenge when defining vibration covariates is to select a single quantity that describes a specific defect most clearly. A specific defect can often be identified by numerous parameters but not all parameters can be used as covariates, since the number of covariates has to be limited. Too many covariates often cause the PIMs to become mathematically unstable or difficult to estimate, especially for small sample sizes.

5.2.3 Description of collected data

The data collected include the pump unit identification, dates of inspection, vibration frequency spectrum at each inspection (covariates), date of failure or suspension and the state at maintenance, i.e. failed or suspended. Accurate inspection data was generally not available for cases where unexpected failures occurred and data was generated by extrapolating available data as appropriately as possible to the date of unexpected failure.

University of Pretoria etd – Vlok, P-J (2006)

A total of 27 lifetimes with condition monitoring information (called *histories*) were compiled over the analysis horizon with 98 inspections (extrapolations included). This gives an average of 3.6 inspections per history. Approximately 50% of all inspections were done on an irregular basis either at the beginning or the end of a pump's life time.

Of the 27 histories, 11 were failures, 8 were suspensions and 8 were calendar suspensions since all 8 pumps were running at the cutoff date of the analysis horizon. The 11 failures were all unexpected and production losses were suffered following these events. The 8 suspensions were all done based on vibration measurements and were considerably cheaper than the unexpected failures. Three of the 8 suspensions were done on very short life times relative to other survival times.

The working age of the pumps was considered to be the same as the calendar age, because the pumps run 24 hours per day, 365 days per year. The pumps are very rarely shut down because of breakdowns on other parts of the plant and these times are considered to be insignificantly small.

Three events were defined for the pumps through their life times: (1) B - Beginning or pump startup; (2) S - Suspension; and (3) F - Failure. Events that occurred to the pumps are listed in Table 5.2 below†.

Table 5.2: Summary of events

Pump ID	Age (days)	Date	Event
PC1131	0	9/1/1996	B
PC1131	397	10/3/1997	S
PC1131	397	10/3/1997	B
PC1131	554	3/9/1998	F
PC1131	554	3/9/1998	B
PC1131	690	7/23/1998	S
PC1131	690	7/23/1998	B
PC1131	765	10/6/1998	F
PC1131	765	10/6/1998	B
PC1131	791	11/1/1998	S
PC1132	0	9/1/1996	B
PC1132	491	1/5/1998	F
PC1132	491	1/5/1998	B
PC1132	544	2/27/1998	S
PC1132	544	2/27/1998	B
PC1132	557	3/12/1998	S

†A graphical illustration of the event data is presented in Figure 5.8.

University of Pretoria etd – Vlok, P-J (2006)

PC1132	557	3/12/1998	B
PC1132	751	9/22/1998	F
PC1132	751	9/22/1998	B
PC1132	791	11/1/1998	S
PC1231	0	9/1/1996	B
PC1231	563	3/18/1998	F
PC1231	563	3/18/1998	B
PC1231	578	4/2/1998	S
PC1231	578	4/2/1998	B
PC1231	791	11/1/1998	S
PC1232	0	9/1/1996	B
PC1232	599	4/23/1998	S
PC1232	599	4/23/1998	B
PC1232	791	11/1/1998	S
PC2131	0	9/1/1996	B
PC2131	184	3/4/1997	F
PC2131	184	3/4/1997	B
PC2131	470	12/15/1997	S
PC2131	470	12/15/1997	B
PC2131	631	5/25/1998	F
PC2131	631	5/25/1998	B
PC2131	774	10/15/1998	F
PC2131	774	10/15/1998	B
PC2131	791	11/1/1998	S
PC3131	0	9/1/1996	B
PC3131	450	11/25/1997	F
PC3131	450	11/25/1997	B
PC3131	791	11/1/1998	S
PC3132	0	9/1/1996	B
PC3132	506	1/20/1998	F
PC3132	506	1/20/1998	B
PC3132	791	11/1/1998	S
PC3232	0	9/1/1996	B
PC3232	563	3/18/1998	F
PC3232	563	3/18/1998	B
PC3232	723	8/25/1998	S
PC3232	723	8/25/1998	B
PC3232	791	11/1/1998	S

University of Pretoria etd – Vlok, P-J (2006)

Detailed inspection data of all the covariate measurements between events is presented in Appendix D. Covariate values immediately after the occurrence of an event were all assumed to be zero. Further detailed comments on the inspection data are presented below:

(i) Covariate RF043H recorded two unusually high values of 250 and 1200 mm/s compared to the normal range of between 0 and 5.6 mm/s. These high values were confirmed by the vibration monitoring database and vibration technicians are confident that these levels were not due to faulty monitoring equipment or human error. A further notable fact is that these values occurred at suspensions.

The most logical explanation for these values lies in the wear mechanism present in the bearings. RF043H is indicative of a particular bearing defect and the bearings that produced these extreme values were probably able to withstand the wear associated with RF043H, i.e. did not abrade with the introduction of the RF043H vibration which would have retrained the vibration levels to within normal limits. The vibration levels continued to rise to the unusually high values, which persuaded management to maintain the pumps preventively.

(ii) Subjective covariates HFD3H, HFD4H, LNF3H and LNF4H indicated the presence of their associated phenomena with a simple "0" or "1". These phenomena appear in different degrees of severity and it is possible to argue that covariates that quantify the severity would lead to more accurate PIMs. It is however very difficult to quantify the severity of these phenomena with a single number (covariate) because it ranges over large frequency bands. In practice, vibration technicians do not attempt to estimate the severity of these phenomena either but only use the presence (or absence) thereof as a supportive argument in decisions. It was hence decided that a simple "0" or "1" would suffice for this study.

It is expected that whenever one of the subjective covariates turns to 1, it will remain 1. This is however not observed in the data, once again due to wear mechanisms present in the pumps. For example, LNF3H or LNF4H is present in certain inspections but absent in following inspections, only to return in subsequent inspections. LNF is indicative of a lack of lubrication. When there is a lack of lubrication asperities induce a lifted noise floor over 600-1200 Hz but the asperities are soon worn off thereby inducing increased levels of unbalance but a reduction in the lifted noise floor. Hence, the LNF covariate appears, diminishes and reappears.

(iii) Failure times are distributed such that 6 failures occurred below 200 days and the remaining 5 failures above 450 days. Suspension times are apparently randomly distributed with some being very short such as 53, 15 and 13 days.

(iv) Covariate RF13H shows comparatively high values in the beginning of histories and then decreases gradually towards events. RF14H has a very similar pattern, although not as distinct. Technical reasons for this would be the same as discussed in (i).

University of Pretoria etd – Vlok, P-J (2006)

Costs associated with failures and suspensions of the pumps could not be disclosed exactly by the Twistdraai plant because of company policy. The Twistdraai plant did provide scaled costs however which is proportional to the true costs. An unexpected failure cost of C_f = R 162 200 will be used and a preventive maintenance cost of C_p = R 25 000. Costs related to production losses suffered due to unavailability are included in C_f and C_p. These costs were average costs sustained by the Twistdraai plant for the two years over which the data was collected.

5.3 Maintenance Strategy Optimization through Proportional Hazards Modeling with Cost Optimization

The decision-model by Makis and Jardine uses the Weibull PHM as PIM to optimize the maintenance strategy. In this section, the selection of covariates and the Weibull PHM fit are described before the decision-model is applied to the Weibull PHM‡. The description mainly focusses on the results since the details are presented in Vlok (1999).

5.3.1 Weibull PHM fit

There is no straightforward procedure to select the most appropriate covariates to obtain an acceptable Weibull PHM. For this data set a combination of backward selection (eliminating covariates with the highest p-values, one at a time), residual graphs, goodness-of-fit tests and technical experience were utilized to obtain the best possible model. Some guidelines for covariate selection proposed by Hosmer and Lemeshow (1999), Sakamoto, Ishiguro, and Kitagawa (1983) and Schwartz (1978) that were also used include:

(i) It is not recommended to exclude several covariates from the model in one step. This may lead to an inaccurate model.

(ii) If two covariates are highly correlated, they can produce very uncertain estimates (large standard errors) which will make them appear as insignificant, even if one of them is a good predictor of failure.

(iii) Some covariates can appear as insignificant, contrary to a technician's opinion, simply because of insufficient data or high variations. It is not recommended to include these in a PIM, because their parameters could be very inaccurate and produce a misleading model. They could be checked again when more data is available.

‡In this section, maintenance actions are referred to as *renewals* because the PHM implicitly makes the GAN assumption.

University of Pretoria etd – Vlok, P-J (2006)

(iv) Positive covariates with negative regression coefficients should be considered with special care, because it indicates that the FOM increases with decreasing covariate values (as is the case with RF13H and RF14H), which is not usually expected. In some cases it could be because some influential events, such as minor repairs, were not recorded.

(v) Some covariates can surprisingly appear as significant, without practical explanation. This almost always indicates some data problem, especially if wrong covariate values are reported at failures.

To be able to recognize all patterns in the data, it was decided to model the data in three phases: (1) By a simple Weibull model, i.e. a Weibull FOM without covariates; (2) By a Weibull PHM where the subjective covariates are temporary excluded; and (3) By a Weibull PHM with all covariates included from the start. This exercise revealed that the second phase produces the most practical model with only two covariates, RF53H and RF54H[§]. The model is given by,

$$h(x, \boldsymbol{z}) = \frac{1.464}{1431.8} \cdot \left(\frac{x}{1431.8}\right)^{0.464} \exp\left(0.127 \cdot \text{RF53H} + 0.143 \cdot \text{RF54H}\right) \tag{5.1}$$

The results of analytical significance tests on the parameters are summarized in Table 5.3. It is clear that both RF53H and RF54H are significant in the failure process although the shape parameter, β, did not prove to be significant. A Kolmogorov-Smirnov (KS) test yielded 0.3180 with a p-value of 0.00628, which is not an excellent model fit.

Table 5.3: Summary of analytical significance tests performed on the model in equation (5.1)

	Parameters		
	β	**RF53H**	**RF54H**
Estimate	1.464	0.1271	0.1414
Standard Error	0.4719	0.0227	0.0569
Wald	0.9678	31.24	6.172
Wald p-value	0.3252	0.000	0.013

Residual analyses were also done on the model. A plot of the residuals in order of appearance is shown in Figure 5.4.

In the case of a perfect model fit, the residuals in Figure 5.4 would all be scattered around the straight line $y = 1$. Note that the residual values of suspended cases will by definition always be greater than 1 (see Schoenfield (1990)). If an upper limit of $y = 3$ (95%) and a lower limit of $y = 0.05$ (5%) are chosen, it is expected that at least 90% of the residuals will fall inside

[§]RF53H and RF54H also proved to be significant covariates in the first and third phases.

University of Pretoria etd – Vlok, P-J (2006)

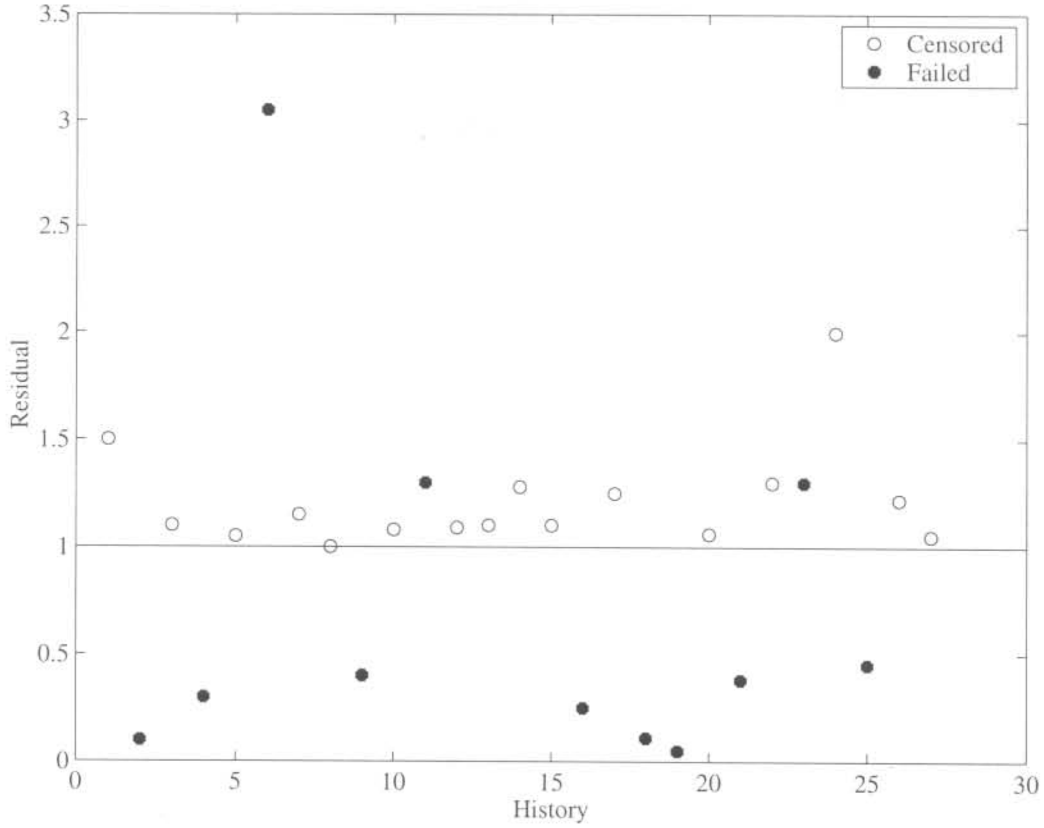

Figure 5.4: Residuals in order of appearance

these limits if the model fits the data. Figure 5.4 shows that 4 of the 6 short failures are not modeled well by the equation (5.1), i.e. the 4 observations close to $y = 3$ and $y = 0.05$. Further analyses of the data showed that no other quantitative covariate contributed significantly to these early failures. RF53H and RF54H proved to be very significant in the other, longer failures.

The model in equation (5.1) is of an average statistical quality but was finally chosen because of its practical value and its relation to the actual situation.

5.3.2 Transition probabilities

Covariates were assumed to be stochastic and transition probabilities were used to estimate future covariate behaviour (see Section 4.2.4). The covariate bands that were selected for RF53H and RF54H are presented in Table 5.4 with the frequency of observations in each band. See Table D.1 and D.2 for the actual data.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.4: Covariate bands for RF53H and RF54H

RF53H		RF54H	
Band	**Frequency**	**Band**	**Frequency**
[0→5]	67	[0→3]	54
(5→10]	15	(3→7]	28
(10→15]	11	(7→11]	11
(15→26.84]	4	(11→15]	4
(26.84→ ∞)	1	(15→ ∞)	1

With the covariate bands in Table 5.4 transition rates were calculated with (4.11) and transition matrices were constructed. For example, the transition probabilities for covariate RF53H for an observation interval of 50 days are given in Table 5.5.

Table 5.5: Transition probability matrix for RF53H for an observation interval of 50 days

BANDS	[0→5]	(5→10]	(10→15]	(15→26.84]	(26.84→ ∞)
[0→5]	0.913	0.068	0.014	0.004	0.001
(5→10]	0.208	0.481	0.173	0.088	0.050
(10→15]	0.063	0.260	0.228	0.216	0.233
(15→26.84]	0.010	0.064	0.104	0.234	0.588
(26.84→ ∞)	0	0	0	0	1

A similar TPMX was calculated for RF54H and is shown in Table 5.6.

Table 5.6: Transition probability matrix for RF54H for an observation interval of 50 days

BANDS	[0→3]	(3→7]	(7→11]	(11→15]	(15→ ∞)
[0→3]	0.893	0.090	0.014	0.0009	0.0004
(3→7]	0.239	0.547	0.184	0.017	0.011
(7→11]	0.108	0.078	0.609	0.96	0.105
(11→15]	0	0	0	0.212	0.787
(15→ ∞)	0	0	0	0	1

With the transition probabilities known, the cost optimization can be performed and it is described in the next section.

University of Pretoria etd – Vlok, P-J (2006)

5.3.3 Renewal decision policy

Makis and Jardine's decision-model was not described when introduced in Section 1.4.2. For the sake of continuity, it is done briefly in this section before the results of the application of the theory on the SASOL data set are presented. Two different maintenance possibilities are considered in the decision-model: (i) Variant 1, where preventive renewal can take place at any moment; and (ii) Variant 2, where preventive renewal can only take place at moments of inspection. Only Variant 1 will be discussed since Variant 2 is a simplification of Variant 1. A basic renewal rule is used: if the FOM is greater than a certain threshold value, preventive renewal should take place otherwise operations can continue. The objective here is thus to calculate this threshold level while taking the working age and covariates into account.

The expected average cost per unit time is a function of the threshold risk level, $\mathbb{D}$, and is given by (see Makis and Jardine (1991) and Makis and Jardine (1992)),

$$\Phi(\mathbb{D}) = \frac{C_p + KQ(\mathbb{D})}{W(\mathbb{D})} \tag{5.2}$$

where $K = C_f - C_p$. $Q(\mathbb{D})$ represents the probability that failure replacement will occur, i.e. $Q(d) = P(X_d \geqslant X)$ with $X_{\mathbb{D}}$ the preventive renewal time at threshold risk level $\mathbb{D}$ or $X_{\mathbb{D}} = \inf\{x \geqslant 0 : h(x, \boldsymbol{z}) \geqslant \mathbb{D}/K\}$. $W(d)$ is the expected time until replacement, regardless whether preventive action or failure, i.e. $W(\mathbb{D}) = E[\min\{X_{\mathbb{D}}, X\}]$. The optimal threshold risk level, $\mathbb{D}^*$, is determined with fixed point iteration to obtain,

$$\Phi(\mathbb{D}^*) = \min_{\mathbb{D}>0} \Phi(\mathbb{D}) = \mathbb{D}^* \tag{5.3}$$

provided that the FOM is non-decreasing, e.g. when $\beta \geqslant 1$, all covariates are non-decreasing and covariate parameters are positive. If covariates are non-monotonic, then fixed point iteration does not work, and $\min_{\mathbb{D}>0} \Phi(\mathbb{D})$ should be calculated by a direct search method. During the calculation of $\mathbb{D}^*$ it is necessary to calculate $Q(\mathbb{D})$ and $W(\mathbb{D})$ which is not a trivial procedure. To do this a covariate vector, $\boldsymbol{z}(x) = [z_1(x), z_2(x), ..., z_m(x)]$, is defined with corresponding vector $\boldsymbol{i}(x) = [i_1(x), i_2(x), ..., i_m(x)]$, the state of every covariate at time x. Thus, for every coordinate l, let $X^l(i_l(x))$ be the value of the l^{th} covariate in state $i_l(x)$ at instant x and $X(\boldsymbol{i}(x)) = [X^1(i_1(x)), ..., X^m(i_m(x))]$. It is hence possible to express the FOM as,

$$h(x, \boldsymbol{i}(x)) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp(\boldsymbol{\gamma} \cdot X(\boldsymbol{i}(x))) \tag{5.4}$$

From Section A.2.1, the conditional reliability function can be defined as $R(j, i, x) = P[X > j\Delta + x | X > j\Delta, \boldsymbol{i}(x)]$, which becomes after substitution,

$$R(j, \boldsymbol{i}(x), x) = \exp\left\{-\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \cdot \left[\left(\frac{j\Delta + x}{\eta}\right)^{\beta} - \left(\frac{j\Delta}{\eta}\right)^{\beta}\right]\right\} \tag{5.5}$$

University of Pretoria etd – Vlok, P-J (2006)

with $0 \leqslant x \leqslant \Delta$. If $h(x, \boldsymbol{i}(x))$ is a non-decreasing function in x, $x_i = \inf\{x \geqslant 0 : h(x, i(x)) \geqslant d/K\}$ and the k_i's are integers such that $(k_i - 1)\Delta \leqslant x_i < k_i\Delta$, the mean sojourn time of the system in each state can be calculated by,

$$S(j, \boldsymbol{i}(x)) = \begin{cases} 0, & j \geqslant k_i \\ S(j, i, a_i), & j = k_i - 1 \\ S(j, i, \Delta), & j < k_i - 1 \end{cases} \tag{5.6}$$

where $a_i = x_i - (k_i - 1)\Delta$ and $S(j, i, s) = \int_0^s R(j, i, x)\mathrm{d}x$. Similarly, the conditional cumulative distribution function for this situation is,

$$F(j, \boldsymbol{i}(x)) = \begin{cases} 0, & j \geqslant k_i \\ 1 - R(j, i, a_i), & j = k_i - 1 \\ 1 - R(j, i, \Delta), & j < k_i - 1 \end{cases} \tag{5.7}$$

Let for each j, $S_j = (S(j,i))_i$ and $F_j = (F(j,i))_i$ be column vectors and $(P_j) = (R(j, i, \Delta)P_{il}(j))_{il}$ be a matrix. The column vectors $W_j = (W(j,i))$ and $Q_j = (Q(j,i))$ can hence be calculated as follows,

$$\begin{aligned} W_j &= S_j + P_j W_{j+1} \\ Q_j &= F_j + P_j Q_{j+1} \end{aligned} \tag{5.8}$$

Following this, $W = W(0, i_0)$ and $Q = Q(0, i_0)$ where i_0 is an initial state of the covariate process, usually $i_0 = 0$. By starting the calculation with a large value for j, with $W_{j+1} = Q_{j+1} = 0$ and working back to 0, it is possible to solve for W and Q from (5.8). The above calculation procedure is described in detail in Makis and Jardine (1992). A forward version of this backward calculation is numerically more convenient and much faster according to Banjevic, Ennis, Braticevic, Makis, and Jardine (1997), which can be suitably adjusted for non-monotonic FOMs as well.

Thus, once the optimal threshold level is determined, the item is renewed at the first moment x when,

$$\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)\right) \geqslant \frac{\mathbb{D}^*}{K} \tag{5.9}$$

or when

$$\boldsymbol{\gamma} \cdot \boldsymbol{z}(x) \geqslant \delta^* - (\beta - 1)\ln x \tag{5.10}$$

where $\delta^* = \ln(\mathbb{D}^*\eta^\beta/K/\beta)$.

A warning level function is defined only in terms of time by,

$$G(x) = \delta^* - (\beta - 1) \cdot \ln x \tag{5.11}$$

University of Pretoria etd – Vlok, P-J (2006)

with $G(x)$ strictly decreasing if $\beta > 1$.

As mentioned earlier, the costs provided by the Twistdraai plant, C_f = R 162 200 and C_p = R 25 000 were based on averages over the two year data horizon. Further details about the cost estimation are not available.

No fixed inspection frequency was used at the plant which made calculations somewhat more difficult. The transition probability matrices were estimated based on transition rates (as described in Section 4.2.4) and a future inspection interval of 50 days was used for the cost model. With all preliminary calculations completed, the cost function of equation (5.2) was hence calculated using the backward recursive procedure. The result is shown graphically in Figure 5.5 in terms of the threshold risk level, $\mathbb{D}$ (or $h(x, \boldsymbol{z}) \cdot K$).

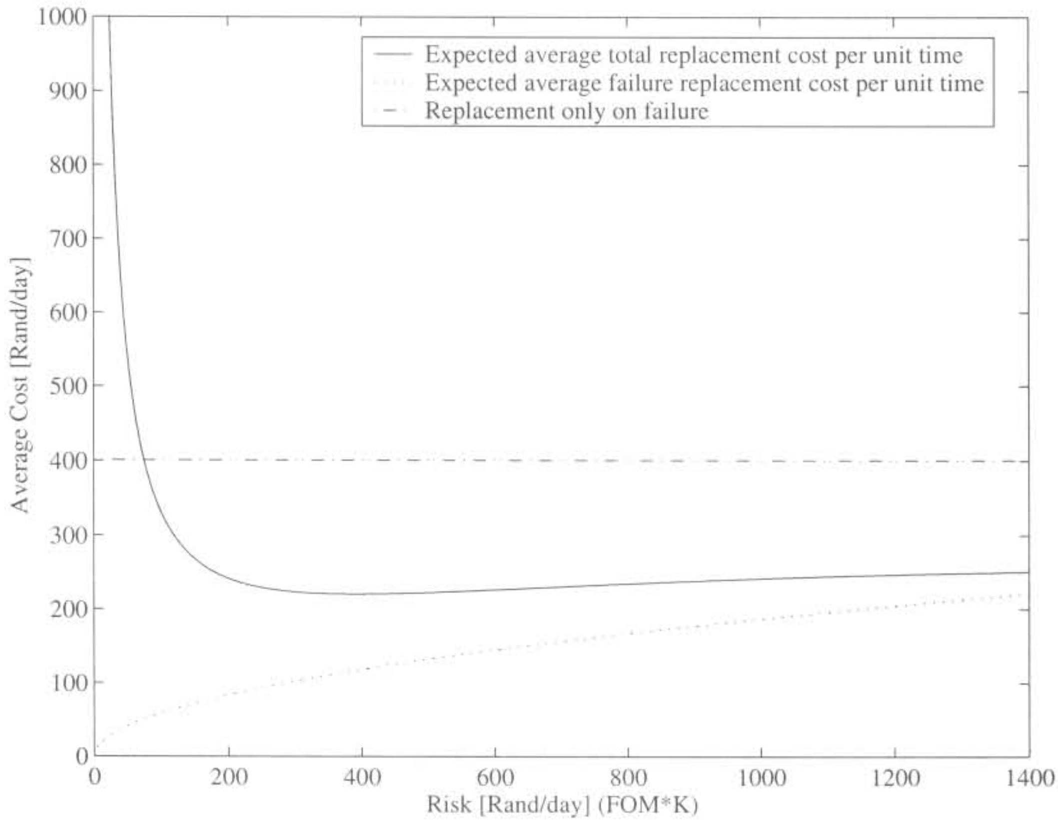

Figure 5.5: Expected cost in terms of risk

Figure 5.5 shows a distinct optimum at a risk of R 401.41 / day or a FOM of $h(x, \boldsymbol{z}) = 0.0029$. If renewal is always performed at this risk, the long term cost is expected to be R 224.04 / day. This optimum is not very sensitive to slight deviations from the decision rule. With the optimal risk known it is also possible to present the renewal rule (equation (5.9)) and warning level function (equation (5.11)) graphically as shown in Figure 5.6.

University of Pretoria etd – Vlok, P-J (2006)

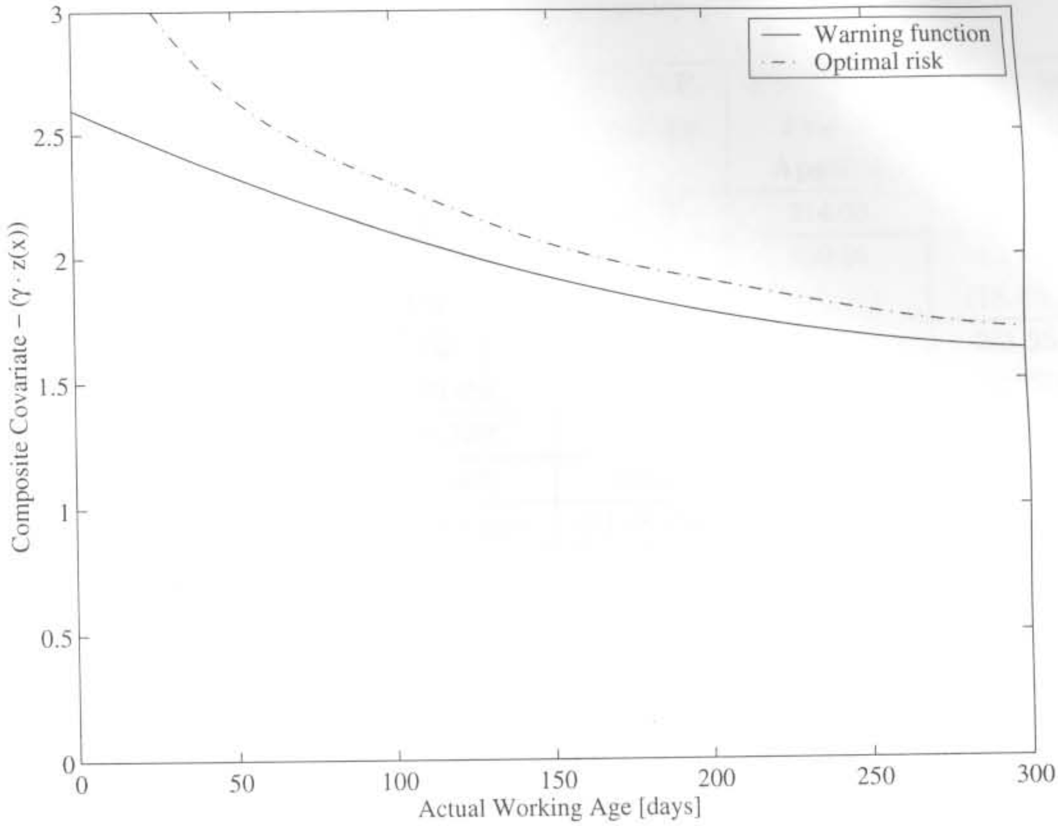

Figure 5.6: Optimal decision policy with warning function

The renewal policy is evaluated in detail in the next section.

5.3.4 Evaluation of renewal policy

A summary of the performance of the decision-model of Makis and Jardine on the SASOL data set is presented in Table 5.7. Four criteria were used to evaluate the decision-model under the following headings:

- (i) *Theoretical optimal policy.* The theoretical average costs.
- (ii) *Renew only on failure (R.O.O.F).* A prediction of costs if a corrective maintenance strategy was followed.
- (iii) *Theoretical policy applied.* An estimation of costs if the decision-model was applied to the observed data.
- (iv) *Observed policy.* The true costs incurred by the plant.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.7: Summary of optimal policy performance

	Theoretical Optimal Policy	R.O.O.F. Strategy	Theoretical Policy Applied	Observed Policy
Cost	224.04	401.41	214.03	345.16
Preventive Renewal Cost	75.31 (33.6%)	0 (0%)	100.56 (47.0%)	63.21 (18.3%)
Failure Renewal Cost	148.73 (66.4%)	401.41 (100%)	113.47 (53.0%)	281.95 (81.7%)
% Preventive Renewals	76.70%	0%	80.00%	42.10%
% Failure Renewals	23.30%	100%	20.00%	57.90%
MTBR	254.49 days	404.08 days	263.6 days	214.6 days

**All costs are in R/day*

The two most important figures in Table 5.7 are the cost per day if the theoretical policy was applied (R 214.03) and the cost per day that was actually observed (R 345.16). It is also important to compare the percentage of preventive renewals with the percentage of failure renewals of the theoretical policy and the observed policy. It is clear that the decision-model of Makis and Jardine is not only considerably less expensive but also more orderly because of 80% of events would have been suspensions if the theoretical policy was applied. Table 5.7 is analyzed in detail in Vlok (1999).

Such coincidence of the theoretical and actual results in some of the above cases should not be expected in general, particularly for a small sample size, but it shows that the selected PHM and decision-model are realistic. The method of comparison could be challenged however because the same data that is used to build the model is used to evaluate it. For this reason a final test of the decision policy was performed by collecting more data from the plant from November 1, 1998 to February 28, 1999. During this period only one of the pumps considered as calendar suspensions in the first data set failed and was renewed. The decision policy's performance for this pump's history is described here, although the data from the other pumps was tested as well.

Pump PC1232 was treated as a calendar suspension after 192 days of working life in the first data set. This was on November 1, 1998. The pump eventually failed unexpectedly 67 days later on January 6, 1999 at an age of 259 days. A total of five inspections were done during this time. The latest inspection data is shown on Figure 5.7, together with the three inspections from the first data set.

Figure 5.7 shows that the unexpected failure could have been prevented if the calculated decision policy was followed. In monetary terms, the unexpected failure cost resulted in

University of Pretoria etd – Vlok, P-J (2006)

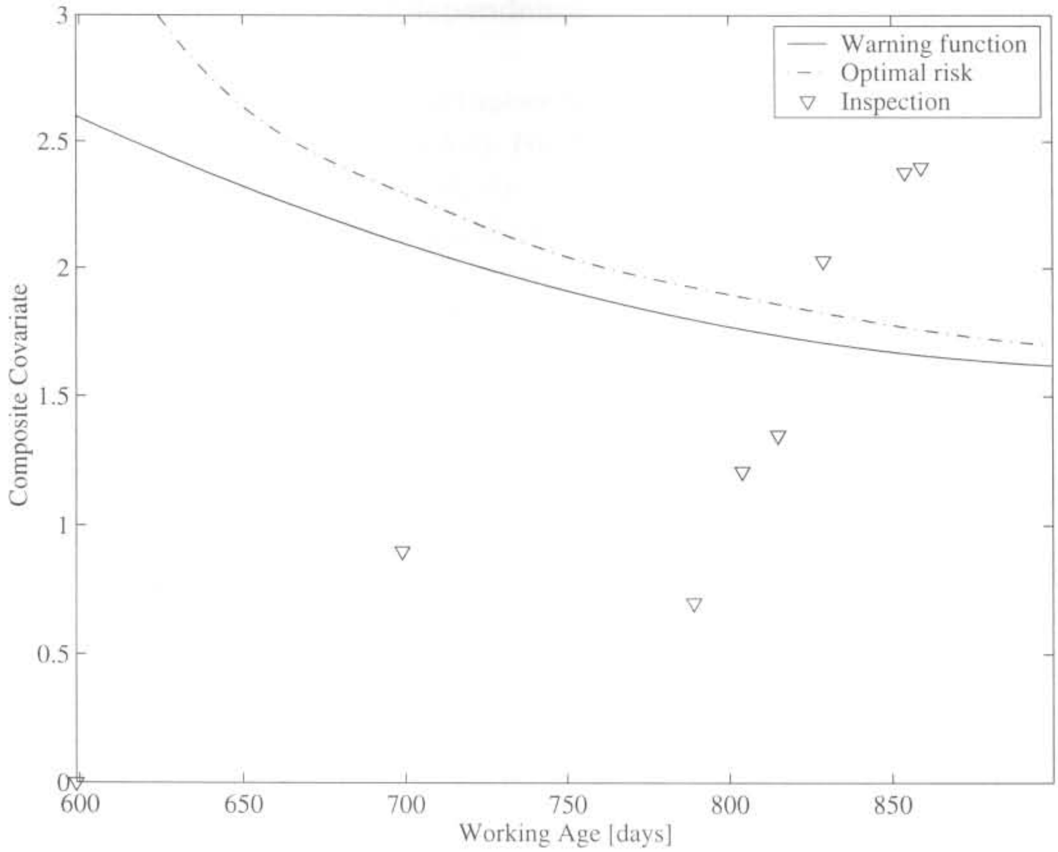

Figure 5.7: Decision-model applied on PC1232

R 162 200/259 days = R 626.25/day. If the calculated policy was available and acted upon, R 25 000/235 days = R 106.38/day, would have been the result. This is another confirmation that the model is relevant and practical.

5.4 Maintenance strategy optimization through combined PIMs and residual life estimation

The modeling methodology proposed in Figure 3.1 is followed in this section to model the SASOL data. It starts off with tests to determine wether non-repairable or repairable systems theory is more applicable for this data set, including tests for trend and dependence. Following this, parametric approximations for the covariates are calculated and different combined PIMs are fitted on the data before the best combined PIM is selected.

University of Pretoria etd – Vlok, P-J (2006)

5.4.1 Testing for trend and dependence

It was motivated in Section 3.1 that the Laplace test will be used to test for trend in the data (Laplace's test is described in Section A.1). Four or more event observations are required to reach a 95% level of confidence of trend. For this reason Laplace's test was only applied on 3 of the 8 pumps, i.e. PC1131, PC1132 and PC2131. The results were as follows: $U_{PC1131} = 1.8043$, $U_{PC1132} = 1.6663$ and $U_{PC2131} = 1.0444$. In all three cases Laplace's test confirmed that the event data is not non-committal and shows signs of reliability degradation, i.e. interarrival times become shorter.

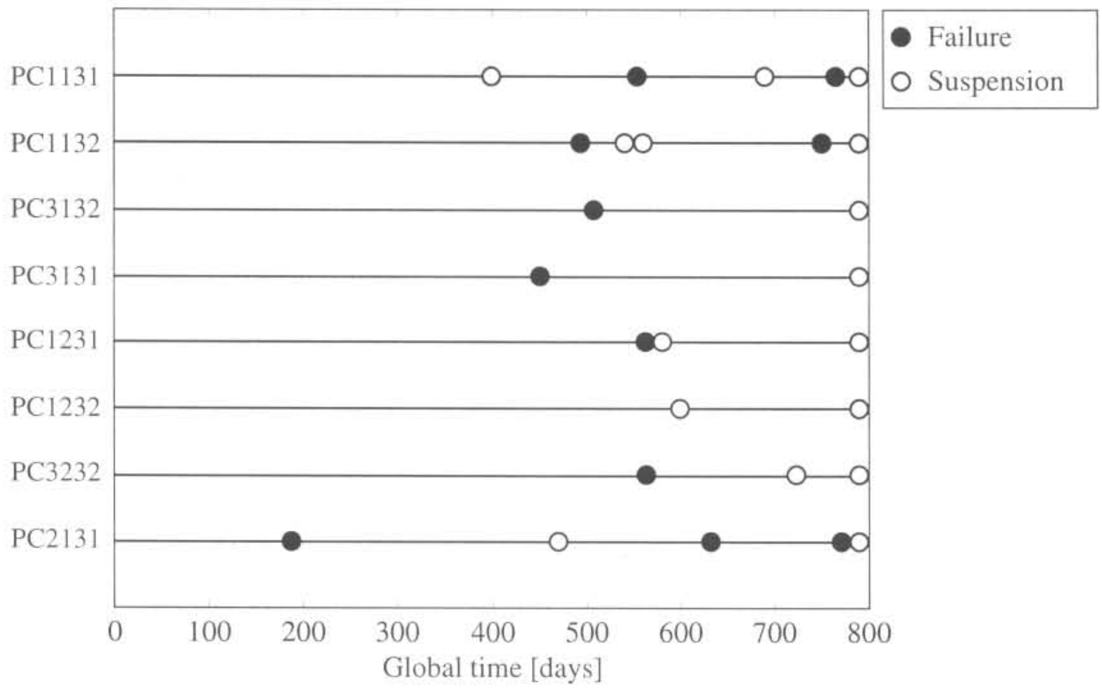

Figure 5.8: Graphic illustration of event times

For the remaining pumps, i.e. PC3132, PC3131, PC1231, PC1232 and PC3232, it is very difficult to proof or deny evidence of a trend mathematically because insufficient data is available. In these cases, simple "eyeball-analysis" is often the most reliable and most effective according to Ascher (1999). The representation of the event data in Figure 5.8 is ideal for eyeball-analysis because each pump's history can not only be evaluated individually but also compared with the others. Eyeball-analysis cannot provide any additional information with regards to trend about PC3132, PC3131 and PC1232 because only one event was observed for each pump (excluding the calendar suspensions). PC1231 and PC3232 appear to be in the process of reliability degradation since interarrival times become shorter as the global time increases.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.8: Summary of functions used to approximate observed covariate behaviour in terms of global time t

Pump	Lifetime	RF53H	RF54H
PC1131	1	$1.67e-6 \cdot t^3 - 0.000812 \cdot t^2 + 0.0938 \cdot t - 0.0399$	$5.71e-007 \cdot t^3 - 0.000254 \cdot t^2 + 0.0262 \cdot t + 0.00841$
	2	$0.000602 \cdot t^2 - 0.47 \cdot t + 91.7$	$0.000793 \cdot t^2 - 0.69 \cdot t + 149$
	3	$0.000983 \cdot t + 0.897$	$-0.00107 \cdot t^2 + 1.33 \cdot t - 411$
	4	$0.0362 \cdot t - 24.9$	$0.00628 \cdot t - 4.16$
	5	$0.463 \cdot t - 354$	$0.27 \cdot t - 207$
PC1132	1	$-6.63e-007 \cdot t^3 + 0.000469 \cdot t^2 - 0.0684 \cdot t + 0.00656$	$0.000277 \cdot t^2 - 0.0818 \cdot t + 0.741$
	2	$0.178 \cdot t - 86.7$	$0.0499 \cdot t - 18.9$
	3	$0.155 \cdot t - 84.4$	$-0.194 \cdot t - 105$
	4	1.63	$8.7e-006 \cdot t^3 - 0.0173 \cdot t^2 + 11.4 \cdot t - 2.49e+003$
	5	$0.0504 \cdot t - 37.8$	$0.192 \cdot t - 145$
PC1231	1	$5.17e-005 \cdot t^2 - 0.0115 \cdot t + 0.125$	$-1.85e-007 \cdot t^3 + 0.000183 \cdot t^2 - 0.0341 \cdot t + 0.0254$
	2	$0.403 \cdot t - 227$	$0.604 \cdot t - 340$
	3	$2.57e-005 \cdot t^2 - 0.0267 \cdot t + 6.92$	$9.54e-005 \cdot t^2 - 0.119 \cdot t + 36.9$
PC1232	1	2.92	1.78
	2	$0.0214 \cdot t - 12.7$	1.05
PC2131	1	$0.000112 \cdot t^2 - 0.0143 \cdot t - 0.00117$	$-2.85e-005 \cdot t^2 + 0.00735 \cdot t + 0.000168$
	2	$0.000197 \cdot t^2 - 0.103 \cdot t + 13.9$	$0.000146 \cdot t^2 - 0.0662 \cdot t + 7.68$
	3	$0.00577 \cdot t^2 - 6.09 \cdot t + 1.59e+003$	$-0.00103 \cdot t^2 + 1.14 \cdot t - 306$
	4	$-0.0127 \cdot t + 10.5$	$-0.0106 \cdot t + 8.41$
	5	$0.218 \cdot t - 169$	$0.331 \cdot t - 257$
PC3131	1	$-0.000107 \cdot t^2 + 0.0566 \cdot t - 0.42$	$-4.89e-005 \cdot t^2 + 0.0315 \cdot t - 0.396$
	2	$-0.000333 \cdot t^2 + 0.429 \cdot t - 128$	$-0.000158 \cdot t^2 + 0.209 \cdot t - 63.3$
PC3132	1	$0.000219 \cdot t^2 - 0.0662 \cdot t + 0.755$	$9e-005 \cdot t^2 - 0.024 \cdot t + 0.305$
	2	$3.11e-005 \cdot t^2 - 0.0288 \cdot t + 6.69$	$-1.95e-005 \cdot t^2 + 0.035 \cdot t - 12.8$
PC3232	1	$0.014 \cdot t - 0.215$	$3.54e-005 \cdot t^2 - 0.00329 \cdot t + 0.0106$
	2	$-0.000685 \cdot t^2 + 0.873 \cdot t - 271$	$-0.00121 \cdot t^2 + 1.54 \cdot t - 477$
	3	$0.0715 \cdot t - 52.2$	$0.0416 \cdot t - 30.3$

University of Pretoria etd – Vlok, P-J (2006)

The parametric functions of Table 5.8 are also shown graphically in Appendix E with the 95% confidence intervals. In some cases confidence intervals could not be calculated because of a lack of data.

5.4.4 Estimation of the PIMs

It was predicted in Section 3.3 that the entire model would probably never be applied to a single situation because of data constraints. In this case the statement is true and certain assumptions for, and simplifications to, the general model of equation (3.30) are required for it to be applicable to the present data set. The most important characteristics of the data set are listed before assumptions and simplifications are made:

(i) Eight system copies have been observed operating in nominally similar conditions over a period of 791 days.

(ii) Eleven failures, eight suspensions and eight calendar suspensions were recorded during the 791 days.

(iii) Two components on each system have been observed, i.e. Bearing 3 and Bearing 4.

(iv) Covariate levels (vibration levels) on Bearing 3 and 4 were recorded at irregular inspection intervals during the 791 days.

(v) The data does not contain any information about the cause of failure, i.e. whether Bearing 3 or Bearing 4 failed or another component that was not observed.

(vi) Following from Section 5.4.1, event data recorded on the systems appear to follow repairable systems theory.

This data set complies to Scenario 4 of Figure 3.6, i.e. a single-component repairable system, even though two components have been observed. The reason for assuming the systems to be single-component repairable systems is because no information is available about the cause of system failure, i.e. whether Bearing 3 or Bearing 4 was responsible for the failure. This implies implicitly that it is assumed that RF53H and RF54H relate to the entire system and not only the two bearings, which is largely true. The use of competing risks is immediately eliminated by this assumption, i.e. summing over individual peril rates of each component is not possible.

Following the summary and assumptions above, the following enhancements are permitted in equation (3.30):

(i) Full stratification and system copy dependency of all coefficients.

(ii) Frailties.

University of Pretoria etd – Vlok, P-J (2006)

(iii) Time jumps or setbacks.
(iv) Acceleration or deceleration of the global time.
(v) Multiplicative or additive functional terms.
(vi) Time-dependent covariates.

In Section 3.1 it was remarked that data sets are often modeled in literature with only one particular enhancement because it is such a laborious task to estimate parameters for a model. It is usually required to develop a virtually unique algorithm to fit a model on any particular data set. For this reason, an algorithm was developed to fit the completely general model of equation (3.30) to the "perfect" data set, i.e. a data set with sufficient definition and observations to satisfy the requirements of (3.30). By restricting the appropriate variables in the algorithm, it is possible to use the algorithm to fit a combined PIM with any combination of enhancements to a data set, similar to what was illustrated in Appendix B.

The generic algorithm made it possible to experiment with countless different combinations of enhancements. In the subsections to follow, the best combinations of the possible enhancements fitted on the data are described. Each model's performance is evaluated by the following criteria:

(i) The sum of squared errors of residual life estimates. Errors on the estimates of calendar suspensions are not included in the sum of squared errors although normal right-censored observations are taken into account.
(ii) The sum (total number of days) of the confidence intervals produced by a particular model.

A model performing well against both the above-mentioned criteria should be a useful tool in practical maintenance decision-making.

5.4.4.1 Combined PIM simplified to the conventional $\rho_1(t)$ model without covariates or stratifications

Model description

This model is the conventional $\rho_1(t)$ model (described in Section 1.2.3.2) often used in reliability literature. Although it is a simplification of the combined PIM, it is not a PIM by definition because it does not rely on intensity proportions. It is presented in this section however, because it performed fairly well and to emphasize the advantages of enhancements in the combined PIM, illustrated later in this section.

Model construction

University of Pretoria etd – Vlok, P-J (2006)

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.9.

Table 5.9: Parameter restrictions on equation (3.30) to obtain a conventional $\rho_1(t)$ model without covariates or stratifications

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1$
s:	$s^l = 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = 0$, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t) = \exp(\Gamma + \Upsilon t) \tag{5.12}$$

When this model was fitted to the SASOL data using Snyman's technique (see Section C.2), the log-likelihood converged at a maximum of $L(\hat{\boldsymbol{\theta}}) = -124.88$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -8.4859$ and $\hat{\Upsilon} = 0.0064$. This model is evaluated in the next section.

Model evaluation

Since there are no covariates present in the model in equation (5.12), dynamic residual life estimates are not possible and estimates remain constant for the entire duration of a system's life. Residual life estimates‖ were calculated at the start of every lifetime of each pump with 2-sided confidence intervals of 95%. Because this model is neither system copy nor stratum specific, predictions for the time to the first event on all pumps are exactly the same. Estimates and actual observations are summarized in Table 5.10. For easy comparison with actual observations, estimated arrival times are reported and not residual life.

A total of five events were observed outside the bounds forecasted by the model. Only one of these five events was not a calendar suspension at 791 days, which shows that this simple model fits the data fairly well.

‖Residual life is exactly equal to the expected time to the next event when the local time is zero.

University of Pretoria etd – Vlok, P-J (2006)

To quantify the quality of the model, squared errors on the estimates were calculated and summed to obtain an indication of the model's accuracy. The width of confidence intervals were also summed to quantify the certainty of the model. Estimates on calendar suspensions do not contribute to squared errors, although normal right-censored observations were taken into account. The sum of the squared errors is 2.3771e5 and the sum of all the widths of confidence intervals is 8201**.

5.4.4.2 Combined PIM simplified to the conventional $\rho_1(t)$ model with stratified time jump/setback coefficients

Model description

The conventional $\rho_1(t)$ model is used again but with the inclusion of τ_s to allow for time jumps/setbacks as a function of the particular stratum that a system is in.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.11.

Table 5.11: Parameter restrictions on equation (3.30) to obtain a conventional $\rho_1(t)$ model with stratified time jump/setback coefficients

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	1
s:	s^l	$=$	$N(t) + 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	τ_s, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t) = \exp(\Gamma + \Upsilon(t - \tau_s)) \tag{5.13}$$

The log-likelihood was maximized using Snyman's technique (see Section C.2) and converged where $L(\hat{\boldsymbol{\theta}}) = -130.43$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -10.5049$,

**These values are compared for each combined PIM in Section 5.4.4.6.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.10: Conventional $\rho_1(t)$ model without covariates or stratifications

Model description: $\rho_1(t) = \Gamma + \Upsilon t$

Estimated parameters: $\hat{\Gamma} = -8.4859$ and $\hat{\Upsilon} = 0.0064$

Pump ID	1st event		2nd event		3rd event		4th event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$149 \leq 467 \leq 711$	554	$414 \leq 559 \leq 727$	690	$560 \leq 640 \leq 757$	765	$692 \leq 735 \leq 807$
PC1132	491	$149 \leq 467 \leq 711$	544	$500 \leq 604 \leq 742$	557	$563 \leq 642 \leq 758$	751	$563 \leq 642 \leq 758$
PC1231	563	$149 \leq 467 \leq 711$	578	$569 \leq 646 \leq 759$	791	$583 \leq 655 \leq 764$	-	-
PC1232	599	$149 \leq 467 \leq 711$	791	$603 \leq 669 \leq 770$	-	-	-	-
PC2131	184	$149 \leq 467 \leq 711$	470	$243 \leq 493 \leq 712$	631	$481 \leq 593 \leq 738$	774	$635 \leq 691 \leq 782$
PC3131	450	$149 \leq 467 \leq 711$	791	$462 \leq 583 \leq 735$	-	-	-	-
PC3132	506	$149 \leq 467 \leq 711$	791	$514 \leq 612 \leq 745$	-	-	-	-
PC3232	563	$149 \leq 467 \leq 711$	723	$569 \leq 646 \leq 759$	791	$725 \leq 761 \leq 823$	-	-

Pump ID	5th event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$766 \leq 795 \leq 847$	3.6862e4	1268
PC1132	791	$752 \leq 783 \leq 839$	2.9528e4	1281
PC1231	-	-	1.3045e4	933
PC1232	-	-	5.0136e3	729
PC2131	791	$775 \leq 803 \leq 852$	1.1567e5	1512
PC3131	-	-	1.7793e4	835
PC3132	-	-	1.1375e4	793
PC3232	-	-	8.4231e3	850
			2.3771e5	8201

University of Pretoria etd – Vlok, P-J (2006)

$\hat{\Upsilon} = 0.0099$, $\hat{\tau_1} = 204.73$, $\hat{\tau_2} = 12.50$, $\hat{\tau_3} = -34.86$, $\hat{\tau_4} = -86.35$ and $\hat{\tau_5} = -80.05$. This model is evaluated in the next section.

Model evaluation

Since there are no covariates present in the model in equation (5.13), dynamic residual life estimates are not possible and estimates remain constant for the duration of a system's lifetime. Residual life estimates were calculated at the start of every lifetime of each pump with 2-sided confidence intervals of 95%. Because this model is not system copy specific, predictions for the time to first event on all pumps are exactly the same. Estimates and actual observations are summarized in Table 5.12. For easy comparison with actual observations, estimated arrival times are reported and not residual life.

A total of twelve events were observed outside the bounds forecasted by the model. A total of eight of these twelve events were not calendar suspensions at 791 days, which is a first indication that this model does not fit the data very well.

To quantify the quality of the model, squared errors on the estimates were calculated and summed to obtain an indication of the model's accuracy. The width of confidence intervals were also summed to quantify the certainty of the model. Estimates on calendar suspensions do not contribute to squared errors, although normal right-censored observations were taken into account. The sum of the squared errors is 3.5171e5 and the sum of all the widths of confidence intervals is 6663[††].

5.4.4.3 Combined PIM simplified to an additive intensity model with stratified regression coefficients

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with an additive term of exponential form containing covariates (similar to the model described in Section 2.3.2). Regression coefficients are stratified into only two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.13.

[††] These values are compared for each combined PIM in Section 5.4.4.6.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.12: $\rho_1(t)$ model with stratified time jump/setback coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau_s))$

Estimated parameters: $\hat{\Gamma} = -10.5049$, $\hat{\Upsilon} = 0.0099$, $\hat{\tau_1} = 204.73$, $\hat{\tau_2} = 12.50$, $\hat{\tau_3} = -34.86$, $\hat{\tau_4} = -86.35$ and $\hat{\tau_5} = -80.05$

Pump ID	1st event		2nd event		3rd event		4th event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$124 \leq 339 \leq 500$	554	$424 \leq 552 \leq 688$	690	$563 \leq 641 \leq 727$	765	$693 \leq 734 \leq 763$
PC1132	491	$124 \leq 339 \leq 500$	544	$501 \leq 582 \leq 749$	557	$554 \leq 637 \leq 749$	751	$572 \leq 672 \leq 749$
PC1231	563	$124 \leq 339 \leq 500$	578	$568 \leq 613 \leq 732$	791	$585 \leq 657 \leq 756$	-	-
PC1232	599	$124 \leq 339 \leq 500$	791	$602 \leq 650 \leq 725$	-	-	-	-
PC2131	184	$124 \leq 339 \leq 500$	470	$313 \leq 528 \leq 729$	631	$491 \leq 608 \leq 757$	774	$638 \leq 706 \leq 772$
PC3131	450	$124 \leq 339 \leq 500$	791	$466 \leq 578 \leq 700$	-	-	-	-
PC3132	506	$124 \leq 339 \leq 500$	791	$515 \leq 600 \leq 706$	-	-	-	-
PC3232	563	$124 \leq 339 \leq 500$	723	$568 \leq 613 \leq 732$	791	$723 \leq 746 \leq 810$	-	-

Pump ID	5th event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$766 \leq 775 \leq 822$	4.2037e3	1074
PC1132	791	$752 \leq 767 \leq 809$	3.6346e4	1042
PC1231	-	-	5.1258e4	711
PC1232	-	-	6.7388e4	499
PC2131	791	$775 \leq 780 \leq 811$	9.6290e4	1450
PC3131	-	-	1.0125e4	709
PC3132	-	-	2.7258e4	567
PC3232	-	-	5.8848e4	611
			3.5171e5	6663

University of Pretoria etd – Vlok, P-J (2006)

Table 5.13: Parameter restrictions on equation (3.30) to obtain an additive intensity model with stratified regression coefficients

Parameter	Restriction		
n:	n	=	1, thus $l = 1$
k:	k	=	1
s:	s^l	=	1 where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$ for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	=	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	=	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	=	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	=	α_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	=	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.14}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -109.02$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -11.1674$, $\hat{\Upsilon} = 0.013$, $\hat{\alpha_{1_1}} = -0.6760$, $\hat{\alpha_{1_2}} = 0.5408$, $\hat{\alpha_{2_1}} = 2.1457$ and $\hat{\alpha_{2_2}} = 3.1665$. This model is evaluated in the next section.

Model evaluation

The covariates present in the model in equation (5.14), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the last inspection of every lifetime of each pump is reported in Table 5.14. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of eight events were observed outside the bounds forecasted by the model. Only two of these eight events were at calendar suspensions, which is an early indication that the model

University of Pretoria etd – Vlok, P-J (2006)

does not fit the data well. This is confirmed by the sum of the squared errors of 1.0599e5 and the sum of the widths of the confidence intervals of 5990.

5.4.4.4 Combined PIM simplified to a multiplicative intensity model with stratified regression coefficients

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with a multiplicative term of exponential form containing covariates, similar to the model by Kumar (1996) that was introduced in Section 2.3.1.2. Regression coefficients are stratified into two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.15.

Table 5.15: Parameter restrictions on equation (3.30) to obtain a multiplicative intensity model with stratified regression coefficients

Parameter	Restriction		
n:	n	=	1, thus $l = 1$
k:	k	=	1
s:	s^l	=	1 where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$ for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l}$	=	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	=	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	=	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	=	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	=	γ_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z}) \tag{5.15}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -142.66$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -6.2011$, $\hat{\Upsilon} = 0.00046$, $\hat{\gamma_{1_1}} = 1.4021$, $\hat{\gamma_{1_2}} = 0.9741$, $\hat{\gamma_{2_1}} = 1.002$ and $\hat{\gamma_{2_2}} = 0.6231$. This model is evaluated in the next section.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.14: Additive intensity model with $\rho_1(t)$ as baseline and stratified regression coefficients

Model description: $\rho(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -11.1674$, $\hat{\Upsilon} = 0.013$, $\hat{\alpha_{1_1}} = -0.6760$, $\hat{\alpha_{1_2}} = 0.5408$, $\hat{\alpha_{2_1}} = 2.1457$ and $\hat{\alpha_{2_2}} = 3.1665$

Pump ID	1st event		2nd event		3rd event		4th event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$91 \leq 400 \leq 562$	554	$414 \leq 490 \leq 602$	690	$576 \leq 595 \leq 670$	765	$740 \leq 810 \leq 1040$
PC1132	491	$182 \leq 316 \leq 420$	544	$532 \leq 560 \leq 721$	557	$566 \leq 612 \leq 790$	751	$628 \leq 725 \leq 824$
PC1231	563	$390 \leq 501 \leq 661$	578	$591 \leq 649 \leq 708$	791	$740 \leq 759 \leq 881$	-	-
PC1232	599	$224 \leq 480 \leq 585$	791	$782 \leq 890 \leq 975$	-	-	-	-
PC2131	184	$112 \leq 241 \leq 410$	470	$390 \leq 422 \leq 577$	631	$544 \leq 718 \leq 785$	774	$673 \leq 770 \leq 890$
PC3131	450	$191 \leq 377 \leq 595$	791	$675 \leq 710 \leq 801$	-	-	-	-
PC3132	506	$279 \leq 409 \leq 500$	791	$721 \leq 845 \leq 921$	-	-	-	-
PC3232	563	$356 \leq 551 \leq 677$	723	$640 \leq 795 \leq 890$	791	$746 \leq 880 \leq 986$	-	-

Pump ID	5th event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$782 \leq 798 \leq 856$	1.5155e4	1127
PC1132	791	$812 \leq 917 \leq 999$	3.4582e4	1034
PC1231	-	-	8.8850e3	529
PC1232	-	-	1.4161e4	554
PC2131	791	$802 \leq 815 \leq 843$	1.3138e4	984
PC3131	-	-	5.3290e3	530
PC3132	-	-	9.4090e3	421
PC3232	-	-	5.3280e3	811
			1.0599e5	5990

University of Pretoria etd – Vlok, P-J (2006)

Model evaluation

The covariates present in the model in equation (5.15), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the second last inspection of every lifetime of each pump is reported in Table 5.16. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump, contrary to the first two models. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of seven events were observed outside the bounds forecasted by the model. Only two of these seven events were at calendar suspensions, which is an early indication that the model does not fit the data very well. The sum of the squared errors is 2.4388e5 and the total width of the confidence bands is 12768.

5.4.4.5 Combined PIM simplified to an additive intensity model with a time jump/setback in the baseline

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with an additive term of exponential form containing covariates. The baseline also allows for a time jump/setback and regression coefficients are stratified into two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.17.

Table 5.17: Parameter restrictions on equation (3.30) to obtain an additive intensity model with stratified coefficients and a time jump/setback in the baseline

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1$
s:	$s^l = 1$ where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$

University of Pretoria etd – Vlok, P-J (2006)

			for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	τ, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	α_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.16}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -128.21$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -9.2212$, $\hat{\Upsilon} = 0.0042$, $\hat{\tau} = -22.02$, $\hat{\alpha_{1_1}} = 2.3061$, $\hat{\alpha_{1_2}} = 1.8036$, $\hat{\alpha_{2_1}} = 0.9261$ and $\hat{\alpha_{2_2}} = 1.5881$. This model is evaluated in the next section.

Model evaluation

The covariates present in the model in equation (5.16), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the second last inspection of every lifetime of each pump is reported in Table 5.18. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump, contrary to the first two models. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of nine events were observed outside the bounds forecasted by the model. Only one of these nine events was observed at a calendar suspension of 791 days. This model generally fits the data very well with the sum of squared errors being 4.8748e4 and the sum of the confidence interval widths being 3475[‡‡].

[‡‡]These values are compared for each combined PIM in Section 5.4.4.6.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.16: Multiplicative intensity model with stratified regression coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -6.2011$, $\hat{\Upsilon} = 0.00046$, $\hat{\gamma_{1_1}} = 1.4021$, $\hat{\gamma_{1_2}} = 0.9741$, $\hat{\gamma_{2_1}} = 1.002$ and $\hat{\gamma_{2_2}} = 0.6231$

	1st event		2nd event		3rd event		4th event	
Pump ID	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$201 \leq 590 \leq 721$	554	$420 \leq 435 \leq 462$	690	$591 \leq 766 \leq 1760$	765	$712 \leq 883 \leq 1051$
PC1132	491	$500 \leq 525 \leq 592$	544	$526 \leq 640 \leq 792$	557	$570 \leq 750 \leq 911$	751	$599 \leq 802 \leq 1096$
PC1231	563	$126 \leq 744 \leq 1009$	578	$622 \leq 746 \leq 918$	791	$615 \leq 791 \leq 983$	-	-
PC1232	599	$165 \leq 662 \leq 1105$	791	$682 \leq 960 \leq 1223$	-	-	-	-
PC2131	184	$186 \leq 290 \leq 407$	470	$202 \leq 530 \leq 751$	631	$507 \leq 655 \leq 817$	774	$709 \leq 787 \leq 915$
PC3131	450	$218 \leq 582 \leq 811$	791	$491 \leq 858 \leq 1179$	-	-	-	-
PC3132	506	$136 \leq 511 \leq 837$	791	$527 \leq 958 \leq 1313$	-	-	-	-
PC3232	563	$82 \leq 434 \leq 722$	723	$610 \leq 812 \leq 989$	791	$770 \leq 942 \leq 1276$	-	-

	5th event		Σ Squared	Σ Confidence
Pump ID	Obs.	Est.	Errors	Intervals
PC1131	791	$795 \leq 894 \leq 1121$	7.1110e4	2396
PC1132	791	$789 \leq 936 \leq 1085$	5.0222e4	1492
PC1231	-	-	6.0985e4	1547
PC1232	-	-	3.9690e3	1481
PC2131	791	$832 \leq 902 \leq 1105$	1.5581e4	1559
PC3131	-	-	1.7424e4	1281
PC3132	-	-	2.5000e1	1487
PC3232	-	-	2.4562e4	1525
			2.4388e5	12768

University of Pretoria etd – Vlok, P-J (2006)

Table 5.18: Additive intensity model with a time jump/setback in the baseline and stratified regression coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t+\tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -9.2212$, $\hat{\Upsilon} = 0.0042$, $\hat{\tau} = -22.02$, $\hat{\alpha_{1_1}} = 2.3061$, $\hat{\alpha_{1_2}} = 1.8036$, $\hat{\alpha_{2_1}} = 0.9261$ and $\hat{\alpha_{2_2}} = 1.5881$

	1st event		2nd event		3rd event		4th event	
Pump ID	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$361 \leq 454 \leq 565$	554	$415 \leq 467 \leq 555$	690	$640 \leq 693 \leq 738$	765	$705 \leq 770 \leq 815$
PC1132	491	$459 \leq 528 \leq 609$	544	$557 \leq 608 \leq 636$	557	$433 \leq 545 \leq 616$	751	$725 \leq 778 \leq 808$
PC1231	563	$516 \leq 555 \leq 624$	578	$592 \leq 612 \leq 659$	791	$621 \leq 694 \leq 775$	-	-
PC1232	599	$550 \leq 581 \leq 613$	791	$662 \leq 815 \leq 894$	-	-	-	-
PC2131	184	$125 \leq 199 \leq 225$	470	$473 \leq 537 \leq 603$	631	$476 \leq 593 \leq 703$	774	$645 \leq 721 \leq 749$
PC3131	450	$339 \leq 401 \leq 432$	791	$731 \leq 753 \leq 773$	-	-	-	-
PC3132	506	$336 \leq 407 \leq 463$	791	$780 \leq 859 \leq 930$	-	-	-	-
PC3232	563	$516 \leq 629 \leq 711$	723	$629 \leq 656 \leq 757$	791	$823 \leq 868 \leq 966$	-	-

	5th event		Σ Squared	Σ Confidence
Pump ID	Obs.	Est.	Errors	Intervals
PC1131	791	$770 \leq 799 \leq 820$	1.0852e4	624
PC1132	791	$763 \leq 868 \leq 903$	6.3380e3	635
PC1231	-	-	1.2200e3	329
PC1232	-	-	3.2400e2	245
PC2131	791	$786 \leq 820 \leq 894$	8.9670e3	764
PC3131	-	-	2.4010e3	135
PC3132	-	-	9.8010e3	277
PC3232	-	-	8.8450e3	466
			4.8748e4	3475

University of Pretoria etd – Vlok, P-J (2006)

5.4.4.6 Comparison of different combined PIMs' performances

In Table 5.19 the performances of the different combined PIMs used in Section 5.4.4 are summarized. The models are sorted by the magnitude of the sum of squared errors in descending order.

Table 5.19: Comparison of different combined PIMs' performance

No.	Combined PIM	Σ Squared errors	Σ Confidence bounds
1.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau_s))$	3.5171e5	6663
2.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z})$	2.4388e5	12768
3.	$\rho_1(t) = \exp(\Gamma + \Upsilon t)$	2.3771e5	8201
4.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$	1.0599e5	5990
5.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$	4.8748e4	3475

Model 5 in Table 5.19 produced the lowest sum of squared errors as well as the lowest sum of confidence bounds. By removing the time setback τ from Model 5, Model 4 is obtained. This model performed significantly worse that Model 5 which stresses the usefulness of the combined PIM. It is also significant to note that the conventional $\rho_1(t)$ model, i.e. Model 3 of Table 5.19, performed better than Models 1 and 2 which are much more sophisticated models. The multiplicative PIM, Model 2 of Table 5.19, had the largest sum of confidence bounds which indicates that the conditional probability densities produced by this model are fairly broad compared to, for example, Model 5 of Table 5.19.

5.5 Comparing the performance of the RLE approach with the combined PIM to the approach of Makis and Jardine

In the introduction of this chapter, the importance of comparing the RLE approach with the established approach of Makis and Jardine was stressed. In this section, the performance of Model 5 of Table 5.19 on the SASOL data is compared to the performance of the policy of Makis and Jardine as described in Section 5.3.4. The criteria for comparison is the "Theoretical Policy Applied" as defined earlier. Table 5.7 of Section 5.3.4 is partially repeated here as Table 5.20, including the comparative values for the RLE approach.

University of Pretoria etd – Vlok, P-J (2006)

Table 5.20: Summary of the comparison between the RLE approach and the approach of Makis and Jardine

	RLE Approach	Makis and Jardine	Observed Policy
Cost	205.22	214.03	345.16
Preventive Action Cost	129.36 (63.03%)	100.56 (47.0%)	63.21 (18.3%)
Corrective Action Cost	75.86 (36.97%)	113.47 (53.0%)	281.95 (81.7%)
% Preventive Action	88.46%	80.00%	42.10%
% Corrective Action	11.54%	20.00%	57.90%
MTBR	248.06 days	263.6 days	214.6 days

**All costs are in R/day*

The most important figure in Table 5.20 is the cost per day of each policy. If the RLE approach was applied to the actual situation, a cost of R 205.22 / day would be the result, which is 4.1% lower than the approach of Makis and Jardine of R 214.03 /day and 40.5% lower than the observed policy of R 345.16 / day. Preventive action is prescribed by the RLE approach in 88.46% of all observed cases which is 8.46% more than the policy of Makis and Jardine. In this particular case the RLE approach was thus a more conservative policy compared to Makis and Jardine's approach. The high percentage of preventive actions leads to a relatively high percentage of preventive action cost as well as a MTBR of 248.06 days, which is 15.54 days shorter than Makis and Jardine's policy. Although the RLE approach produced marginally better results than Makis and Jardine's policy, are both significant improvements on the observed policy.

To further compare the the RLE approach's performance to the policy of Makis and Jardine, Model 5 of Table 5.19 was also applied to the second data set that was compiled for Pump PC1232 (see Section 5.3.4). The result is shown graphically in Figure 5.9 in the format proposed in Section 4.5.

Figure 5.9 shows the entire history of the second lifetime of PC1232. It was put back into service after 599 days and a vibration measurement was taken and recorded. A second vibration measurement was taken after 699 days and then again at 791 days when the pump was calendar-suspended. Five more measurements were taken up to 857 days of working life and the pump failed unexpectedly one day later on 858 days. Model 5 of Table 5.19 estimated the pump's residual life to be $662 \leq 815 \leq 894$ after the CM inspection on 599 days. This is also the value reported in Table 5.18. The residual life estimate increases significantly at 791 days but then start to decrease rapidly to $5 \leq 18 \leq 49$ at 829 days. Action should thus

University of Pretoria etd – Vlok, P-J (2006)

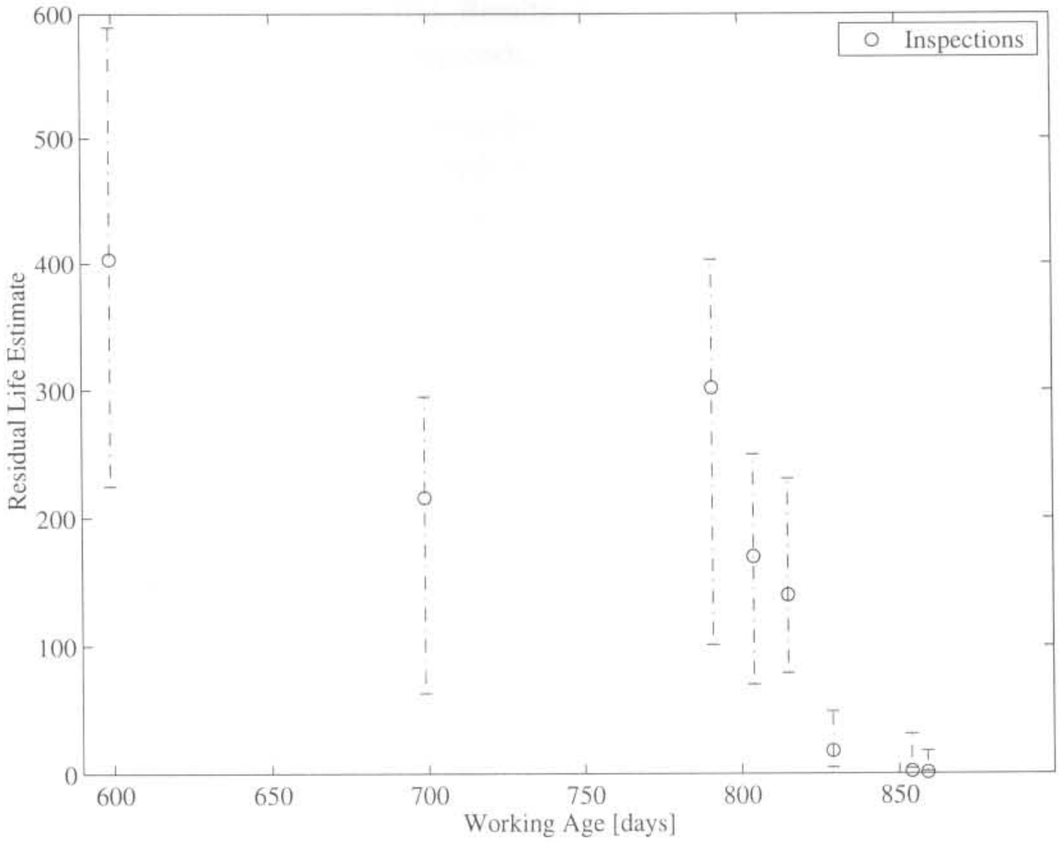

Figure 5.9: RLE approach applied on PC1232

have been taken after 834 days but the pump was left to run to over 850 days. The estimates after 850 days are also consistent with the estimate at 829 days in that action should have been taken immediately.

Figure 5.9 shows that the unexpected failure could have been prevented if the RLE approach with Model 5 of Table 5.19 was followed. In monetary terms, the unexpected failure cost resulted in R 162 200/(858−599) days = R 626.25/day. If the RLE approach was acted upon, R 25 000/(829−599) days = R 108.69/day, would have been the result. If Makis and Jardine's approach was followed, the result would have been R 25 000/(834−599) days = R 106.38/day. This is another confirmation that the model is relevant and practical.

5.6 Conclusion

The objective of Chapter 5 is to test the theory developed in Chapters 3 and 4 on an actual data set obtained from industry and to compare the the results with a similar approach. The only other maintenance decision support technique that uses a PIM as basis, is that of Makis

University of Pretoria etd – Vlok, P-J (2006)

and Jardine (1991) that uses the PHM. Results of the RLE approach were hence compared to results from Makis and Jardine's approach.

A data set was obtained from SASOL Secunda in South Africa. This data set has the typical shortcomings of an industrial data set such as missing observations and irregular inspection intervals. The data set contains a total of 27 histories of which eight are calendar suspensions, eleven are failures and eight are suspensions. Twelve vibration covariates were recorded with each history.

When Vlok (1999) applied the PHM to the data set, only two covariates (RF53H and RF54H which are both related to cavitation) proved to be significant. The final PHM obtained by Vlok (1999) is repeated here as equation (5.17):

$$h(x,\boldsymbol{z}) = \frac{1.464}{1431.8} \cdot \left(\frac{x}{1431.8}\right)^{0.464} \exp\left(0.127 \cdot \mathrm{RF53H} + 0.143 \cdot \mathrm{RF54H}\right) \tag{5.17}$$

Makis and Jardine's approach was used to optimize the plant's vibration monitoring maintenance strategy with (5.17) and the results are briefly repeated in the first part of this chapter. Covariate behaviour was assumed to be stochastic and semi-homogeneous Markov chains were used to predict future covariate behaviour. A cost of unexpected failure of C_f = R 162 200 and a preventive maintenance cost of C_p = R 25 000 were used in Makis and Jardine's policy. If this policy was applied to the actual data set, it would have resulted in a cost of R 214.03 /day which is considerably lower than the observed policy of R 345.16 / day.

In the second part of Chapter 5, the theory developed in this thesis is applied to the data set. Trends of reliability degradation were detected in the interarrival times of the data set and it was hence decided to use repairable systems theory (see Figure 1.3). It was assumed that the same covariates found to be significant in the PHM are significant in the combined PIM of equation (3.30) and covariate behaviour was predicted with parametric functions, thereby assuming covariates to be non-stochastic provided that these covariates were observed up to a certain time t. No formal methodology was followed to obtain the best possible combined PIM. Instead, a totally generic algorithm was developed to fit any combination of enhancements in (3.30) and the best combined PIM was found by trial and error to be:

$$\rho(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t-\tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.18}$$

If (5.18) was used to estimate residual life and action was taken on the lower confidence bound of the prediction, it would have resulted in a cost of R 205.22 / day, which is 4.1% lower than the approach of Makis and Jardine of R 214.03 /day and 40.5% lower than the observed policy. The RLE approach was evaluated further by applying it on a second data set that was collected for pump PC1231. In this evaluation the RLE approached performed

University of Pretoria etd – Vlok, P-J (2006)

slightly worse than the approach of Makis and Jardine but still considerably better that the actual situation.

Chapter 5 proves that the RLE approach is valid and practical. It compares well with the approach of Makis and Jardine and even performed marginally better in certain areas. There are several possible improvements in the RLE approach. These possible improvements are discussed in Chapter 6.

University of Pretoria etd – Vlok, P-J (2006)

Chapter 6

Closure

6.1 Overview

At present, there is a world-wide drive to optimize maintenance decisions in an increasingly competitive manufacturing industry. Preventive maintenance is often the most organized and cost efficient strategy to follow, but a decision still has to be made on the optimal instant to perform preventive maintenance. Use based preventive maintenance decisions have been optimized through statistical analysis of failure data while predictive preventive maintenance (condition monitoring) has been optimized by utilizing more sophisticated technology. Very little work has however been done to combine the advantages of the two schools of thought. This thesis originated from a realization of the potential improvement in maintenance practice by combining use based preventive maintenance optimization techniques with high technology condition monitoring.

A literature survey showed that only one established technique exists to optimize preventive maintenance decisions by considering failure time data and condition monitoring information. That is the approach followed by Makis and Jardine (1991) where the PHM is utilized to describe the failure process and decisions are then made by performing cost trade-offs in terms of risk. Although this technique has a sound theoretical base and has produced many successful results, it is not always well accepted by maintenance practitioners. The technique produces results that are difficult to understand and the underlying model, the PHM, has certain limitations.

Following the literature study, it was decided to pursue an approach that produces results that are much easier to understand, i.e. residual life estimates. The RLE approach developed in this thesis is similar to that of Makis and Jardine in that it also bases estimations on a PIM (the PHM is a special case of a PIM) but the limitations of the PHM are largely overcome. A combined PIM for non-repairable systems and a combined PIM for repairable systems were

University of Pretoria etd – Vlok, P-J (2006)

developed that contains the majority of the enhancements of conventional PIMs in literature as special cases. Any data set under consideration dictates which enhancements are applicable in the combined PIM and the combined PIM can be simplified to fit the attributes of the particular data set.

A data set was obtained from SASOL and both Makis and Jardine's approach and the RLE approach were applied to it. The two techniques produced very similar results with the RLE approach performing marginally better in certain cases.

6.2 Recommendations for future research

Clear objectives for this thesis have been set in Chapter 1 and although these objectives were largely achieved there are still areas where further research can improve the results obtained. A few recommendations for future research are discussed in this section.

6.2.1 Upper and lower bounds on residual life estimates

Upper and lower residual life bounds in Chapter 4 and 5 were calculated directly from the conditional expectation of an event produced by the combined PIM. In doing this it was implicitly assumed that the covariate behaviour was predicted without error. By evaluating the graphs of actual vs. estimated covariate values of Appendix E, it is clear that this implicit assumption is questionable. The influence of the quality of covariate behaviour predictions on residual life estimates is also a function of a a particular combined PIM. Combined PIMs with relatively high regression coefficients would be more sensitive to the quality of covariate behaviour predictions than models with relatively low regression coefficients.

Although the combined PIM used in the Chapter 5 produced relatively good results, the influence of the quality of covariate behaviour predictions is not known and and was not formally taken into account in upper and lower confidence bounds. Further research on this aspect could make residual life estimates more reliable.

6.2.2 Covariate and combined PIM selection

Two of the most difficult steps in a proportional intensity analysis such as this are the selection of appropriate covariates for a particular combined PIM and the selection of the most relevant combined PIM. These steps were addressed as follows in the case study of Chapter 5:

(i) It was assumed that covariates RF53H and RF54H are good descriptors of the failure

University of Pretoria etd – Vlok, P-J (2006)

process of the pumps based on the tests of significance of these covariates in a Weibull PHM done by Vlok (1999). The validity of this assumption was not verified for every model that was evaluated. It was however decided to sustain with these covariates because their physical significance was confirmed by technician experience.

(ii) A trial and error method was used to determine the most applicable combined PIMs for the SASOL data set. This was possible because the generic algorithm that was developed to fit any simplification of the combined PIM in equation (3.30) could easily be adjusted to evaluate numerous different combinations of enhancements. This is one of the biggest advantages of the combined PIM.

Reasonable results were obtained by dealing with the two steps in the manner outlined above but a formal mathematical methodology confirming these selections could benefit future application of the combined PIMs.

6.2.3 Using variable regression coefficients to limit the number of parameters in models

It was on numerous occasions pointed out that in practice reliability data is often very limited. Small data sets are desirable because that indicates that a system is performing well (Ascher (1999)). For the complete combined PIMs in equation (3.13) and (3.30), large data sets are required to be able to estimate parameters with reasonable certainty. Simplifications of the combined PIMs requires much smaller data sets but fewer regression coefficient are always desirable. For this reason regression coefficient elimination techniques should also be implemented.

One method that could be used to reduce the number of parameters in models (especially stratified models), is to define regression coefficients as functions of external influences. This concept is illustrated by the following example. Suppose $k = 1, 2, ..., w$ single part system copies are studied and the position of each copy, $p_k = 1, 2, ..., w$, is expected to influence the survival time of a particular system. Allocate position 1, i.e. p_1, to the system that is least affected by its position and p_w to the system that is most affected by its position. If a stratified combined PIM is used to model event data and the number of parameters should be reduced, regression coefficients can be defined as a function of p_k. For example, if a linear relationship between the regression coefficients and the position exists, $\boldsymbol{\gamma}_s$ can be defined as $\boldsymbol{\gamma}_s(k) = ap_k + b$. In such a case, only a and b need to be determined.

This method is appropriate provided that there is a valid reason to define a certain relationship and in practice such relationships often exist. It is very difficult to formulate a single technique or procedure that would lead to an optimal combined PIM with the minimum number of parameters. Each data set should be modeled on merit.

University of Pretoria etd – Vlok, P-J (2006)

6.3 Conclusion

The RLE approach followed in this thesis originated directly from an industrial need and was constructed in a formal and structured manner after a thorough literature survey. Results obtained from the RLE approach compares favourably with those of an established approach. This verifies that the RLE approach is relevant and practical.

University of Pretoria etd – Vlok, P-J (2006)

Appendix A

Reliability Statistics Preliminaries

A.1 Laplace's trend test

De Laplace (1773) makes use of the fact that under the HPP assumption, the first $m-1$ arrival times, $T_1, T_2, ..., T_{m-1}$ are the order statistics from a uniform distribution on $(0, T_m]$ and hence is,

$$U = \frac{\frac{\sum_{i=1}^{m-1} T_i}{m-1} - \frac{T_m}{2}}{T_m\sqrt{\frac{1}{12(m-1)}}} \tag{A.1}$$

U approximates a standardized normal variate at a 5% level of significance as soon as $m \geq 4$.

In the case where $U \geq 2$ there is strong evidence of reliability degradation while $U \leq -2$ indicates reliability improvement. If $1 \geq U \geq -1$, there is no evidence of an underlying trend and it is referred to as a non-committal data set.

A.2 Renewal theory

A.2.1 Basic concepts

Only IID data sets can be used meaningfully in renewal theory. Data sets of this types are very often, but not necessarily, generated by parts (as defined in Section 1.2.1).

Suppose the interarrival times are part of a distribution $f_X(x)$ with cumulative distribution $F_X(x)$. $F_X(x)$ is referred to as the *unreliability* function since it gives the probability of failure up to a certain age x, i.e. $F_X(x) = \Pr[X \leq x]$. Similarly, is the *reliability* function, $R_X(x)$, defined as $R_X(x) = \Pr[X \geq x]$ or $R_X(x) = 1 - F_X(x)$, i.e. the probability of survival

University of Pretoria etd – Vlok, P-J (2006)

up to age x. From this it is possible to define the force of mortality (FOM) or hazard rate of an item that gives the probability of failure within a short time, provided that the item lived up to that time, i.e. $h_X(x) = \Pr[x < X \leq x + dx | X > x]$. The FOM can also be expressed as,

$$h_X(x) = \frac{f_X(x)}{1 - F_X(x)} \tag{A.2}$$

The FOM is often described as a conditional probability density function. This is not true because,

$$R_X(x) = e^{-\int_0^x h_X(\tau)d\tau} \tag{A.3}$$

and since $R_X(\infty) = 0$ it implies that

$$\lim_{x \to \infty} \int_0^x h_X(\tau)d\tau = \infty \tag{A.4}$$

A.2.2 Distributions

Some distributions often used to model renewal situations are summarized in Table A.1 below.

Table A.1: Distributions often used in renewal theory

Distribution	**Probability Density Function** ($f_X(x)$)	**FOM** ($h_X(x)$)
Exponential $\lambda > 0, x \geq 0$	$\lambda \exp(-\lambda x)$	λ
Weibull $\beta, \eta > 0, x \geq 0$	$\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^{\beta}}$	$\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1}$
Log-normal $\sigma > 0, x \geq 0$	$\frac{\exp\left[-\frac{1}{2}(\ln x - \mu)^2\right]}{\sqrt{2\pi} x \sigma}$	$f_X(x)/\left[1 - \int_0^x f_T(\tau)d\tau\right]$
Log-logistic $\alpha, \lambda > 0, x \geq 0$	$\frac{\alpha x^{\alpha-1}\lambda}{[1+\lambda x^{\alpha}]^2}$	$\frac{\alpha x^{\alpha-1}\lambda}{1+\lambda x^{\alpha}}$
Normal $\sigma > 0, x \geq 0$	$\frac{\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}{\sqrt{2\pi}\sigma}$	$f_X(x)/\left[1 - \int_0^x f_T(\tau)d\tau\right]$

University of Pretoria etd – Vlok, P-J (2006)

A.2.3 Incomplete observations in renewal situations

Very often only partial information is available on an item's survival time. These are referred to as *censored* or *truncated* information. In many cases, this type of information is the only type available in reliability modeling.

A.2.3.1 Censoring

Type I Right Censoring occurs where an event is observed only if it happens prior to some prespecified time. *Progressive Type I Right Censoring* occurs where specimens have different, fixed-sacrifice censoring times, predetermined by the observer. This has the advantage that the sacrificed specimens give information on the natural history of nonlethal events.

Type II Right Censoring occurs where a study continues until the failure of the first r individuals, with $r < n$ and n the total number of individuals. This type of censoring scheme may save time and money if equipment is tested. *Progressive Type II Right Censoring* is a natural extension of *Progressive Type I Right Censoring*.

Left censored observations occur when the event of interest has occurred to the specimen before the period of observation. A good example is a study on the time to first use of marijuana by boys, where the question was asked: "When did you first use marijuana?" and the response "I have used it but I cannot recall just when the time was".

A data set contains *doubly censored* observations where some are left censored and some right censored. If an event is only known to have occurred within a certain interval, the observation is called *interval censored*.

A.2.3.2 Truncation

A *truncation* is defined as a condition where certain subjects are screened so that the investigator is not aware of their existence. If Y is the time of the event which truncates individuals, then, for left-truncated samples, only individuals with $X \geq Y$ are observed. For example, if survival times in an old age home are studied where the age of 60 is a prerequisite ($Y = 60$).

It is also possible to define *right truncations*. This situation is encountered where an event has to occur first before a specimen is included in the sample. A good example is a mortality study on AIDS infected people.

University of Pretoria etd – Vlok, P-J (2006)

A.2.3.3 Contribution of incomplete observations to the likelihood

The maximum likelihood method is most often used to estimate model parameters in survival analysis and it is thus important to note incomplete observations' respective contributions to the likelihood.

Table A.2: The contributions of incomplete observations to the likelihood

Observation type	Contribution to likelihood
Exact lifetimes	$f_X(x)$
Right-censored	$R_X(r_i)$
Left-censored	$1 - R_X(l_i)$
Left-truncations	$f_X(x)/R_X(Y)$
Right-truncations	$f_X(Y)/[1 - R_X(Y)]$
Interval-censored	$[R_X(l_i) - R_X(r_i)]$

In Table A.2, l_i and r_i refer to the left and right margin of an observation interval respectively. Klein and Moeschberger (1990) discuss incomplete information in survival analysis in detail.

A.3 Point Process Theory

A *point process* is a mathematical model that describes a physical phenomenon occurring as highly localized events, distributed randomly in a continuum. In this case, the events are failures and the continuum is time. Brillinger (1978) gives a formal definition.

A.3.1 Basic concepts

Counting process. A counting process, $N(t)$, counts the number of events that have occurred up to time t, where $N(t) \in \mathbb{Z}^+$ and $t \in \mathbb{R}^+$.

Independent increments. A counting process $N(t), t \geq 0$, has independent increments if $N(t_1) - N(0), ..., N(t_k) - N(t_{k-1})$ for $0 < t_1 < ... < t_k$, $k = 2, 3, ..,$ are independent random variables.

Stationary increments. A counting process $N(t), t \geq 0$, has stationary increments if for any two points $t > s \geq 0$ and any $\Delta > 0$, the random variables $(N(t) - N(s))$ and $(N(t + \Delta) - N(s + \Delta))$ are identically distributed.

University of Pretoria etd – Vlok, P-J (2006)

Stationarity of a point process. If a point process has stationary increments, it is said to be stationary.

Intensity. The intensity of a counting process is defined as:

$$\iota(t) \equiv \lim_{\Delta t \to 0} \frac{\Pr\{N(t+\Delta t) - N(t) \geq 1 | H_t\}}{\Delta t} \tag{A.5}$$

where $N(t)$ is the observed number of failures in $(0, t]$ and H_t is the history up to, but not including, time t. Thus, $\iota(t)\Delta t$ is, for a small Δt, the approximate probability of an event in $[t, t + \Delta t)$, given the process history.

When simultaneous failures cannot occur (when the process is orderly) and also stationary, then $\iota(t) = v(t)$, where $v(t)$ is the so called ROCOF, i.e.

$$v(t) = \frac{d}{dt} E\{N(t)\} \tag{A.6}$$

The ROCOF of an NHPP is referred to as the *peril* rate and is denoted by $\rho(t)$.

A.3.2 Homogeneous Poisson Process (HPP)

The HPP is a non-terminating sequence of independent and identically exponentially distributed X_i's. A counting process, $N(t)$, is said to be an HPP if:

(i) $N(0) = 0$

(ii) $\{N(t), t \geq 0\}$ has independent increments, i.e. $N(t_2) - N(t_1) \perp N(t_1)$.

(iii) The number of events in any interval of length $t_2 - t_1$ has a Poisson distribution with mean $\rho(t_2 - t_1)$. This implies that for $t_2 > t_1 \geq 0$,

$$\Pr\{N(t_2) - N(t_1) = j\} = \frac{e^{-\rho(t_2 - t_1)}\{\rho(t_2 - t_1)\}^j}{j!} \tag{A.7}$$

for $j \geq 0$

A.3.3 Non-homogeneous Poisson Process (NHPP)

The NHPP is a non-terminating sequence of independent and identically exponentially distributed X_i's. A counting process, $N(t)$, is said to be an NHPP if:

(i) $N(0) = 0$

(ii) $\{N(t), t \geq 0\}$ has independent increments, i.e. $N(t_2) - N(t_1) \perp N(t_1)$

University of Pretoria etd – Vlok, P-J (2006)

(iii) The number of events in any interval of length $t_2 - t_1$ has a Poisson distribution with mean $\int_{t_1}^{t_2} \rho(t)dt$. This implies that for $t_2 > t_1 \geq 0$,

$$\Pr\{N(t_2) - N(t_1) = j\} = \frac{e^{-\rho(t_2-t_1)}\{\int_{t_1}^{t_2} \rho(t)dt\}^j}{j!} \tag{A.8}$$

for $j \geq 0$.

Bain, Engelhardt, and Wright (1985) proposed some methods to test for the validity of either the NHPP or HPP assumption. Two popular parametric forms for the peril rate of an NHPP are (1) $\rho_1(t) = \alpha e^{\gamma t}$ (log-linear); and (2) $\rho_2(t) = \alpha\gamma t^{\gamma-1}$ (power-law). The latter is often referred to as a Weibull process because the distribution of times to first failure of processes of this kind will be Weibull. To avoid confusion, this term will not be used.

A.3.4 Branching Poisson Process (BPP)

The BPP is discussed in detail in Cox and Lewis (1966) and a summary of their discussion in the present notation is given here. For this process a series of primary events is generated by an HPP and each primary event has positive probability of generating a series of subsidiary events according to a finite renewal process. It is also assumed that the two series of events are indistinguishable. As before, the interarrival times to events (primary or subsidiary) are denoted by X_i, while the interarrival times between primary events are described by Z_i. The interarrival time between a primary and subsidiary event or between two subsidiary events is called Y_i.

Let q be the probability that a primary event triggers a series of a subsidiary events. From this it follows that the expected number of subsidiary events, given that at least 1 subsidiary event occurs, is a/q. Also, if it is assumed that times between subsidiary events will tend to be small relative to Z_i's, it is possible to calculate $\hat{E}[Z]$ with,

$$\hat{E}[Z] = \frac{\sum\limits_{j=1}^{l}(G_j - y)}{l} \tag{A.9}$$

where G_j is the j^{th} excess time over j and l is the total number of observed intervals. (y should be interpreted in the same way as x, defined in Figure 1.2).

A.3.5 Likelihood construction for PMIM applied on Poisson Process data

Define the PMIM as,

$$\iota_u(t, \boldsymbol{z}) = \iota_{u_0}(t) \cdot \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}) \tag{A.10}$$

University of Pretoria etd – Vlok, P-J (2006)

The corresponding cumulative intensity function is:

$$I_u(t, \boldsymbol{z}) = I_{u_0}(t) \cdot \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}) \tag{A.11}$$

where $I_{u_0}(t) = \int_0^t \iota_0(u)\mathrm{d}u$.

Suppose m individuals are under observation. Individual i is observed over the time interval (S_i, T_i) and n_i events are observed at times $t_{i1} < ... < t_{in_i}$. For simplicity suppose S_i is equal to zero. Now, let $\iota_{u_0}(t)$ be specified by parameters in the vector $\boldsymbol{\theta}$. The likelihood function is then,

$$L(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{n_i} \iota_u(t_{ij}) \right\} \exp\{-I_u(T_i)\} \tag{A.12}$$

which can be decomposed as,

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\gamma}) &= \prod_{i=1}^{m} \prod_{j=1}^{n_i} \frac{\iota_{u_0}(t_{ij}; \boldsymbol{\theta})}{I_{u_0}(T_i; \boldsymbol{\theta})} \cdot \prod_{i=1}^{m} \exp\left[-I_{u_0}(T_i; \boldsymbol{\theta}) e^{\boldsymbol{\gamma} \cdot \boldsymbol{z}_i}\right] \left[I_{u_0}(T_i; \boldsymbol{\theta}) e^{\boldsymbol{\gamma} \cdot \boldsymbol{z}_i}\right]^{n_i} \\ &= L_1(\boldsymbol{\theta}) \cdot L_2(\boldsymbol{\theta}, \boldsymbol{\gamma}) \end{aligned} \tag{A.13}$$

The likelihood kernel $L_2(\boldsymbol{\theta}, \boldsymbol{\gamma})$ arises from the Poisson distribution of the counts $n_1, n_2, ..., n_m$, and the kernel $L_1(\boldsymbol{\theta})$ arises from the conditional distribution of the event times, given the counts. Lawless (1987) has shown that if the failure times are not too different, the two kernels can be solved individually to obtain a result fairly close to the full maximum likelihood estimate.

If it assumed that all the T_i's are equal to T, then $L_2(\boldsymbol{\theta}, \boldsymbol{\gamma})$ can be decomposed as,

$$L_2(\boldsymbol{\theta}, \boldsymbol{\gamma}) \propto \Pr(n_1, ..., n_m | \sum_{i=1}^{m} n_i = n) \cdot \Pr(\sum_{i=1}^{m} n_i = n) \tag{A.14}$$

or,

$$L_2(\boldsymbol{\theta}, \boldsymbol{\gamma}) = L_3(\boldsymbol{\gamma}) \cdot L_4(\boldsymbol{\theta}, \boldsymbol{\gamma}) \tag{A.15}$$

where

$$L_3(\boldsymbol{\gamma}) = \prod_{i=1}^{m} \left[\frac{\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_i)}{\sum_{l=1}^{m} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_l)} \right]^{n_i} \tag{A.16}$$

and

$$L_4(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \exp\left[-I_{u_0}(T; \boldsymbol{\theta}) \sum_{l=1}^{m} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_l) \right] \cdot \left[I_{u_0}(T; \boldsymbol{\theta}) \sum_{l=1}^{m} \exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}_l) \right]^{n} \tag{A.17}$$

Williams (1981) indicated that $L_3(\boldsymbol{\gamma})$ is precisely Cox's partial likelihood. Lawless (1987) used data from Gail, Santner, and Brown (1980) to illustrate the convenience of the theory above.

University of Pretoria etd – Vlok, P-J (2006)

Appendix B

Simplifications of the Combined Models

B.1 The non-repairable case

In this section, it is shown that equation (3.8) (repeated below as equation (B.1) for convenience) can be reduced to the majority of models considered in the literature survey on advanced intensity models (Section 2.3).

$$h(x,\boldsymbol{\theta}) = \sum_{l=1}^{n} \varsigma_s^{k_l} \left(g_s^{k_l}(x, \tau_s^{k_l}, \psi_s^{k_l}) \cdot \lambda(\boldsymbol{\gamma}_s^{k_l} \cdot \boldsymbol{z}_s^{k_l}) + \nu(\boldsymbol{\alpha}_s^{k_l} \cdot \boldsymbol{z}_s^{k_l}) \right) \tag{B.1}$$

It is assumed that covariate values are always positive.

B.1.1 Proportional Hazards Model

Restrictions are summarized in Table B.1.

Table B.1: Parameter restrictions for equation (B.1) to obtain a Proportional Hazards Model

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	γ_j, for $j = 1, 2, ..., m$ and all values of s, k and l

University of Pretoria etd – Vlok, P-J (2006)

Equation (B.1) reduces to,

$$h(x,\boldsymbol{\theta}) = g(x) \cdot \lambda(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \tag{B.2}$$

which is similar to (2.8) of Section 2.3.1.1, if λ is chosen to be exponential. For the fully parametric Weibull PHM, $g(x)$ should be substituted with the FOM of a Weibull distribution. To obtain a stratified PHM, s should not be fixed to 1 but, for example, used to denote the previous number of failures, i.e. $s^k = 1$ if $x \leq X_1^k$, $s^k = 2$ if $X_1^k < x \leq X_2^k$, etc. This leads to,

$$h(x,\boldsymbol{\theta}) = g_s(x) \cdot \lambda(\gamma_s \cdot \boldsymbol{z}(x)) \tag{B.3}$$

which was introduced in (2.11).

B.1.2 Proportional Odds Model for Non-repairable Systems

Equation (B.1) should be reduced to $g(x)$ only. Restrictions are summarized in Table B.2.

Table B.2: Parameter restrictions for equation (B.1) to obtain a Proportional Odds Model for non-repairable systems

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1, 2, ..., w$
s:	$s^l = 1$, for all values of i^{k_l}
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions in Table B.2 lead to $h(x,\boldsymbol{\theta}) = g(x)$. To obtain the effect of diminishing covariates, $g(x)$ should be substituted with the FOM of a log-logistic distribution, i.e.

$$h(x;\boldsymbol{\theta}) = \frac{\delta}{x \cdot (1 + x^{-\delta} \cdot \exp(-\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)))} \tag{B.4}$$

as explained in Section 2.3.1.3.

B.1.3 Additive Hazards Model

Restrictions are summarized in Table B.3.

University of Pretoria etd – Vlok, P-J (2006)

Table B.3: Parameter restrictions for equation (B.1) to obtain an Additive Hazards Model

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1, 2, ..., w$
s:	$s^l = 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = \alpha_j$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.1) reduces to,

$$h(x, \boldsymbol{\theta}) = g(x) + \nu(\boldsymbol{\alpha} \cdot \boldsymbol{z}(x)) \tag{B.5}$$

If s is not fixed to 1, the model can be stratified as Pijnenburg (1991) suggested.

B.1.4 PWP Model 2

Restrictions are summarized in Table B.4.

Table B.4: Parameter restrictions for equation (B.1) to obtain a PWP Model 2

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1, 2, ..., w$
s:	$s^l = i^{k_l}$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = \gamma_{s_j}$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.1) reduces to,

$$h(x, \boldsymbol{\theta}) = g_s(x) \cdot \lambda(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}(x)) \tag{B.6}$$

University of Pretoria etd – Vlok, P-J (2006)

which is similar to the PWP Model 2 presented in (2.37). The combined model is not able to reduce to the model proposed by Prentice et al. in (2.42). To have (2.42) as a special case of (3.1), a second stratification variable would be required.

B.1.5 Accelerated Failure Time Model for Non-repairable Systems

Restrictions are summarized in Table B.5.

Table B.5: Parameter restrictions for equation (B.1) to obtain an Accelerated Failure Time Model for non-repairable systems

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	ψ_s^k	$=$	$\phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(x))$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.1) reduces to,

$$h(x, \boldsymbol{\theta}) = g(x \cdot \phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(x)) \tag{B.7}$$

which allows for implementation of (2.44) to (2.47). By lifting the restriction that $\gamma_{s_j}^{k_l} = -\infty$, for all values of s, k and $j \in \{1, 2, ..., m\}$, the Extended Hazard Regression Model of Ciampi and Etezadi-Amoli (1985) and Etezadi-Amoli and Ciampi (1987) can be obtained, i.e.

$$h(x, \boldsymbol{\theta}) = g(x \cdot \phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(x)) \cdot \lambda(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \tag{B.8}$$

as presented in (2.48).

B.1.6 Proportional Age Reduction

Restrictions are summarized in Table B.6.

University of Pretoria etd – Vlok, P-J (2006)

Table B.6: Parameter restrictions for equation (B.1) to obtain an Proportional Age Reduction Model

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1, 2, ..., w$
s:	$s^l = 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = \tau$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.1) reduces to,

$$h(x, \boldsymbol{\theta}) = g(x, \tau) \tag{B.9}$$

The FOM in (B.9) is only a function of x and the factor τ that allows for a jump or setback in time. This model can be used to formulate any PAR model discussed in Section 2.3.3.4.

B.1.7 The model of Lawless and Thiagarajah (1996)

For this model the baseline function g is chosen to be 1. Further restrictions are summarized in Table B.7.

Table B.7: Parameter restrictions for equation (B.1) to obtain an Proportional Age Reduction Model

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1, 2, ..., w$
s:	$s^l = 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\boldsymbol{\gamma}_s^{k} = [\ln \frac{\beta}{\eta^\beta} \quad \beta - 1]$, for $j = 1, 2, ..., m$ and all values of s, k and l
$z_s^{k_l}$:	$\boldsymbol{z}_s^{k_l} = [1 \quad \ln x]$, for all values of s, k and l

University of Pretoria etd – Vlok, P-J (2006)

Equation (B.1) reduces to,

$$
\begin{aligned}
h(x, \boldsymbol{\theta}) &= \frac{\beta}{\eta^{\beta}} x^{\beta-1} \\
&= \frac{\beta}{\eta} \left(\frac{x}{\eta} \right)^{\beta-1}
\end{aligned}
\tag{B.10}
$$

which is a simple Weibull FOM. Following the same argument, it is also possible to obtain the models proposed by Calabria and Pulcini (2000), which are special cases of the model by Lawless and Thiagarajah (1996).

B.2 The repairable case

It is shown in this section that equation (3.25) (repeated below as equation (B.11) for convenience) can be reduced to the majority of models considered in the literature survey on advanced intensity models (Section 2.3).

$$
v(t, \boldsymbol{\theta}) = \sum_{l=1}^{n} \varsigma_s^{k_l} \left(g_s^{k_l}(t, \tau_s^{k_l}, \psi_s^{k_l}) \cdot \lambda(\boldsymbol{\gamma}_s^{k_l} \cdot \boldsymbol{z}_i^{k_l}) + \nu(\boldsymbol{\alpha}_s^{k_l} \cdot \boldsymbol{z}_i^{k_l}) \right) \tag{B.11}
$$

It is assumed that covariate values are always positive.

B.2.1 Proportional Mean Intensity Model

Restrictions are summarized in Table B.8.

Table B.8: Parameter restrictions for equation (B.11) to obtain a Proportional Mean Intensity Model

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	γ_j, for $j = 1, 2, ..., m$ and all values of s, k and l

University of Pretoria etd – Vlok, P-J (2006)

Equation (B.11) reduces to,

$$v(t,\boldsymbol{\theta}) = g(t) \cdot \lambda(\boldsymbol{\gamma} \cdot \boldsymbol{z}(t)) \tag{B.12}$$

which is similar to (2.19) of Section 2.3.1.2, if λ is chosen to be exponential. If the PMIM in (B.12) is parameterized with a log-linear representation of a NHPP, i.e. $g(t)$ is chosen to be log-linear, the model in (2.23) is obtained. Equation (B.12) can also be stratified as described in Section 2.3.1.2.

B.2.2 Proportional Odds Model

No reference was found where the POM was applied on repairable systems, but a similar approach as in Section B.1.2 can be followed where (B.11) is reduced to $g(t)$ only. The restrictions are summarized in Table B.9.

Table B.9: Parameter restrictions for equation (B.11) to obtain a Proportional Odds Model for repairable systems

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Following the argument of B.1.2, the restrictions in Table B.9 lead to $v(t,\boldsymbol{\theta}) = g(t)$. To obtain the effect of diminishing covariates, $g(t)$ could be substituted with a function where if $t \to \infty$, $\gamma \to 0$. One such function is,

$$v(t;\boldsymbol{\theta}) = \frac{\delta}{t \cdot (1 + t^{-\delta} \cdot \exp(-\boldsymbol{\gamma} \cdot \boldsymbol{z}(t)))} \tag{B.13}$$

where δ is a measure of precision as before.

B.2.3 Additive Mean Intensity Model (Additive ROCOF Model)

Restrictions are summarized in Table B.10.

University of Pretoria etd – Vlok, P-J (2006)

Table B.10: Parameter restrictions for equation (B.11) to obtain an Additive Mean Intensity Model (Additive ROCOF Model)

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	α_j, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.11) reduces to,

$$v(t, \boldsymbol{\theta}) = g(t) + \nu(\boldsymbol{\alpha} \cdot \boldsymbol{z}(t)) \tag{B.14}$$

If s is not fixed to 1, the model can be stratified.

B.2.4 PWP Model 1

Restrictions are summarized in Table B.11.

Table B.11: Parameter restrictions for equation (B.11) to obtain a PWP Model 1

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	i^{k_l}, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	γ_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l

University of Pretoria etd – Vlok, P-J (2006)

Equation (B.11) reduces to,

$$v(x, \boldsymbol{\theta}) = g_s(t) \cdot \lambda(\boldsymbol{\gamma}_s \cdot \boldsymbol{z}(t)) \tag{B.15}$$

which is similar to the PWP Model 1 presented in (2.36). The combined model is not able to reduce to the model proposed by Prentice et al. in (2.41). To have (2.41) as a special case of (B.11), a second stratification variable would be required.

B.2.5 Accelerated Failure Time Model for Repairable Systems

Restrictions are summarized in Table B.12.

Table B.12: Parameter restrictions for equation (B.11) to obtain an Accelerated Failure Time Model for repairable systems

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	$1, 2, ..., w$
s:	s^l	$=$	1, for all values of i^{k_l}
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	ψ_s^{k}	$=$	$\phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(x))$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l

Equation (B.11) reduces to,

$$v(t, \boldsymbol{\theta}) = g(t \cdot \phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(t))) \tag{B.16}$$

which allows for implementation of (2.44) to (2.47). By lifting the restriction that $\gamma_{s_j}^{k_l} = -\infty$, for all values of s, k and $j \in \{1, 2, ..., m\}$, the Extended Hazard Regression Model of Ciampi and Etezadi-Amoli (1985) and Etezadi-Amoli and Ciampi (1987) can be obtained, i.e.

$$h(x, \boldsymbol{\theta}) = g(x \cdot \phi(\boldsymbol{\omega} \cdot \boldsymbol{z}(x)) \cdot \lambda(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \tag{B.17}$$

as presented in (2.48).

B.2.6 Proportional Age Reduction Model

Restrictions are summarized in Table B.13.

University of Pretoria etd – Vlok, P-J (2006)

Table B.13: Parameter restrictions for equation (B.11) to obtain an Proportional Age Reduction Model

Parameter	Restriction
n:	$n \;=\; 1$, thus $l = 1$
k:	$k \;=\; 1, 2, ..., w$
s:	$s^l \;=\; 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} \;=\; (1 - \epsilon_k)$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k} \;=\; 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} \;=\; 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} \;=\; -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} \;=\; \gamma_j$, for $j = 1, 2, ..., m$ and all values of s, k and l

In this case $g(t)$ should be selected such that $g(t) = t$. Equation (B.11) reduces to,

$$v(t, \boldsymbol{\theta}) = (1 - \epsilon_k) \cdot \lambda(\boldsymbol{\gamma} \cdot \boldsymbol{z}(t)) \cdot t \tag{B.18}$$

which is similar to the model proposed in (2.56).

University of Pretoria etd – Vlok, P-J (2006)

Appendix C

Numerical Optimization Techniques

C.1 Introduction

Four optimization techniques were implemented successfully to solve the objective functions described in Chapter 3, i.e. converged to the point where all the objective function's partial derivatives were zero, namely:

(i) A Nelder-Mead type simplex search method. (See Buchanan and Turner (1992)).

(ii) A Standard Broyden-Fletcher-Goldfarb-Shanno (BFGS) Quasi-Newton method with a mixed quadratic and cubic line search procedure. (See Wismer and Chattergy (1978)).

(iii) Snyman's dynamic trajectory optimization method. (See Snyman (1982) and Snyman (1983)).

(iv) A modified Newton-Raphson procedure. (See Klein and Moeschberger (1990) and Press, Teukolsky, Vetterling, and Flannery (1993)).

The performance of each one of the methods was measured according to their economy (number of iterations needed before convergence, number of objective function evaluations and number of partial derivative evaluations) and robustness (the accuracy of initial values required for convergence and its ability to handle steep valleys and discontinuities in the objective function). Methods (i) and (ii) maximized the objective functions successfully but performed mediocre. Snyman's method was found to be expensive but extremely robust which is a very valuable attribute. The modified Newton-Raphson method proved to be the most economical and fairly robust as well. For the above mentioned reasons, only Snyman's method and the modified Newton-Raphson method are considered in this discussion on numerical optimization procedures.

University of Pretoria etd – Vlok, P-J (2006)

C.2 Snyman's Dynamic Trajectory Optimization Method

Snyman's method models a conservative force field in m-dimensions (the number of variables in the objective function) with the objective function and then monitors the trajectory of a particle of unit mass (released from rest) as it 'rolls' down the objective function to the point of least potential energy, which is the minimum of the objective function. In this brief presentation of Snyman's technique, the objective function is $l(x, \boldsymbol{\theta})$, the log-likelihood function as presented in (3.15).

The attributes of Snyman's technique can be summarized as follows:

(i) It uses only gradient information, i.e. $\nabla[l(x, \boldsymbol{\theta})]$.

(ii) No explicit line searches are performed.

(iii) It is extremely robust and handles steep valleys and discontinuities in the objective function or gradient with ease.

(iv) This algorithm seeks a low local minimum and it can be used as a basic component in a methodology for global optimization.

(v) The method is not as efficient on smooth and near quadratic functions as classical methods.

The basic dynamic model assumes a particle of unit mass in a m-dimensional conservative force field with potential energy at $\boldsymbol{\theta}$ given by $l(x, \boldsymbol{\theta})$, then the force experienced by the particle at $\boldsymbol{\theta}$ is given by $m\boldsymbol{a} = \ddot{\boldsymbol{\theta}} = -\nabla[l(x, \boldsymbol{\theta})]$. From this it follows that for the time interval $[0, x]$,

$$\frac{1}{2}\left\|\dot{\boldsymbol{\theta}}|_{x=x}\right\|^2 - \frac{1}{2}\left\|\dot{\boldsymbol{\theta}}|_{x=0}\right\|^2 = l(0, \boldsymbol{\theta}) - l(x, \boldsymbol{\theta}) \quad \text{(C.1)}$$

Equation (C.1) can be simplified by expressing it in terms of kinetic energy, T, as $T(x) - T(0) = l(0, \boldsymbol{\theta}) - l(x, \boldsymbol{\theta})$. It is clear that $l(x, \boldsymbol{\theta}) + T(x)$ is constant, which indicates conservation of energy in the conservative force field. It should also be noted that $\Delta l = -\Delta T$, therefore as long as T increases, l decreases, which is the basis of the dynamic algorithm.

Suppose $l(x, \boldsymbol{\theta})$ has to be minimized from a starting point $\boldsymbol{\theta}|_{x=0} = \boldsymbol{\theta}_0$, then the dynamic algorithm is as follows:

(i) Compute the dynamic trajectory by solving the initial value problem, $\ddot{\boldsymbol{\theta}}|_{x=x} = -\nabla[l(x, \boldsymbol{\theta})]$, $\dot{\boldsymbol{\theta}}|_{x=0} = 0$ and $\boldsymbol{\theta}|_{x=0} = \boldsymbol{\theta}_0$. In practice the numerical integration of the initial value problem is often done by the "leap-frog" method. Compute for $k = 1, 2, ..$ and time step Δx, the following: $\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^k + \dot{\boldsymbol{\theta}}^k \Delta x$ and $\dot{\boldsymbol{\theta}}^{k+1} = \dot{\boldsymbol{\theta}}^k + \ddot{\boldsymbol{\theta}}^k \Delta x$, where $\ddot{\boldsymbol{\theta}}^k = -\nabla[l(x, \boldsymbol{\theta}^k)]$ and $\dot{\boldsymbol{\theta}}_0 = {}^1\!/\!_2 \ddot{\boldsymbol{\theta}}_0 \Delta x$.

University of Pretoria etd – Vlok, P-J (2006)

(ii) Monitor $\dot{\boldsymbol{\theta}}|_{x=x}$, the velocity of the particle. As long as the kinetic energy $T = {}^1\!/_2 \left\|\dot{\boldsymbol{\theta}}|_{x=x}\right\|$ increases, the potential energy decreases, i.e. $l(x, \boldsymbol{\theta})$ decreases.

(iii) As soon as T decreases, the particle is moving uphill and the objective function is increasing, i.e. $||\dot{\theta}^{k+1}|| \leqslant ||\dot{\theta}^{k}||$. Some interfering strategy should be applied to extract energy from the particle to increase the likelihood of decent. A typical interfering strategy is to let $\dot{\boldsymbol{\theta}}^k = {}^1\!/_4(\dot{\boldsymbol{\theta}}^{k+1} + \dot{\boldsymbol{\theta}}^k)$ and $\boldsymbol{\theta}^{k+1} = {}^1\!/_2(\boldsymbol{\theta}^{k+1} + \boldsymbol{\theta}^k)$ after which a new $\boldsymbol{\theta}^{k+1}$ is calculated and the algorithm is continued.

(iv) To accelerate convergence of the method, the algorithm should allow for magnification and reduction of the step size, Δx , depending on the particle's position.

The method is extremely robust and particularly useful when variables in the objective function is totally unknown.

C.3 Modified Newton-Raphson Optimization Method

The objective of the numerical procedure is to find the value of $\boldsymbol{\theta}$ where all the partial derivatives of $l(x, \boldsymbol{\theta})$ are zero. Suppose $(F(x))$ and $(G(x))$ are matrices containing the first and second partial derivatives of $l(x, \boldsymbol{\theta})$, respectively. An approximation often used for $(F(x))$ is $(F(\boldsymbol{\theta})) \approx (F(\boldsymbol{\theta})) + (G(\boldsymbol{\theta}_0)) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ where $\boldsymbol{\theta}_0$ is an initial estimate. It is required to solve $(F(\boldsymbol{\theta}_0)) + (G(\boldsymbol{\theta}_0)) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0) = 0$ to determine the optimal value of $\boldsymbol{\theta}$.

The conventional Newton-Raphson procedure would solve for $\boldsymbol{\theta}$ as follows:

(i) Estimate a meaningful initial value for $\boldsymbol{\theta}_0$, i.e. $\boldsymbol{\theta}$.

(ii) Calculate $(F(x))$ and $(G(x))$.

(iii) Solve for Δ_0 in the system $(G(\boldsymbol{\theta}_0))\Delta_0 = -(F(\boldsymbol{\theta}_0))$.

(iv) Set $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_0 + \Delta_0$ and repeat the procedure until convergence.

Instead of the conventional Newton-Raphson method, a variable metric method (quasi-Newton method) can be used to overcome some numerical difficulties. In this modified Newton-Raphson method, $(G(x))$ is not calculated directly but an approximation of $(G(x))$ is used that is chosen to be always positive definite, thereby eliminating the possibility of singular matrices. The approximation of $(G(x))$ is explained in detail in Press, Teukolsky, Vetterling, and Flannery (1993). Press et al. also describe methods to vary step sizes in the procedure as well as stopping rule procedures. Vlok (1999) discusses methods to accelerate convergence and increase the accuracy of the procedure by transforming the data before iterations start.

University of Pretoria etd – Vlok, P-J (2006)

Appendix D

SASOL Data

D.1 Inspection data for Bearing 3

The inspection data for Bearing 3 is presented in Table D.1 on the next page, where the columns have the following meanings:

Pump ID: Pump identification number.

Age: Global age of the pump measured in days.

Date: Actual date of inspection.

A: RF043H, i.e 0.4× rotational frequency amplitude, measured on horizontally on Bearing 3 in mm/s, indicative of a bearing defect.

B: RF13H, i.e. 1× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of unbalance in the pump.

C: RF23H, i.e. 2× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of misalignment in the pump.

D: RF53H, i.e. 5× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of cavitation in the pump.

E: HFD3H, i.e. high frequency domain components between 1200-2400 Hz, measured on Bearing 3, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.

F: LNF3H, i.e. lifted noise floor in 600-1200 Hz range, measured on Bearing 3, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.

University of Pretoria etd – Vlok, P-J (2006)

Table D.1: Inspection data for Bearing 3

Pump ID	Age (Days)	Date	A [mm/s]	B [mm/s]	C [mm/s]	D [mm/s]	E [0/1]	F [0/1]
PC1131	159	02/07/97	0	0.7	0.3	0.8	1	0
PC1131	295	06/23/97	0.15	0.3	0.25	0.55	0	1
PC1131	387	09/23/97	0.3	3	0.9	8	1	0
PC1131	394	09/30/97	0.8	2.4	1	12.3	1	0
PC1131	397	10/03/97	250	175	20	17	1	0
PC1131	530	02/13/98	0.1	11.5	3.2	11	0	0
PC1131	533	02/16/98	0.3	8.8	3.5	13	1	0
PC1131	554	03/09/98	0.5	7	3.8	16	0	0
PC1131	578	04/02/98	1	19.5	1.5	2	1	0
PC1131	597	04/21/98	0.3	27.5	1.5	1.6	1	0
PC1131	639	06/02/98	0.5	31	6	4	1	0
PC1131	689	07/22/98	0	9	2	0.8	0	0
PC1131	690	07/23/98	0	8.27	1.82	0.67	0	0
PC1131	703	08/05/98	0.05	1.2	0.95	0.2	1	0
PC1131	712	08/14/98	0.05	0.5	0.8	1.4	1	0
PC1131	765	10/06/98	0.05	0.4	0.7	2.7	1	0
PC1131	791	11/01/98	0.5	9	2	12	0	0
PC1132	239	04/28/97	0	0.9	0.3	1.5	0	0
PC1132	386	09/22/97	0.1	7	0.6	2.1	1	0
PC1132	394	09/30/97	0.2	8	0.5	11	1	0
PC1132	397	10/03/97	0.1	6.2	0.2	3	0	0
PC1132	491	01/05/98	0.1	5	0.5	1	0	0
PC1132	499	01/13/98	0.1	27.5	2	2.5	0	0
PC1132	533	02/16/98	0.1	35	2.5	12	0	0
PC1132	543	02/26/98	5	19	26	9	0	0
PC1132	544	02/27/98	5.61	16.94	28.93	8.56	0	0
PC1132	557	03/12/98	3	43	9	2	0	0
PC1132	558	03/13/98	1	41	14	3	0	0
PC1132	597	04/21/98	4	29	3.7	2.6	0	1
PC1132	689	07/22/98	0.1	5.6	1.7	0.3	0	1
PC1132	712	08/14/98	0.1	3.4	0.6	0.9	0	1
PC1132	751	09/22/98	0.99	3.01	0.3	2.99	0	1
PC1132	791	11/01/98	0.08	4.65	0.17	2.01	0	0
PC1231	239	04/28/97	0.3	5.5	1.9	1	0	0
PC1231	295	06/23/97	1.3	10.4	2.2	1	0	0
PC1231	390	09/26/97	1	56	12	3	0	0

University of Pretoria etd – Vlok, P-J (2006)

PC1231	530	02/13/98	0.3	18.1	6.1	8.5	1	0
PC1231	563	03/18/98	0.09	12	1.18	10.24	1	0
PC1231	578	04/02/98	1	33	18	6	1	1
PC1231	653	06/16/98	0.22	3.57	0.98	0.57	0	0
PC1231	698	07/31/98	0.68	8.11	1.47	0.61	0	0
PC1231	791	11/01/98	0.73	38.64	7.68	1.86	0	0
PC1232	583	04/07/98	0.5	56	9	4	0	0
PC1232	592	04/16/98	0.4	54	4	6.5	0	0
PC1232	597	04/21/98	0.6	48	9	3.5	0	0
PC1232	599	04/23/98	0.05	7	2.1	0.6	1	1
PC1232	699	08/01/98	0.33	34.16	5.76	2.48	0	0
PC1232	791	11/01/98	0.24	32.4	2.44	4.09	0	0
PC2131	156	02/04/97	0	9	1.2	0.4	0	0
PC2131	159	02/07/97	0.1	5.8	2.2	0.6	0	1
PC2131	178	02/26/97	0.2	4	3.3	1.35	0	1
PC2131	179	02/27/97	0	8.3	2	0.9	0	0
PC2131	184	03/04/97	0	36.39	2	1	0	1
PC2131	239	04/28/97	0.09	3.65	1.6	1.55	1	0
PC2131	241	04/30/97	0.05	3.1	0.75	1.7	1	0
PC2131	295	06/23/97	0.1	2.55	2.2	1.4	1	0
PC2131	386	09/22/97	0.4	5.6	7.5	0.7	1	0
PC2131	470	12/15/97	1200	120	30	10	0	0
PC2131	535	02/18/98	0.2	20.9	1.6	4.8	0	0
PC2131	583	04/07/98	2	77	46	11	0	0
PC2131	597	04/21/98	2	66	43	6	0	0
PC2131	604	04/28/98	1	74	37.5	5	1	0
PC2131	611	05/05/98	0.01	20	4.1	11.6	1	0
PC2131	631	05/25/98	0.1	18	10	72.33	1	0
PC2131	640	06/03/98	0.6	10.5	2.8	5.9	1	0
PC2131	689	07/22/98	0.09	1.7	0.4	0.5	1	0
PC2131	768	10/09/98	0.1	1.92	0.55	0.66	1	0
PC2131	774	10/15/98	0.14	2.66	0.76	1.12	1	0
PC2131	791	11/01/98	0.16	13.37	1.08	3.69	0	0
PC3131	241	04/30/97	0.1	6.8	3.9	1.3	1	0
PC3131	295	06/23/97	0.8	29	17	14	1	0
PC3131	386	09/22/97	0.5	37	6.5	4	1	0
PC3131	450	11/25/97	0.2	20.52	6	3	1	0
PC3131	550	03/05/98	0.09	7.2	3.74	1.27	1	0
PC3131	651	06/14/98	0.96	33.06	17.34	16.8	1	0

University of Pretoria etd – Vlok, P-J (2006)

PC3131	750	09/21/98	0.59	40.33	6.43	4.16	1	0
PC3131	791	11/01/98	0.2	19.48	5.82	3.39	1	0
PC3132	239	04/28/97	0.1	2.4	0.15	0.39	1	0
PC3132	295	06/23/97	0.2	9.6	1.8	1.6	1	1
PC3132	386	09/22/97	0.2	24	3	3.5	1	1
PC3132	450	11/25/97	0.5	32	21	13	0	0
PC3132	506	01/20/98	0.97	37.56	48.37	26.84	0	0
PC3132	566	03/21/98	0.12	2.44	0.16	0.45	1	1
PC3132	711	08/13/98	0.19	11.04	1.92	1.82	1	1
PC3132	791	11/01/98	0.2	27.6	3.27	3.39	1	1
PC3232	239	04/28/97	0.3	11.5	3.8	0.6	1	0
PC3232	295	06/23/97	1	43	8	6	1	0
PC3232	386	09/22/97	2	39	6	6	1	0
PC3232	535	02/18/98	0	66	44	7	0	0
PC3232	563	03/18/98	0	75.72	56.86	7.33	1	0
PC3232	591	04/15/98	0	235	22	10	0	0
PC3232	604	04/28/98	2	175	18	7	0	0
PC3232	639	06/02/98	3	74	9	3	0	0
PC3232	722	08/24/98	0	20.5	14.8	1.9	1	1
PC3232	723	08/25/98	0	21.45	15.1	1.96	1	1
PC3232	748	09/19/98	0.18	7.59	2.96	0.39	1	0
PC3232	783	10/24/98	0.62	26.66	5.44	4.5	1	0
PC3232	791	11/01/98	1.28	28.08	3.72	4.08	1	0

In Chapter 5 it is shown that RF53H is a good predictor of failure and plays an significant role in the maximum likelihood. For the sake of completeness, the data for this covariate is also displayed graphically in Figures D.1 to D.8 for each lifetime.

University of Pretoria etd – Vlok, P-J (2006)

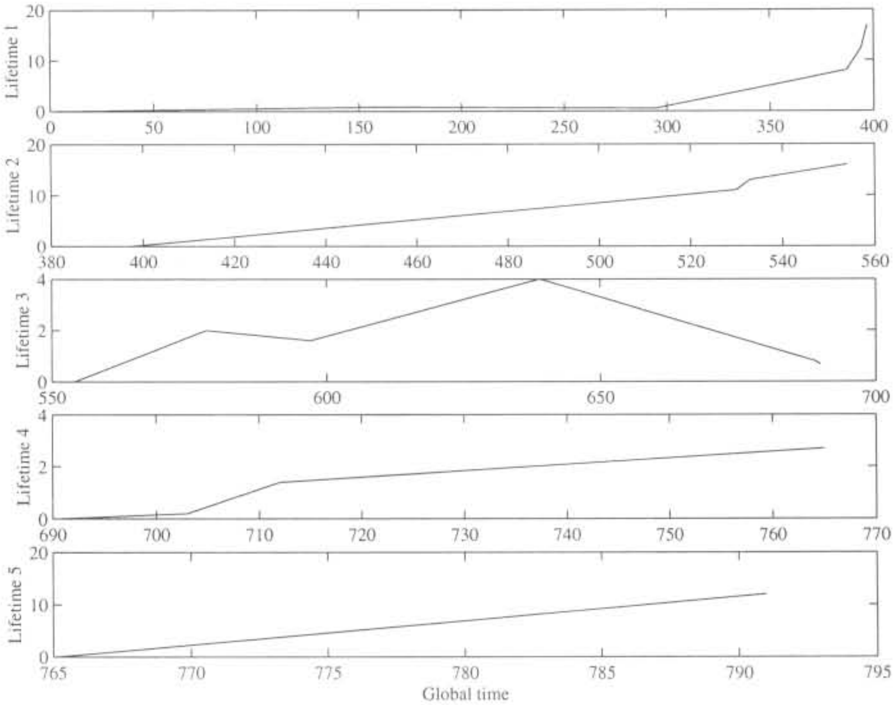

Figure D.1: Observed values of RF53H for PC1131

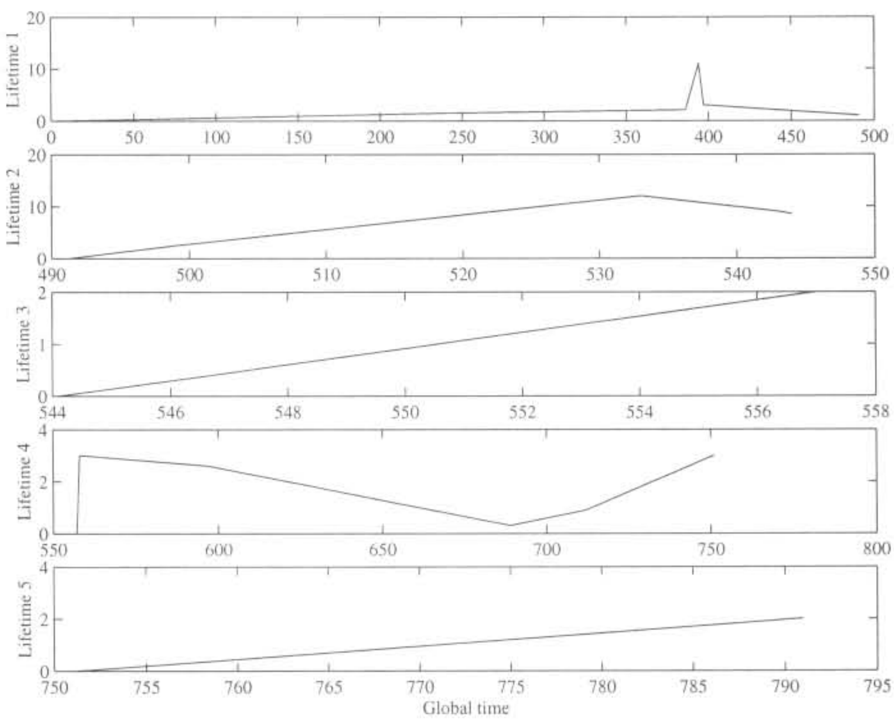

Figure D.2: Observed values of RF53H for PC1132

University of Pretoria etd – Vlok, P-J (2006)

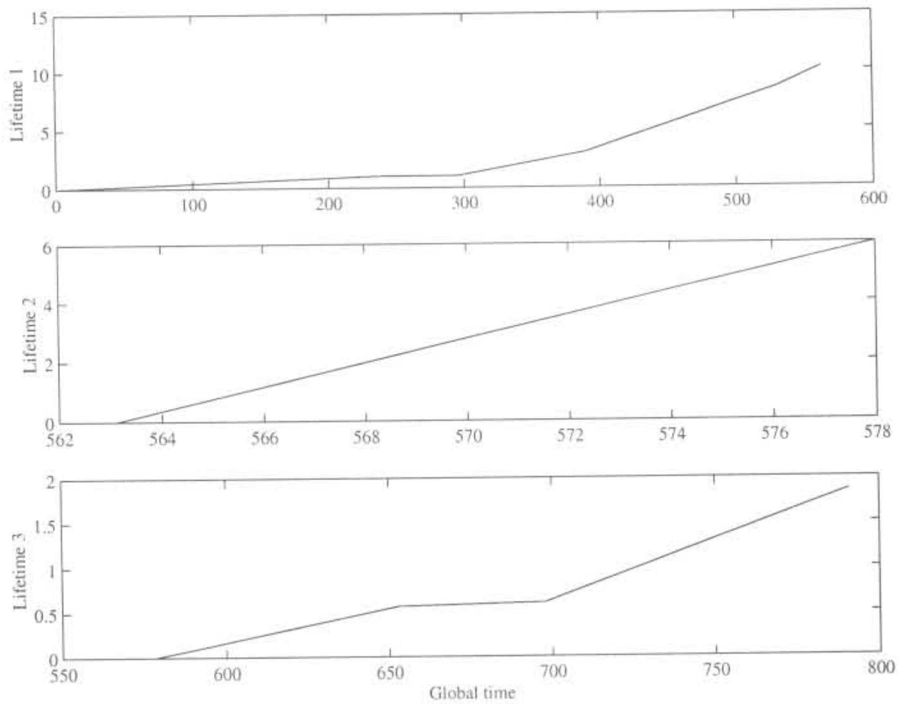

Figure D.3: Observed values of RF53H for PC1231

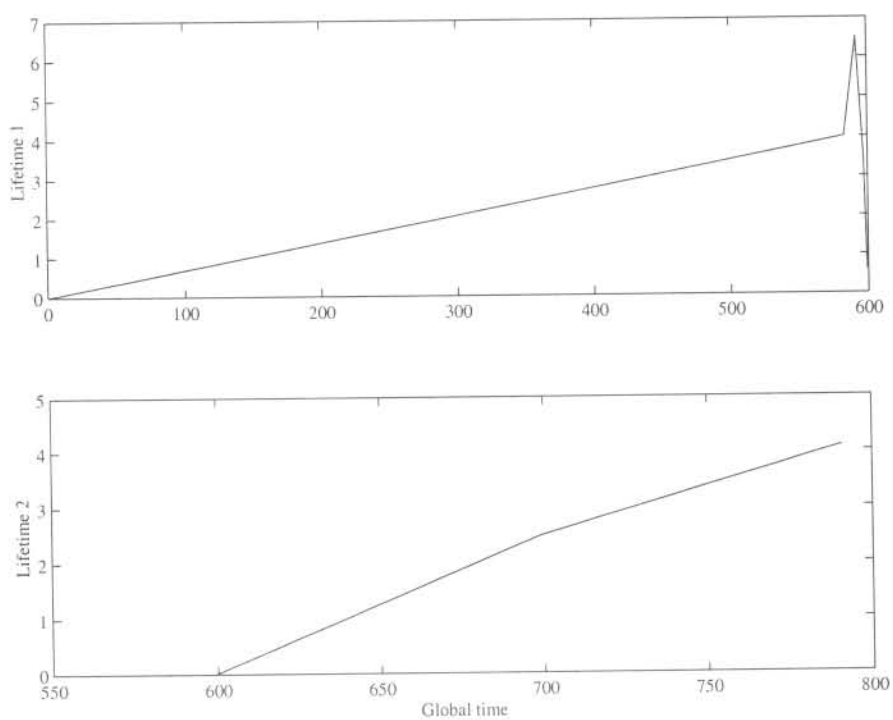

Figure D.4: Observed values of RF53H for PC1232

University of Pretoria etd – Vlok, P-J (2006)

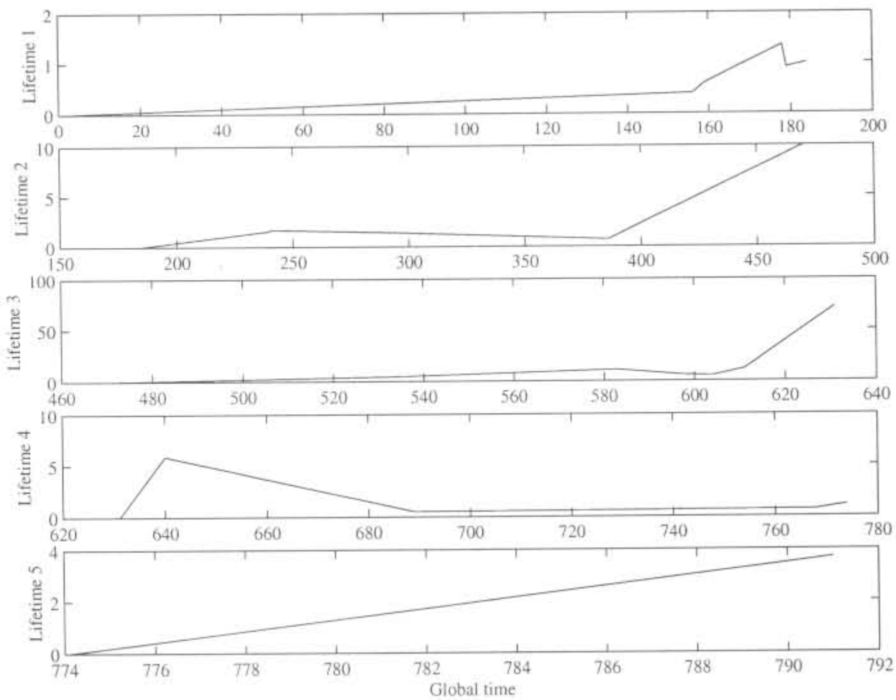

Figure D.5: Observed values of RF53H for PC2131

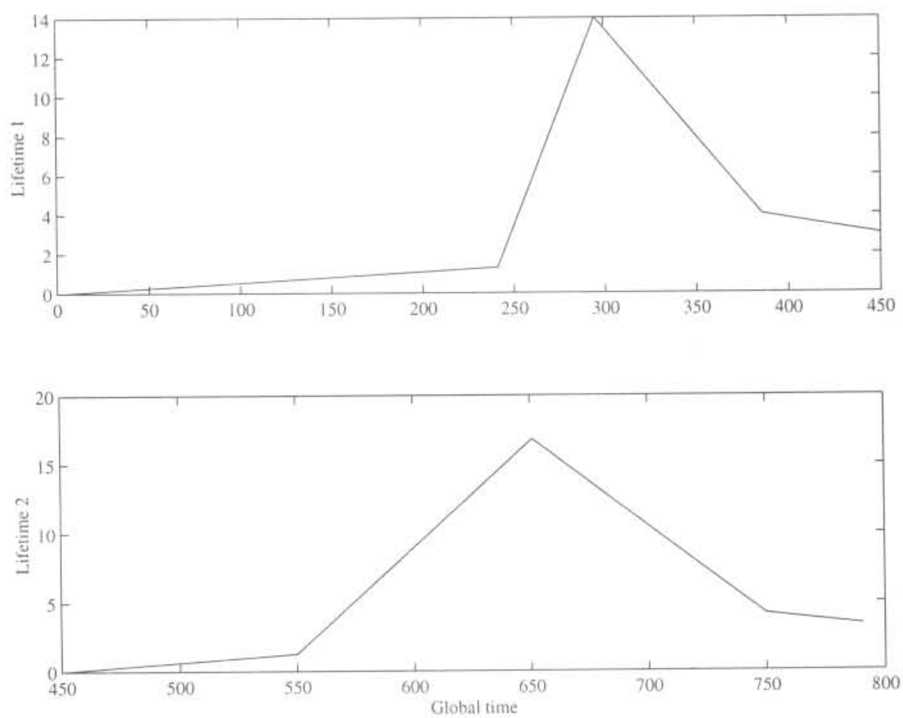

Figure D.6: Observed values of RF53H for PC3131

University of Pretoria etd – Vlok, P-J (2006)

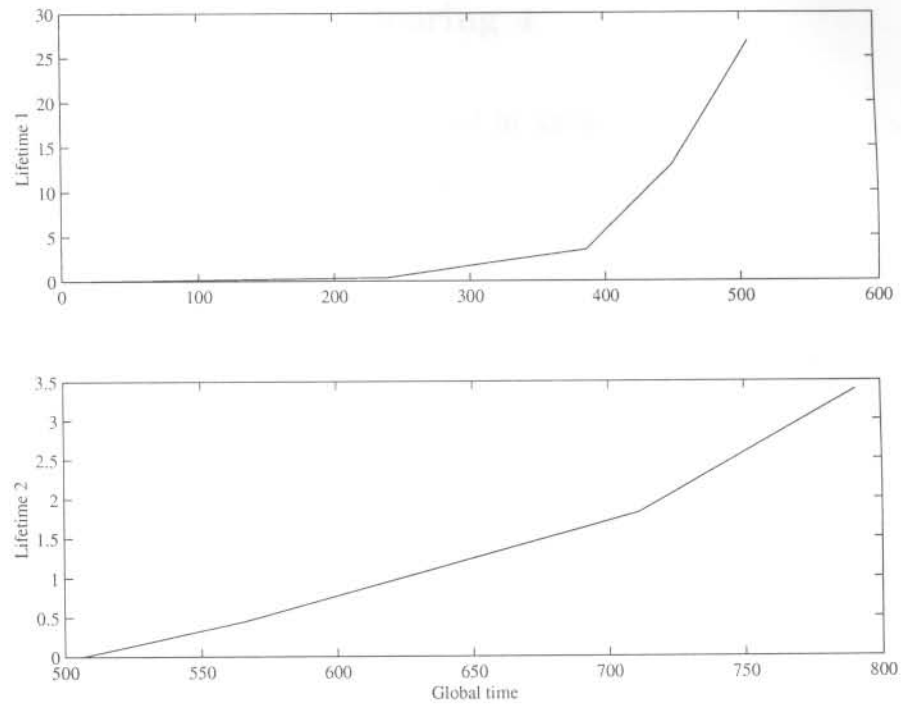

Figure D.7: Observed values of RF53H for PC3132

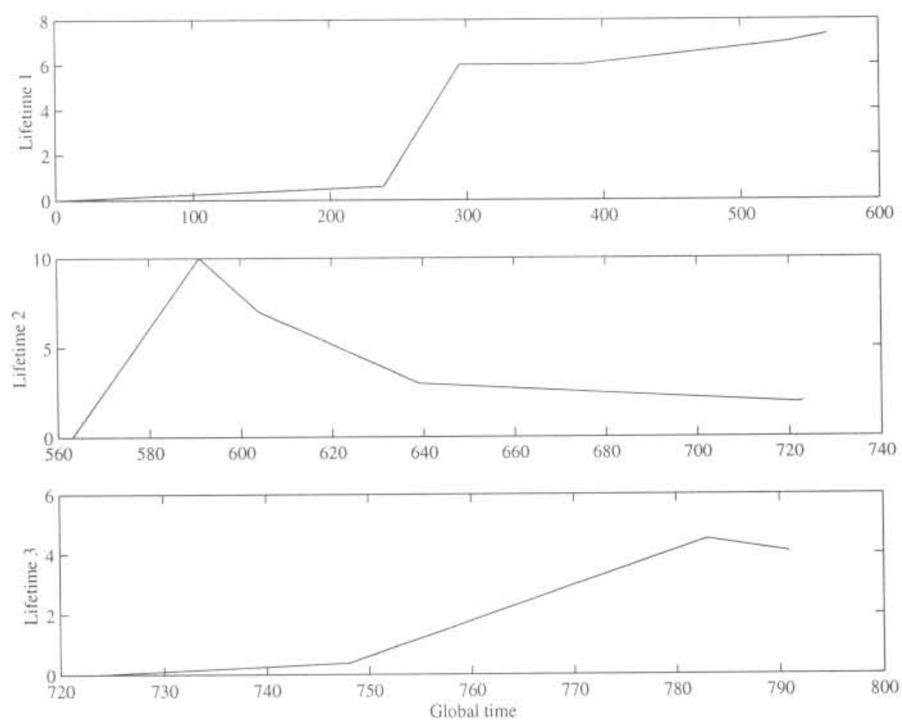

Figure D.8: Observed values of RF53H for PC3232

University of Pretoria etd – Vlok, P-J (2006)

D.2 Inspection data for Bearing 4

The inspection data for Bearing 4 is presented in Table D.2 on the next page, where the columns have the following meanings:

Pump ID: Pump identification number.
Age: Global age of the pump measured in days.
Date: Actual date of inspection.
A: RF044H, i.e 0.4× rotational frequency amplitude, measured on horizontally on Bearing 4 in mm/s, indicative of a bearing defect.

B: RF14H, i.e. 1× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of unbalance in the pump.
C: RF24H, i.e. 2× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of misalignment in the pump.
D: RF54H, i.e. 5× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of cavitation in the pump.
E: HFD4H, i.e. high frequency domain components between 1200-2400 Hz, measured on Bearing 4, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.
F: LNF4H, i.e. lifted noise floor in 600-1200 Hz range, measured on Bearing 4, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.

Table D.2: Inspection data for Bearing 4

Pump ID	Age (Days)	Date	A [mm/s]	B [mm/s]	C [mm/s]	D [mm/s]	E [0/1]	F [0/1]
PC1131	159	02/07/97	0.05	0.85	0.3	0.1	1	0
PC1131	295	06/23/97	0.2	0.45	0.25	0.12	0	1
PC1131	387	09/23/97	0.1	4	1.7	6.2	1	0
PC1131	394	09/30/97	2.3	4	2.1	5	0	0
PC1131	397	10/03/97	4	4.6	2.8	6	1	0
PC1131	530	02/13/98	0.1	13.2	3.5	5.5	0	0
PC1131	533	02/16/98	0.2	10	3.8	7	1	0
PC1131	554	03/09/98	0.3	5	4.2	10	0	0
PC1131	578	04/02/98	0.7	42	3	3	1	0
PC1131	597	04/21/98	0.5	52	2	5	1	0
PC1131	639	06/02/98	0.5	47	8	5	1	0
PC1131	689	07/22/98	0	14	2	1.2	0	0
PC1131	690	07/23/98	0	13.04	1.73	1.08	0	0

University of Pretoria etd – Vlok, P-J (2006)

PC1131	703	08/05/98	0.2	2.25	0.9	0.4	1	0
PC1131	712	08/14/98	0.05	0.58	1.3	0.41	1	1
PC1131	765	10/06/98	0.05	0.4	2.1	0.6	1	1
PC1131	791	11/01/98	0.2	12	2	7	0	0
PC1132	239	04/28/97	0	1.65	0.3	0.72	0	1
PC1132	386	09/22/97	0.1	12.2	0.7	7.8	1	0
PC1132	394	09/30/97	0.1	14	0.9	8.2	1	0
PC1132	397	10/03/97	0.2	12	0.9	12	1	0
PC1132	491	01/05/98	1	10	0.8	30	1	0
PC1132	499	01/13/98	0.1	66	4	12	0	0
PC1132	533	02/16/98	0	65	3	10	0	0
PC1132	543	02/26/98	1	120	38	7	0	0
PC1132	544	02/27/98	1.13	126.88	42.38	6.64	0	0
PC1132	557	03/12/98	1	34	5	2.5	1	0
PC1132	558	03/13/98	2	27.5	6.5	1	0	0
PC1132	597	04/21/98	1	24	4.2	5.4	0	1
PC1132	689	07/22/98	0.1	4.8	0.7	0.4	0	0
PC1132	712	08/14/98	0.05	2.7	0.3	0.4	0	0
PC1132	751	09/22/98	0.13	1.61	0.06	1.54	0	1
PC1132	791	11/01/98	0.15	7.8	0.56	7.68	1	0
PC1231	239	04/28/97	0	9	0.6	0.4	0	0
PC1231	295	06/23/97	0.3	16.5	2.3	0.3	0	0
PC1231	390	09/26/97	0	67	6	4	0	0
PC1231	530	02/13/98	0	21	6	6	1	1
PC1231	563	03/18/98	0.08	10	5.05	5.87	1	1
PC1231	578	04/02/98	2	51	16	9	1	1
PC1231	653	06/16/98	0	6.75	0.41	0.27	0	0
PC1231	698	07/31/98	0.22	10.72	1.35	0.15	0	0
PC1231	791	11/01/98	0	46.9	4.14	2.64	0	0
PC1232	583	04/07/98	0	71	8	3	0	0
PC1232	592	04/16/98	0.05	53	3	2	0	0
PC1232	597	04/21/98	1	57	6	3	0	0
PC1232	599	04/23/98	0.15	7.9	3.5	0.9	0	1
PC1232	699	08/01/98	0	49.7	5.28	1.92	0	0
PC1232	791	11/01/98	0.03	36.57	2.04	1.24	0	0
PC2131	156	02/04/97	0	15.5	2.1	0.5	0	1
PC2131	159	02/07/97	0	7	1.8	0.4	0	1
PC2131	178	02/26/97	0.05	6.7	2.3	0.4	0	0
PC2131	179	02/27/97	0	12.2	2.2	0.4	0	0

University of Pretoria etd – Vlok, P-J (2006)

PC2131	184	03/04/97	0	47.97	1.51	0.4	0	1
PC2131	239	04/28/97	0.05	9.6	1.1	0.7	0	0
PC2131	241	04/30/97	0.1	8.1	1	0.7	1	0
PC2131	295	06/23/97	0.2	6.1	1.5	0.4	1	0
PC2131	386	09/22/97	1.7	21	1.4	3.7	1	0
PC2131	470	12/15/97	78	48	12	9	0	0
PC2131	535	02/18/98	0.5	27	7.4	7	0	0
PC2131	583	04/07/98	2	62	39	6	0	0
PC2131	597	04/21/98	2	64	38	4	0	0
PC2131	604	04/28/98	2	61	37	5	1	0
PC2131	611	05/05/98	0.01	24	6	1.4	1	0
PC2131	631	05/25/98	0.01	10	10	1	1	0
PC2131	640	06/03/98	0.2	26	1	4	1	0
PC2131	689	07/22/98	0.05	4.6	0.25	0.33	1	0
PC2131	768	10/09/98	0.05	4.2	0.3	0.2	1	0
PC2131	774	10/15/98	0.06	5.89	0.37	0.48	1	0
PC2131	791	11/01/98	0.34	17.55	4.66	5.6	0	0
PC3131	241	04/30/97	0.1	8	1.7	1	1	0
PC3131	295	06/23/97	0.7	35	10	7	1	0
PC3131	386	09/22/97	2	33	5	7	1	0
PC3131	450	11/25/97	3.13	20	4	2	1	0
PC3131	550	03/05/98	0.1	8.08	1.81	1.2	1	0
PC3131	651	06/14/98	0.71	39.2	9.8	7.7	1	0
PC3131	750	09/21/98	2.4	36.3	4.9	6.58	1	0
PC3131	791	11/01/98	3.47	21.4	4.08	1.8	1	0
PC3132	239	04/28/97	0.2	3.6	0.25	0.55	1	0
PC3132	295	06/23/97	0.3	12.2	0.9	2.2	1	1
PC3132	386	09/22/97	0.05	35	2.5	2.4	1	1
PC3132	450	11/25/97	0	81	8	6.5	0	0
PC3132	506	01/20/98	0.04	141.55	15.78	12.77	0	0
PC3132	566	03/21/98	0.23	4.32	0.25	0.59	1	0
PC3132	711	08/13/98	0.37	15.61	1.06	2.35	1	1
PC3132	791	11/01/98	0.06	39.9	3.25	2.61	1	1
PC3232	239	04/28/97	0.01	16	2.3	0.3	1	0
PC3232	295	06/23/97	1	48	9	4	1	0
PC3232	386	09/22/97	1	52	4	3	1	0
PC3232	535	02/18/98	0	91	26	8	0	0
PC3232	563	03/18/98	0	102.83	34.32	9.86	0	0
PC3232	591	04/15/98	0	280	10	15	0	0

University of Pretoria etd – Vlok, P-J (2006)

PC3232	604	04/28/98	0	150	9	8	0	0
PC3232	639	06/02/98	5	73	6	6	0	0
PC3232	722	08/24/98	0	27	10	0.8	0	0
PC3232	723	08/25/98	0	27.62	10.14	0.73	0	0
PC3232	748	09/19/98	0	12	1.84	0.23	1	0
PC3232	783	10/24/98	0.73	30.72	5.85	3.2	1	0
PC3232	791	11/01/98	0.72	31.2	2.96	1.95	1	0

In Chapter 5 it is shown that RF54H is a good predictor of failure and plays an significant role in the maximum likelihood. For the sake of completeness, the data for this covariate is also displayed graphically in Figures D.9 to D.16 for each lifetime.

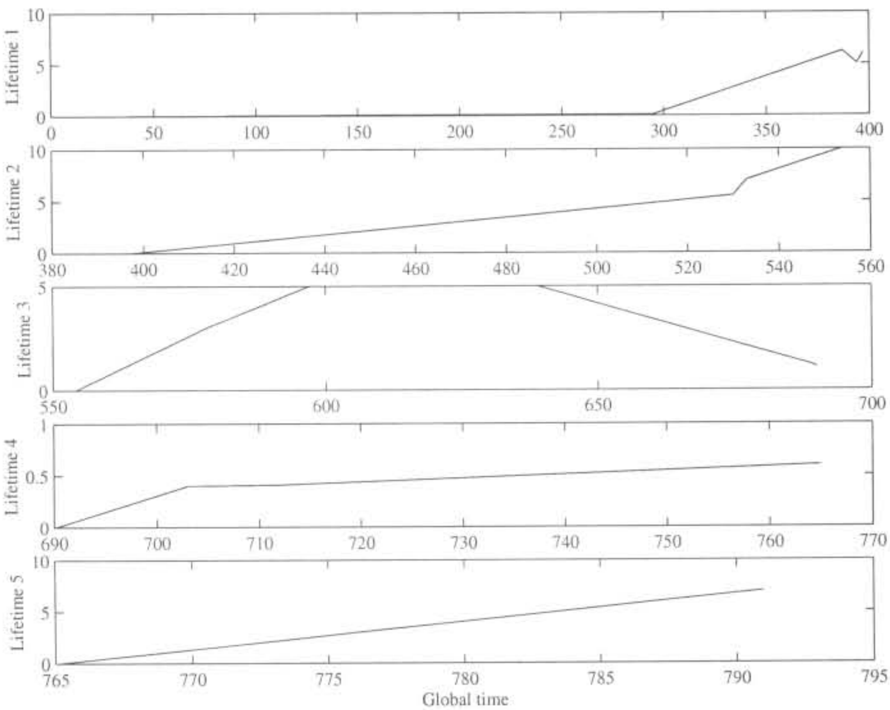

Figure D.9: Observed values of RF54H for PC1131

University of Pretoria etd – Vlok, P-J (2006)

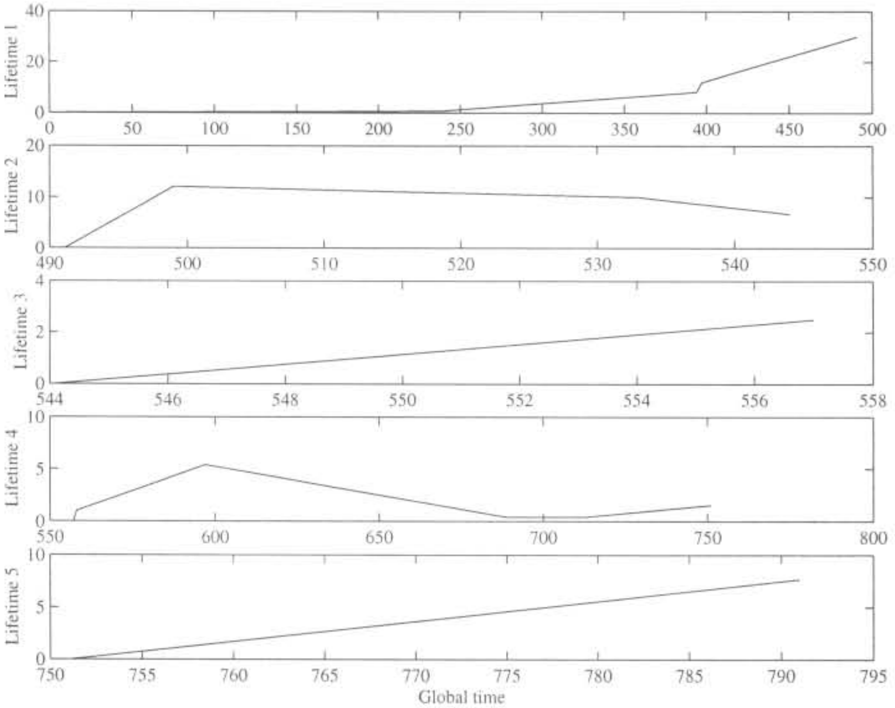

Figure D.10: Observed values of RF54H for PC1132

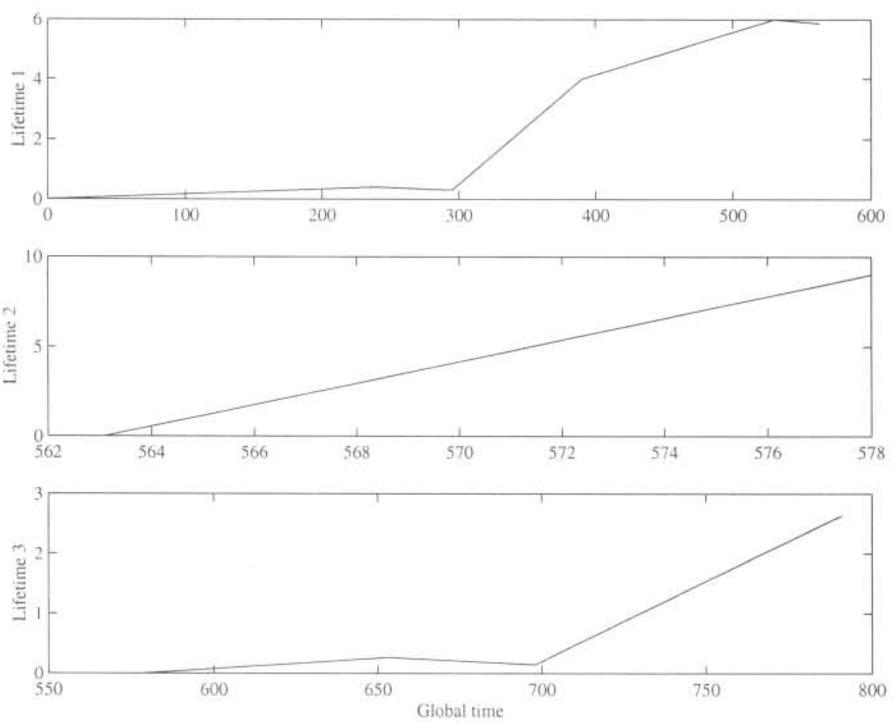

Figure D.11: Observed values of RF53H for PC1231

University of Pretoria etd – Vlok, P-J (2006)

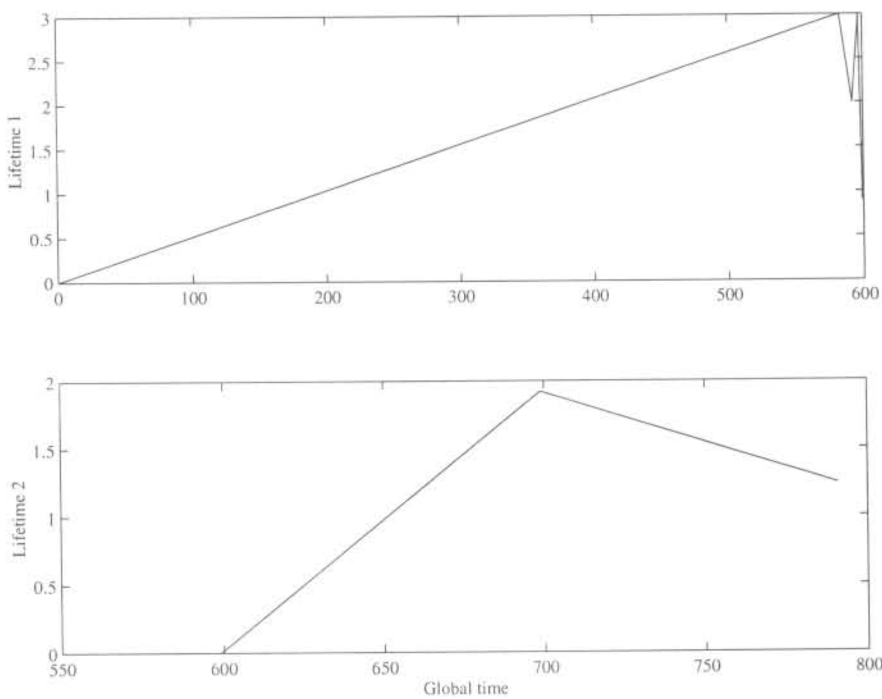

Figure D.12: Observed values of RF54H for PC1232

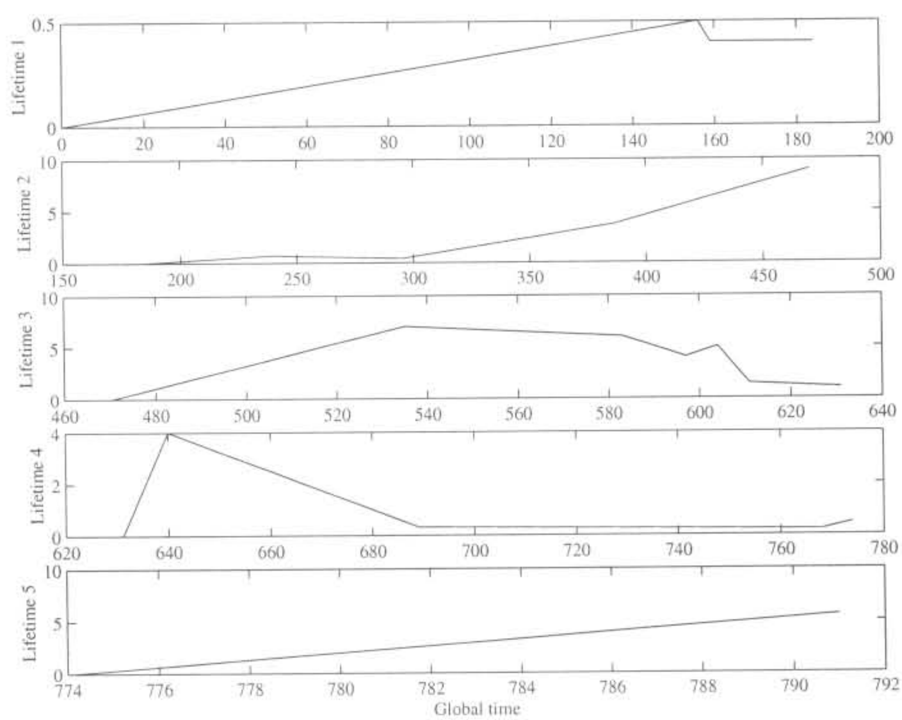

Figure D.13: Observed values of RF54H for PC2131

University of Pretoria etd – Vlok, P-J (2006)

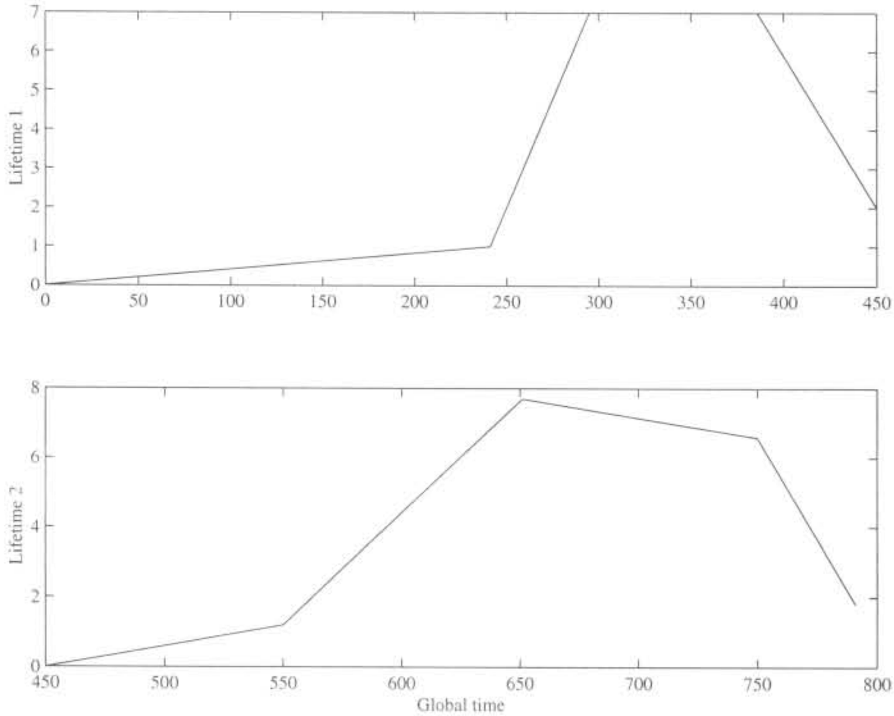

Figure D.14: Observed values of RF54H for PC3131

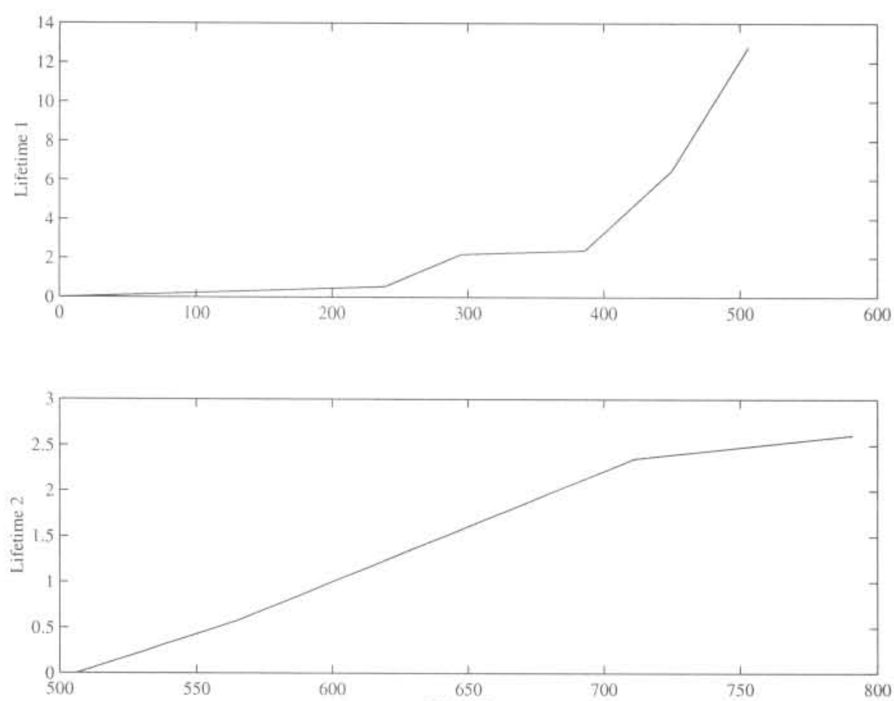

Figure D.15: Observed values of RF54H for PC3132

University of Pretoria etd – Vlok, P-J (2006)

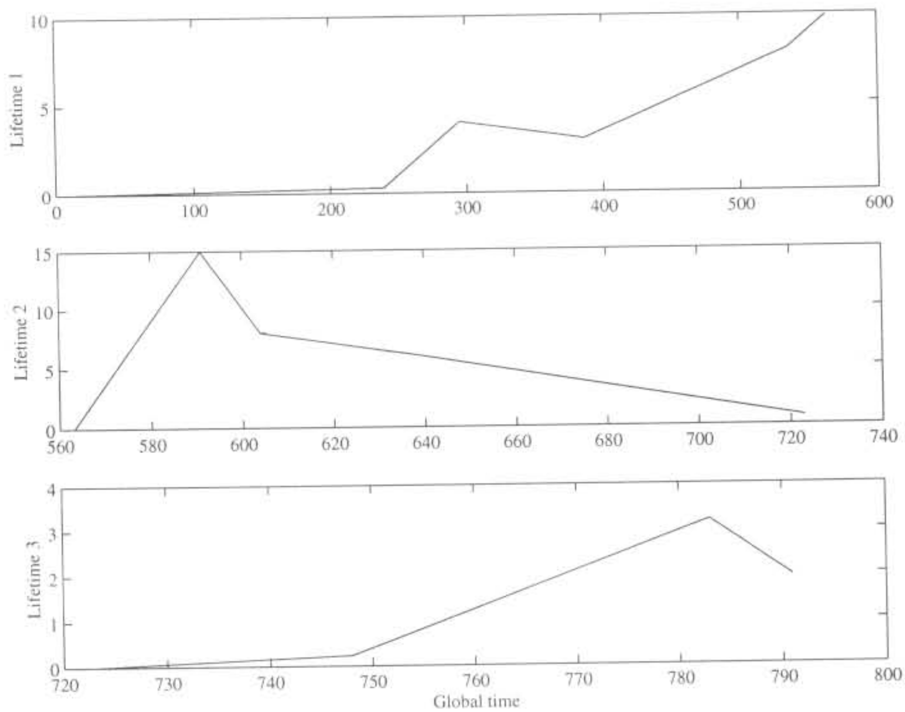

Figure D.16: Observed values of RF54H for PC3232

University of Pretoria etd – Vlok, P-J (2006)

Appendix E

Approximations for Covariates RF53H and RF54H

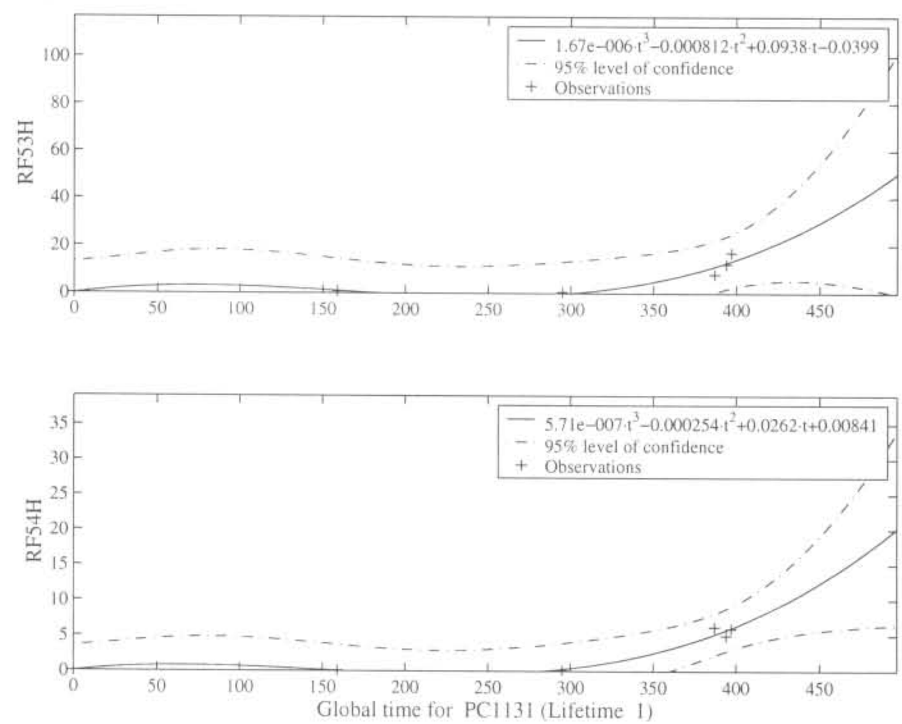

Figure E.1: Approximation of RF53H and RF54H measured on PC1131 during Lifetime 1

University of Pretoria etd – Vlok, P-J (2006)

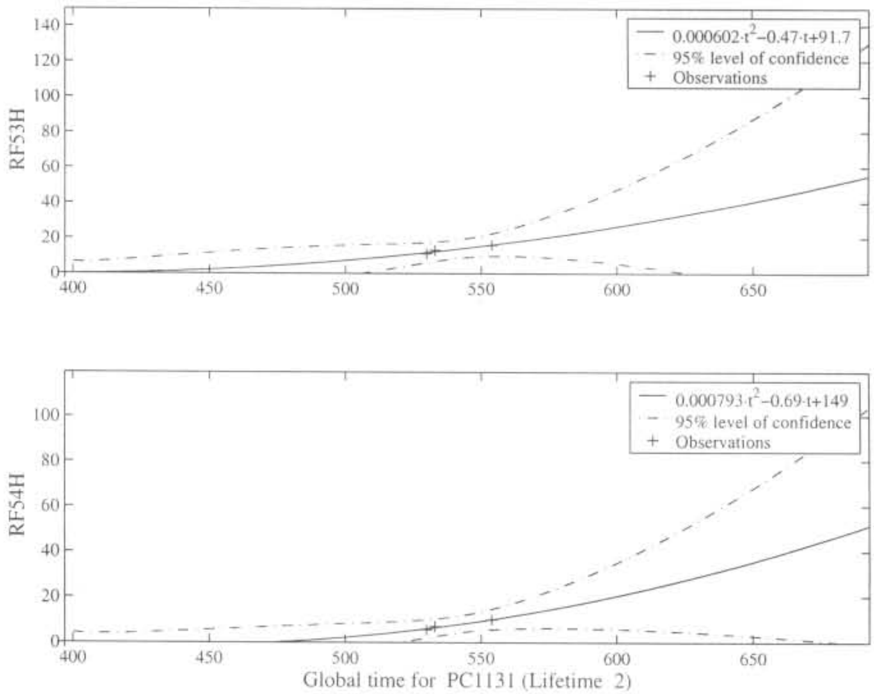

Figure E.2: Approximation of RF53H and RF54H measured on PC1131 during Lifetime 2

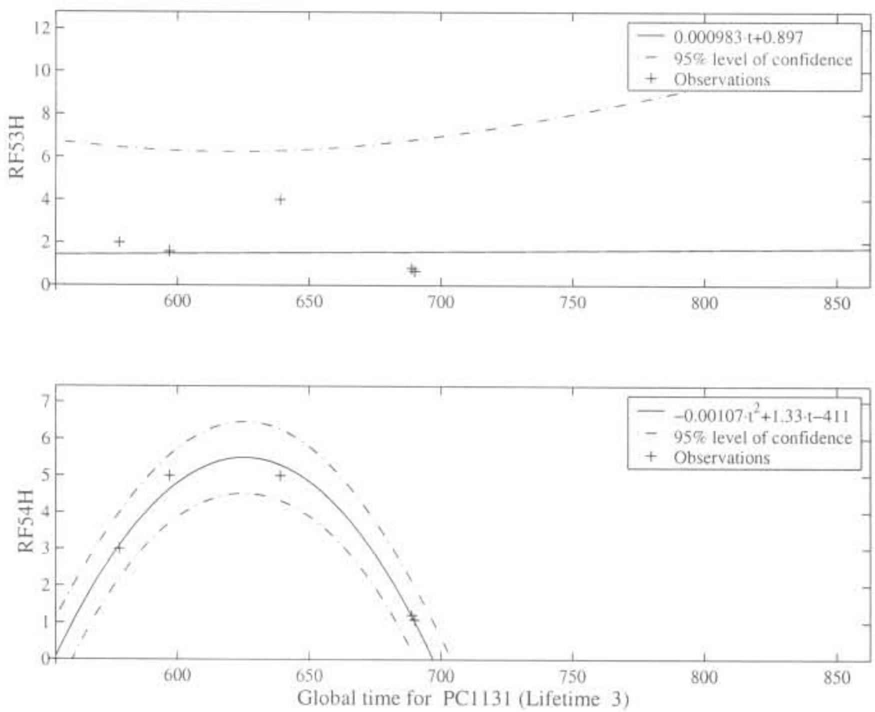

Figure E.3: Approximation of RF53H and RF54H measured on PC1131 during Lifetime 3

University of Pretoria etd – Vlok, P-J (2006)

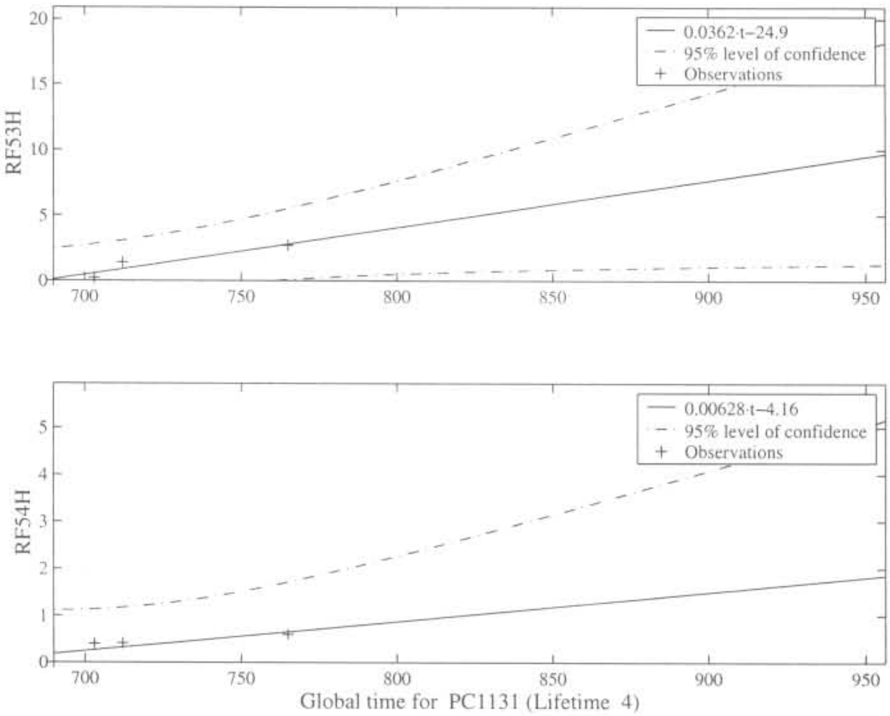

Figure E.4: Approximation of RF53H and RF54H measured on PC1131 during Lifetime 4

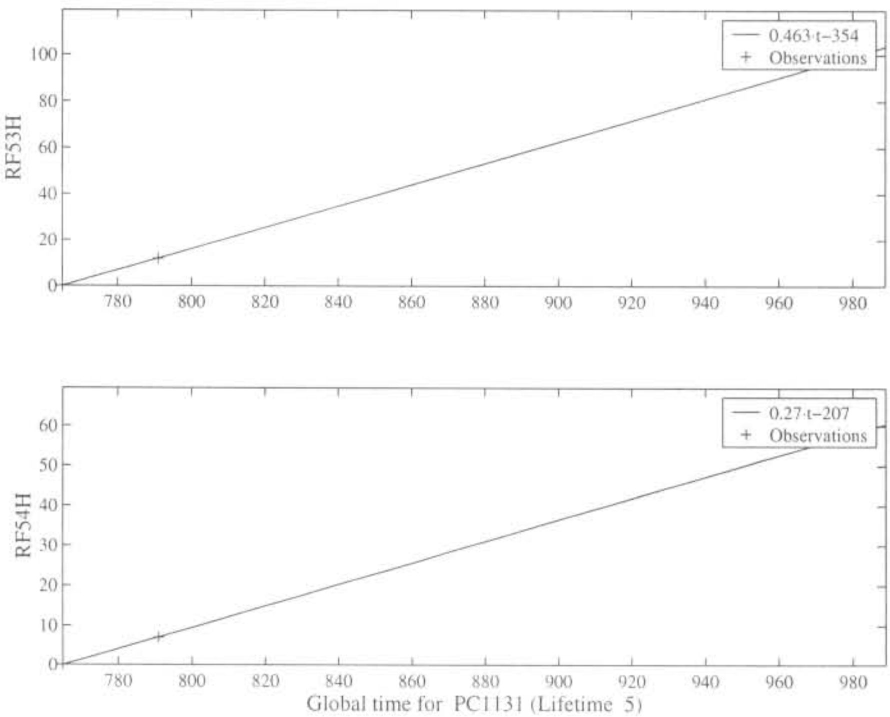

Figure E.5: Approximation of RF53H and RF54H measured on PC1131 during Lifetime 5

University of Pretoria etd – Vlok, P-J (2006)

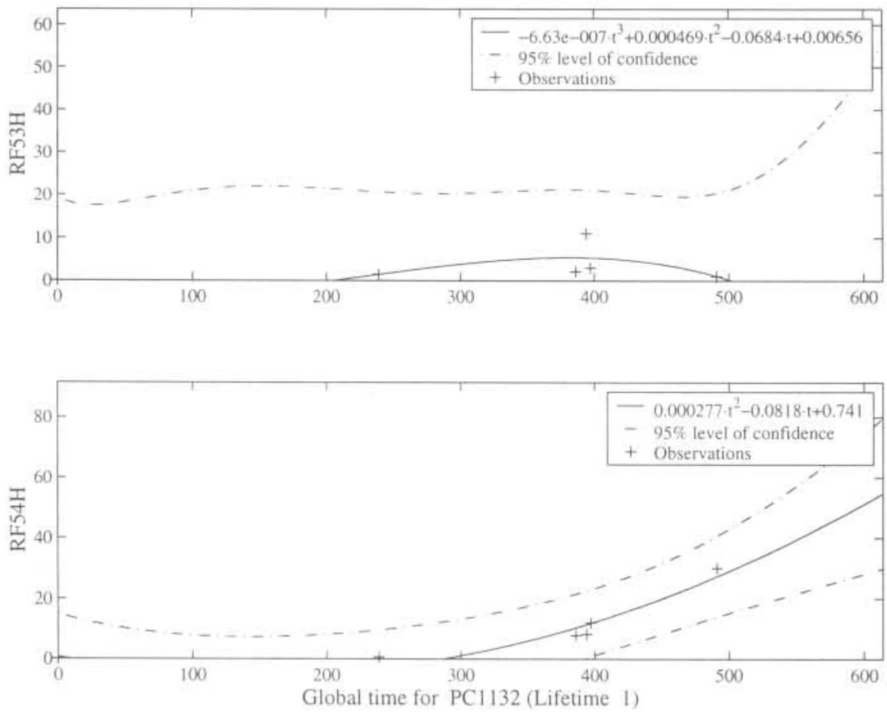

Figure E.6: Approximation of RF53H and RF54H measured on PC1132 during Lifetime 1

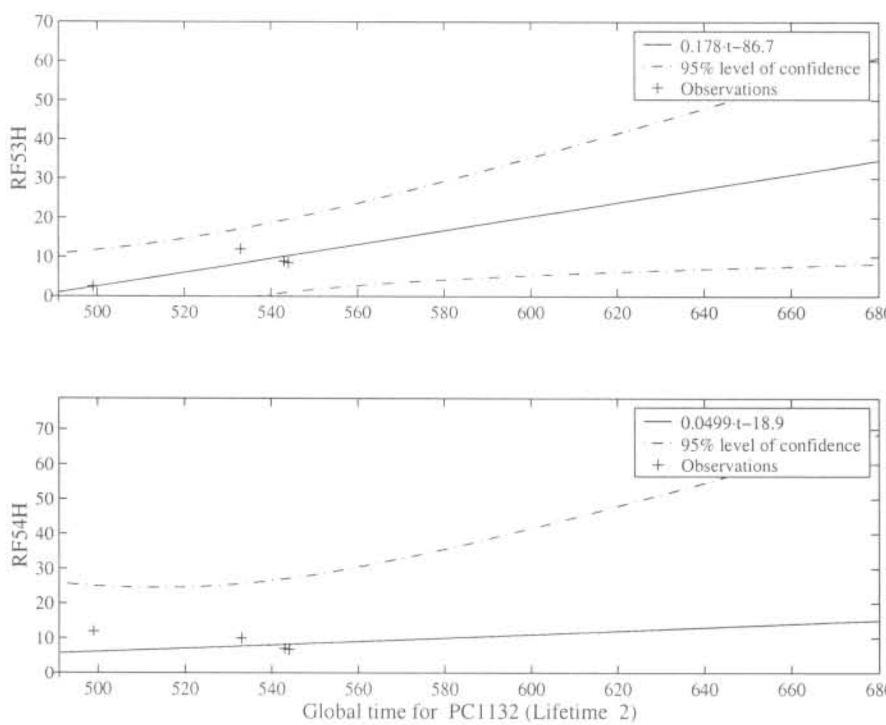

Figure E.7: Approximation of RF53H and RF54H measured on PC1132 during Lifetime 2

University of Pretoria etd – Vlok, P-J (2006)

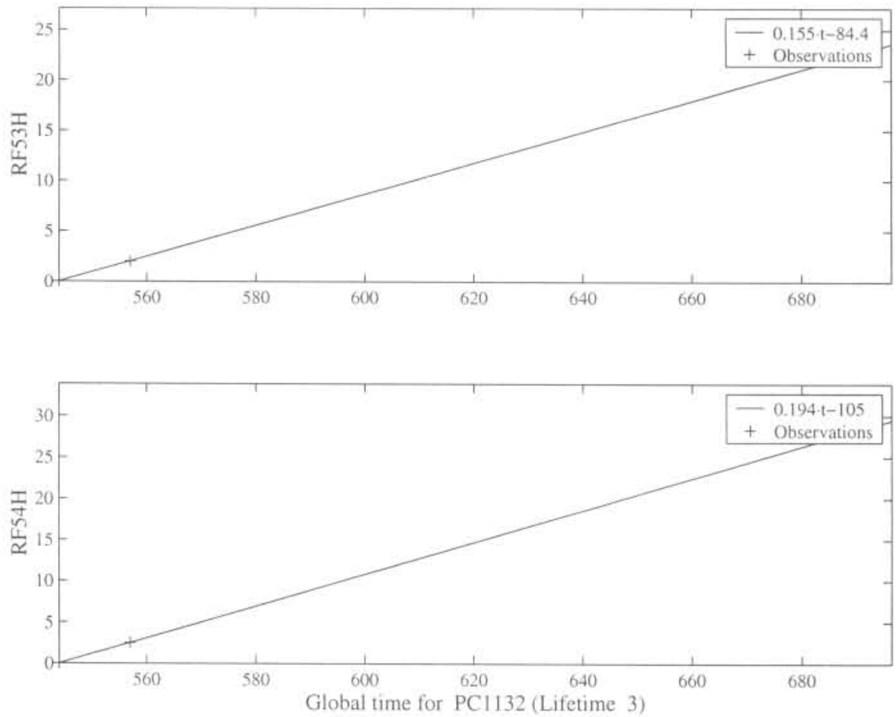

Figure E.8: Approximation of RF53H and RF54H measured on PC1132 during Lifetime 3

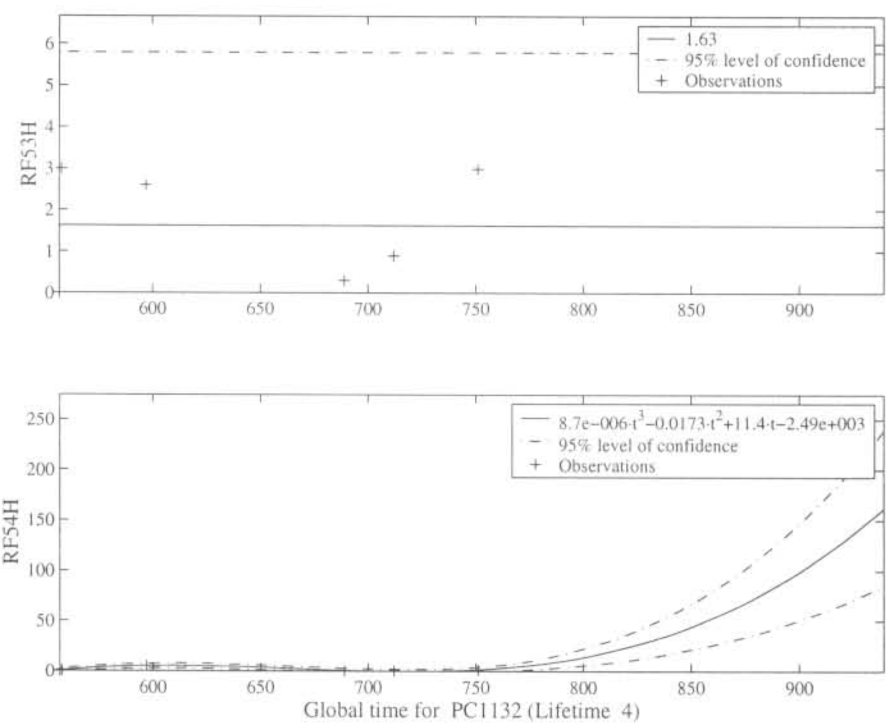

Figure E.9: Approximation of RF53H and RF54H measured on PC1132 during Lifetime 4

University of Pretoria etd – Vlok, P-J (2006)

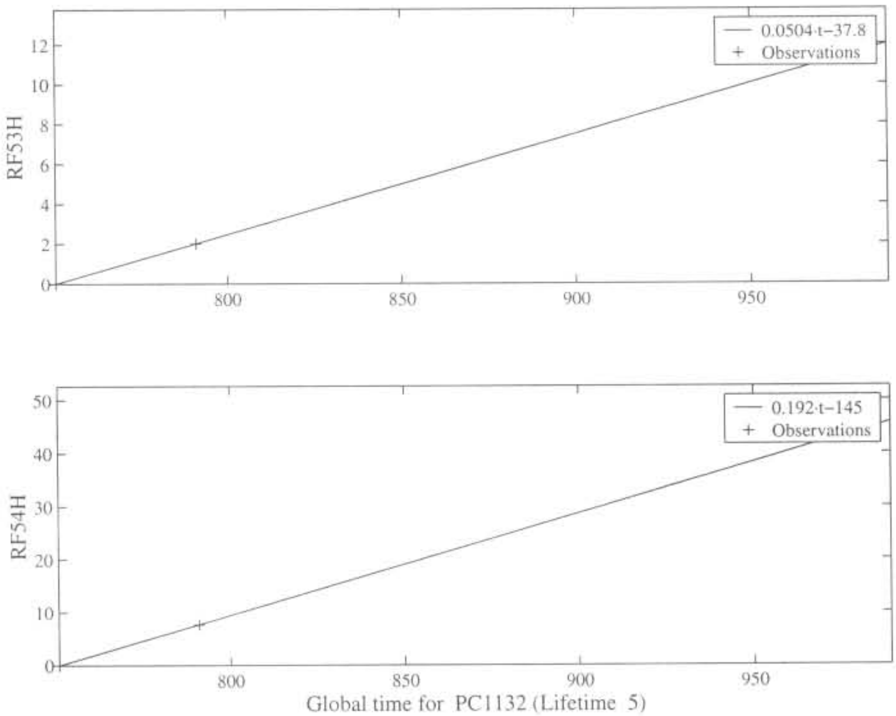

Figure E.10: Approximation of RF53H and RF54H measured on PC1132 during Lifetime 5

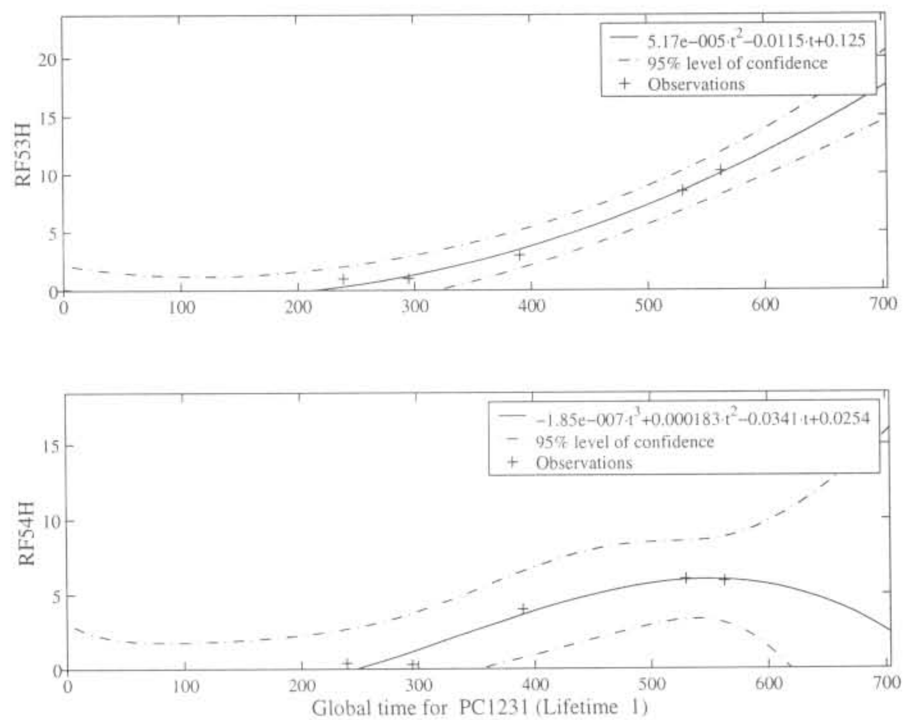

Figure E.11: Approximation of RF53H and RF54H measured on PC1231 during Lifetime 1

University of Pretoria etd – Vlok, P-J (2006)

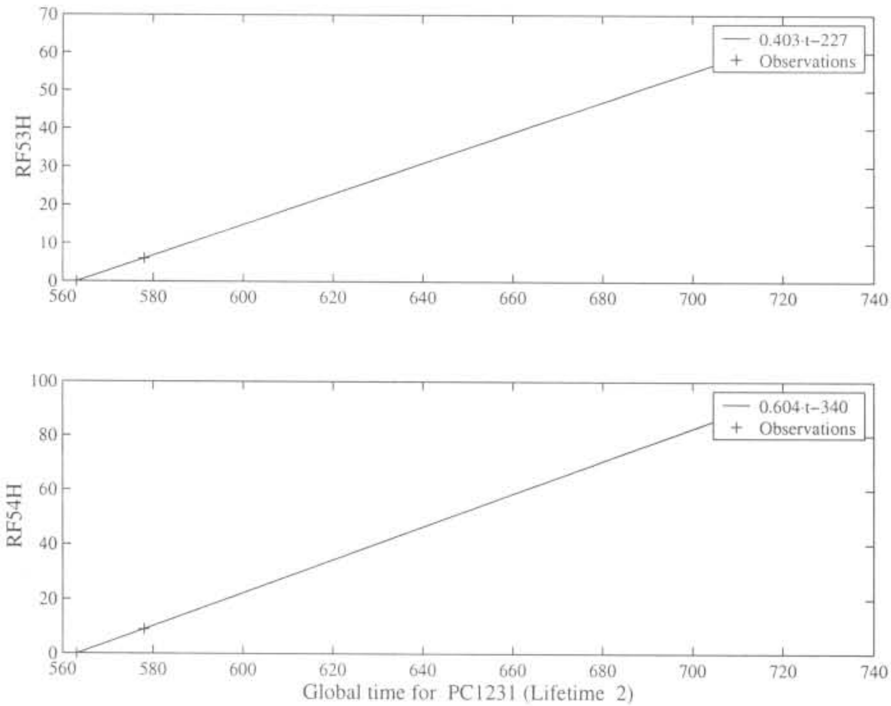

Figure E.12: Approximation of RF53H and RF54H measured on PC1231 during Lifetime 2

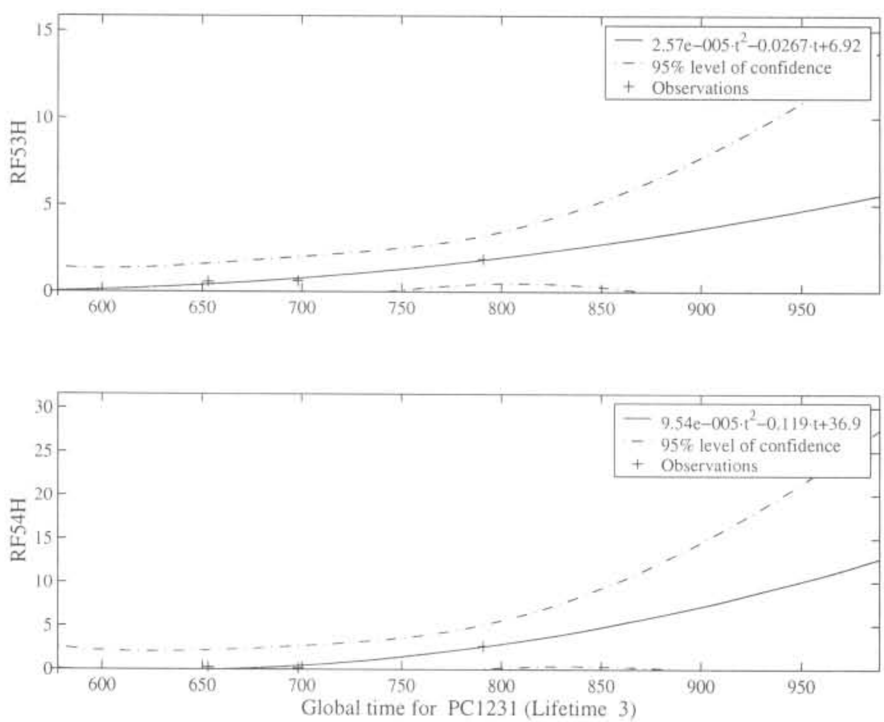

Figure E.13: Approximation of RF53H and RF54H measured on PC1231 during Lifetime 3

University of Pretoria etd – Vlok, P-J (2006)

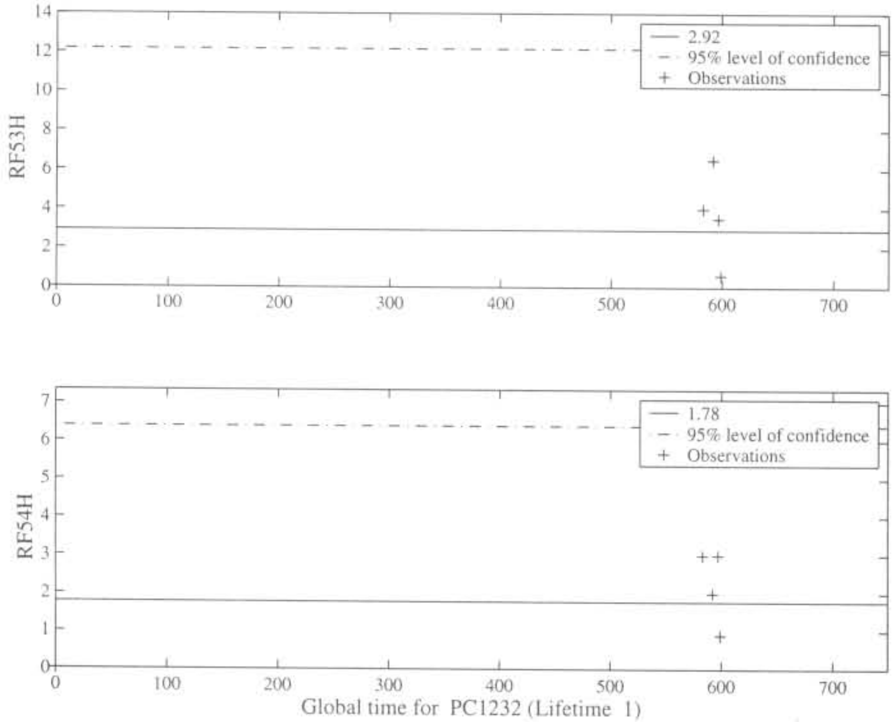

Figure E.14: Approximation of RF53H and RF54H measured on PC1232 during Lifetime 1

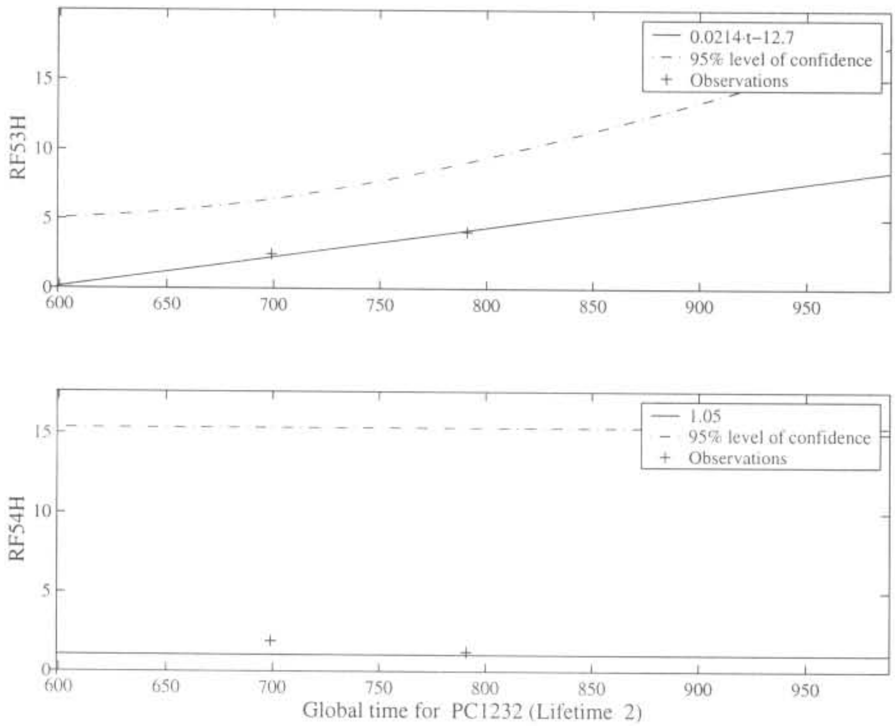

Figure E.15: Approximation of RF53H and RF54H measured on PC1232 during Lifetime 2

University of Pretoria etd – Vlok, P-J (2006)

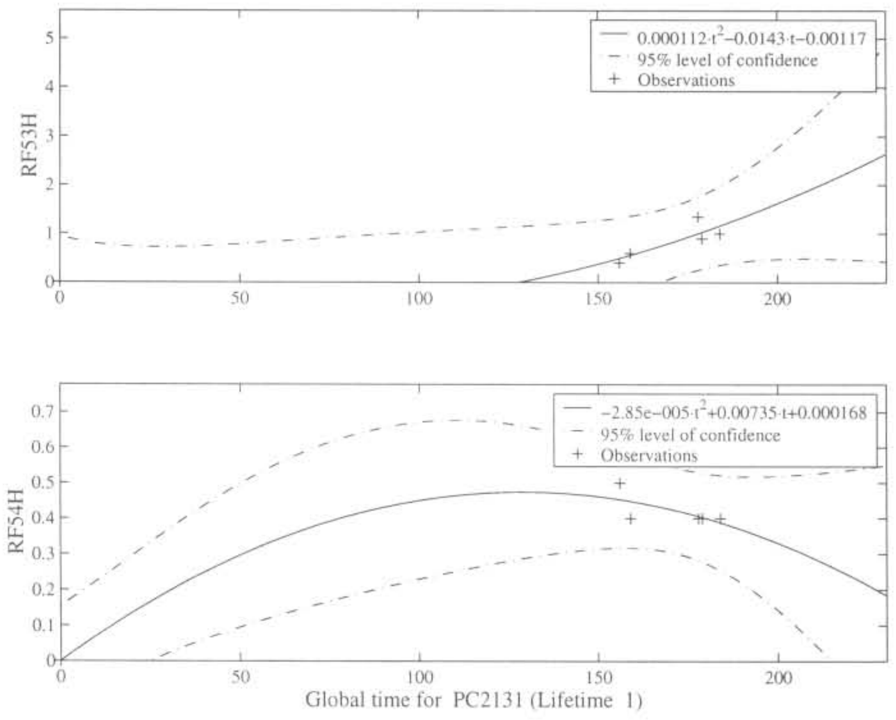

Figure E.16: Approximation of RF53H and RF54H measured on PC2131 during Lifetime 1

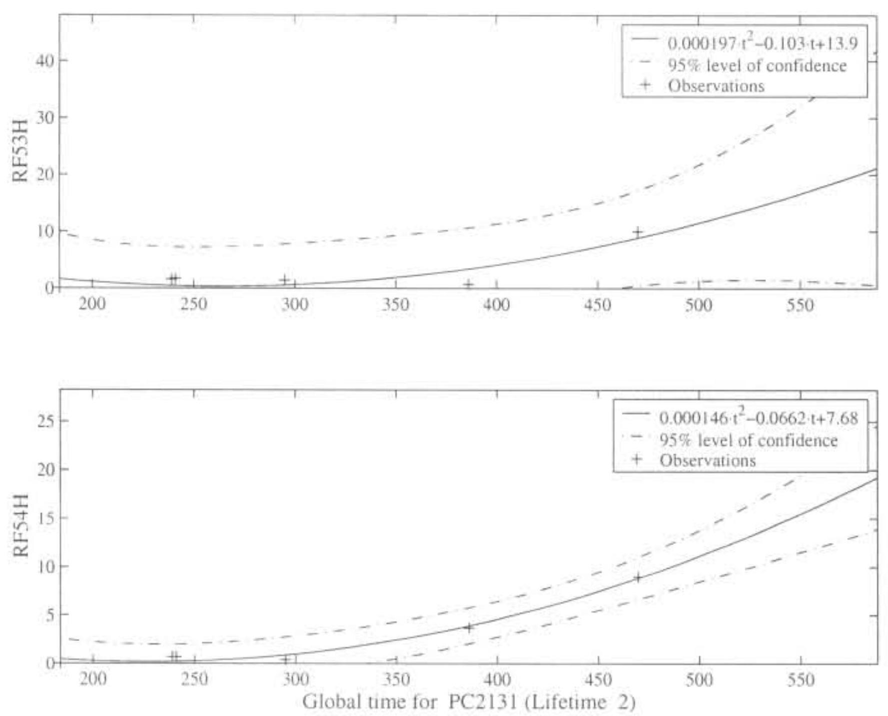

Figure E.17: Approximation of RF53H and RF54H measured on PC2131 during Lifetime 2

University of Pretoria etd – Vlok, P-J (2006)

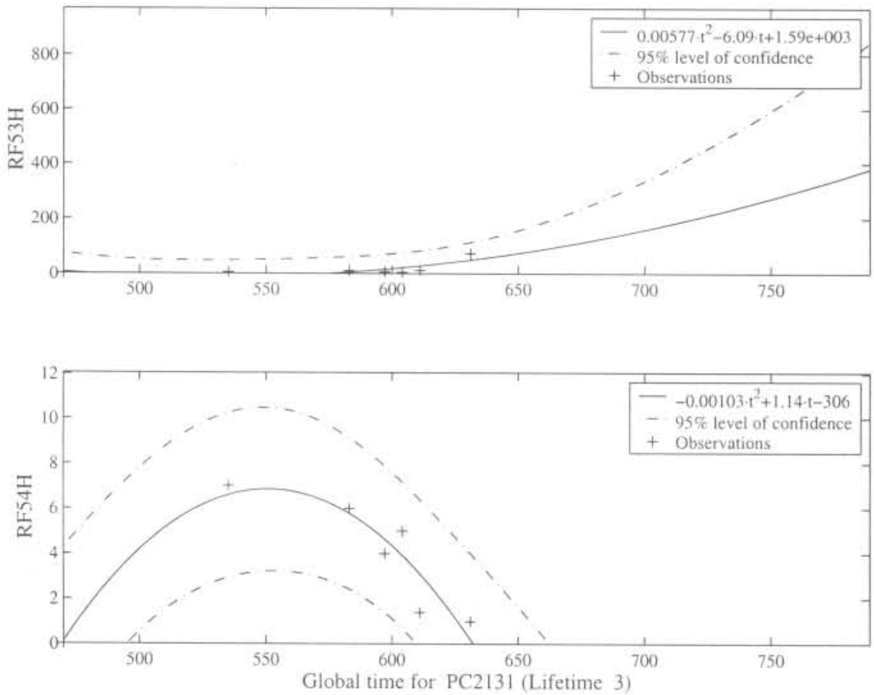

Figure E.18: Approximation of RF53H and RF54H measured on PC2131 during Lifetime 3

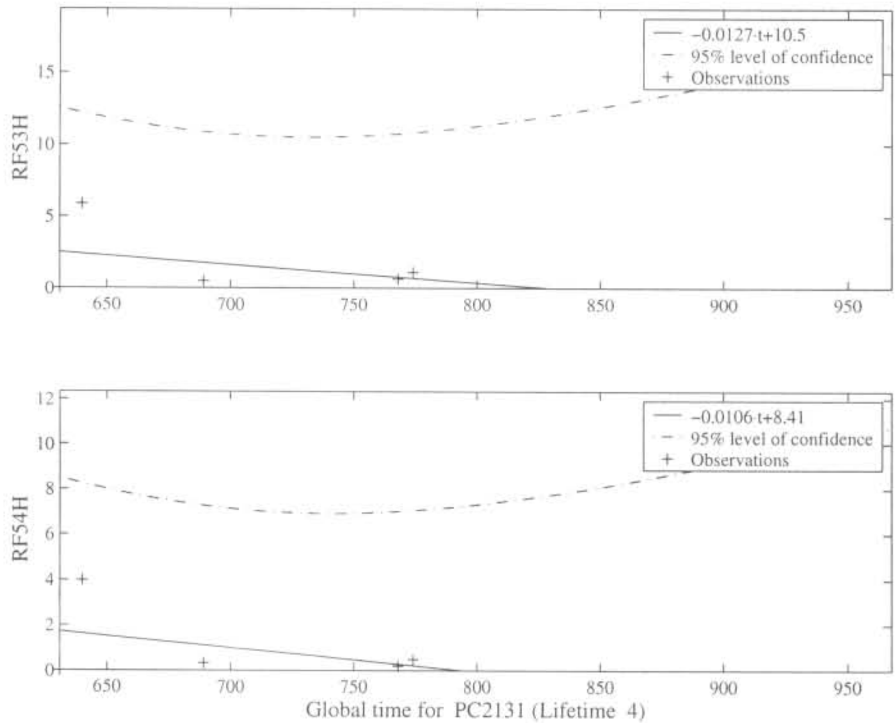

Figure E.19: Approximation of RF53H and RF54H measured on PC2131 during Lifetime 4

University of Pretoria etd – Vlok, P-J (2006)

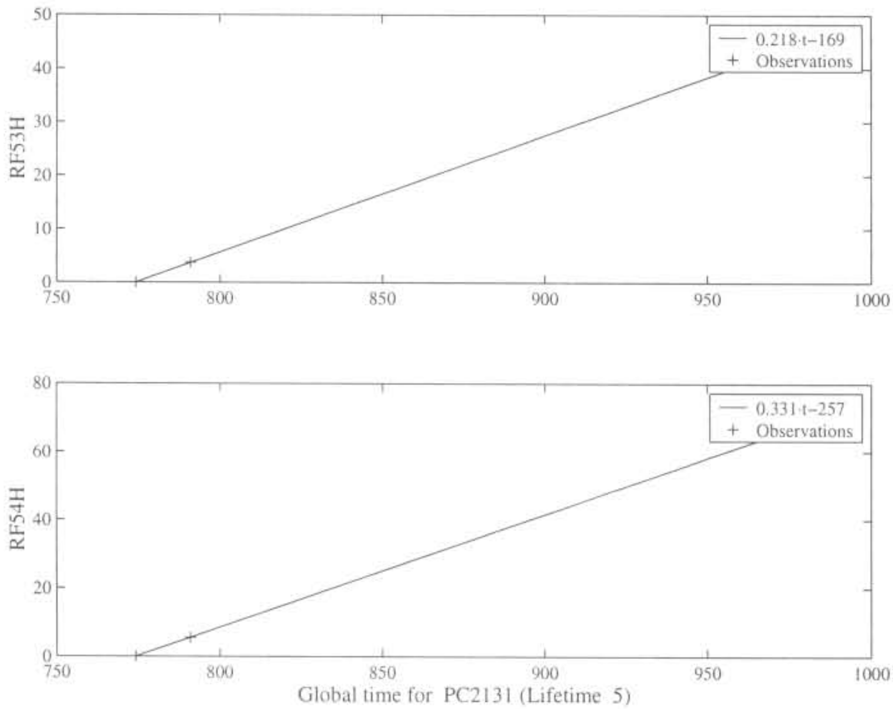

Figure E.20: Approximation of RF53H and RF54H measured on PC2131 during Lifetime 5

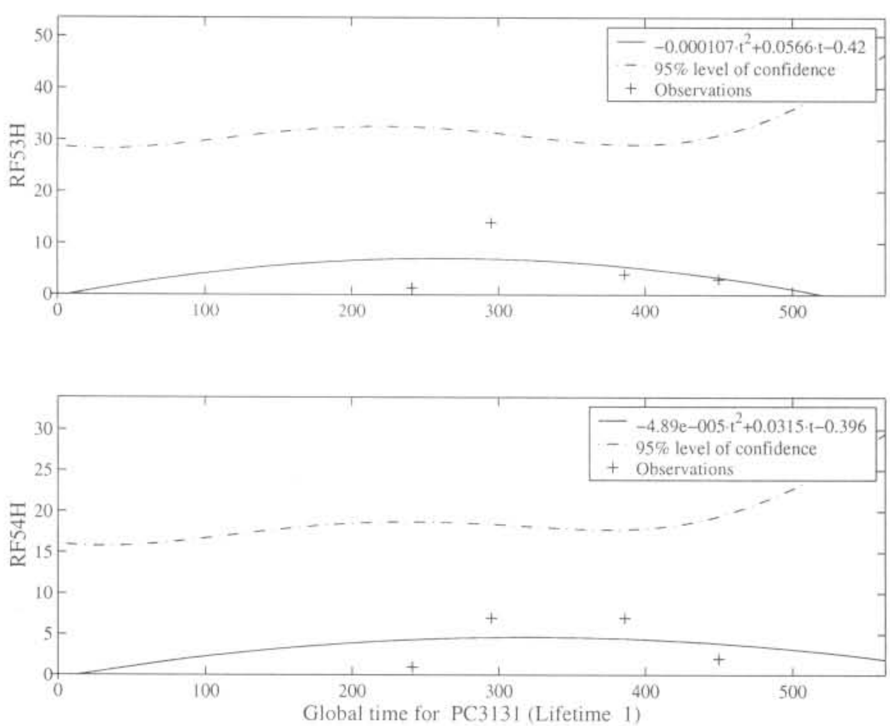

Figure E.21: Approximation of RF53H and RF54H measured on PC3131 during Lifetime 1

University of Pretoria etd – Vlok, P-J (2006)

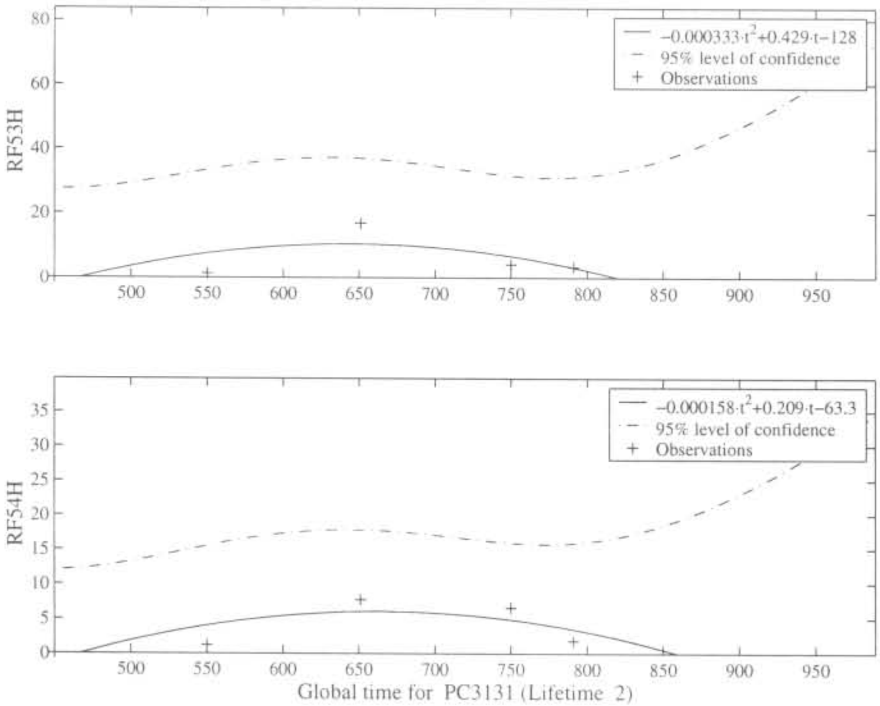

Figure E.22: Approximation of RF53H and RF54H measured on PC3131 during Lifetime 2

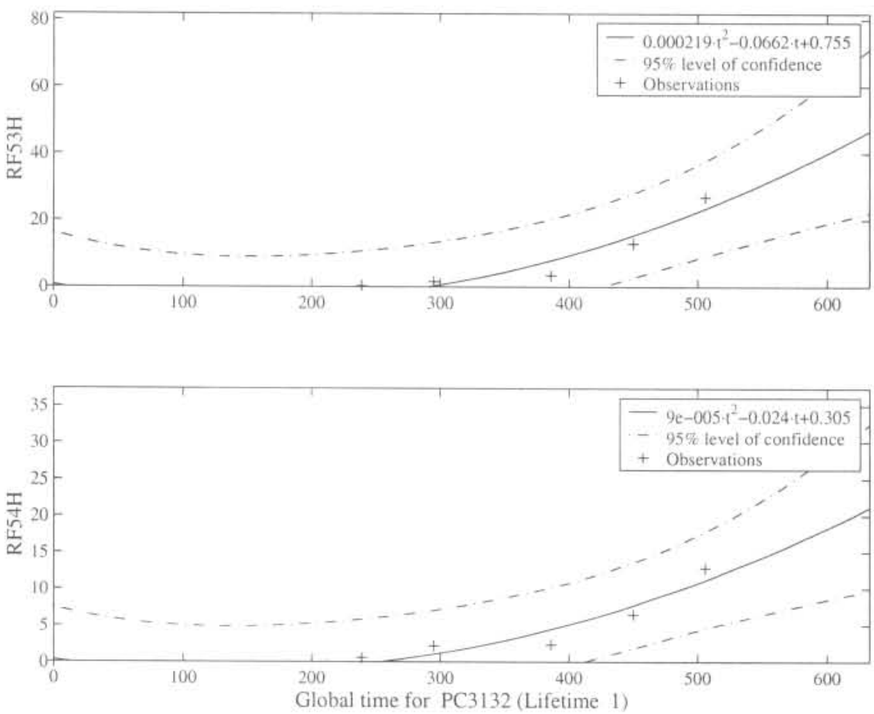

Figure E.23: Approximation of RF53H and RF54H measured on PC3132 during Lifetime 1

University of Pretoria etd – Vlok, P-J (2006)

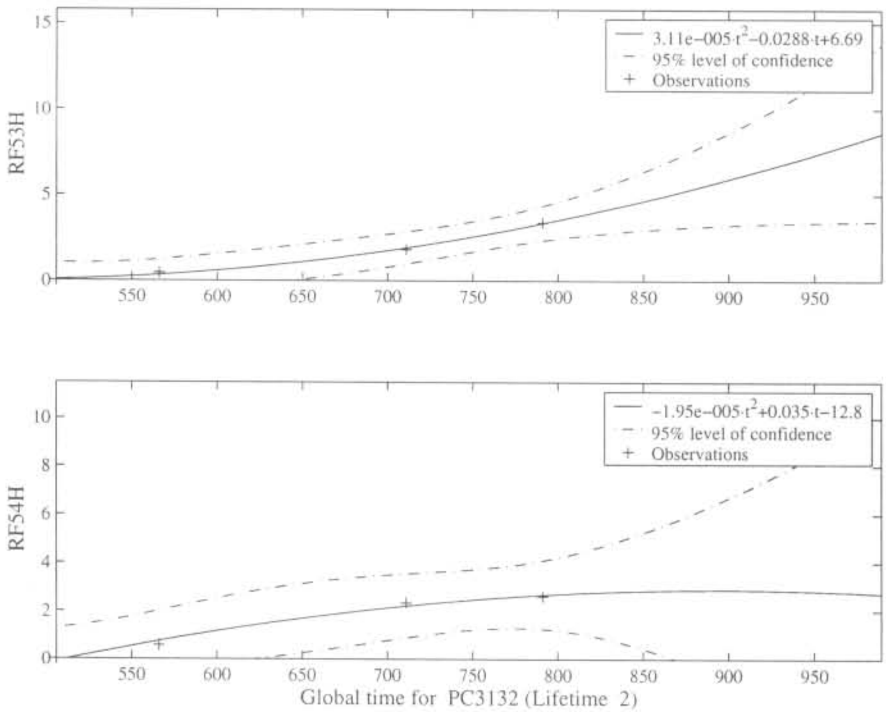

Figure E.24: Approximation of RF53H and RF54H measured on PC3132 during Lifetime 2

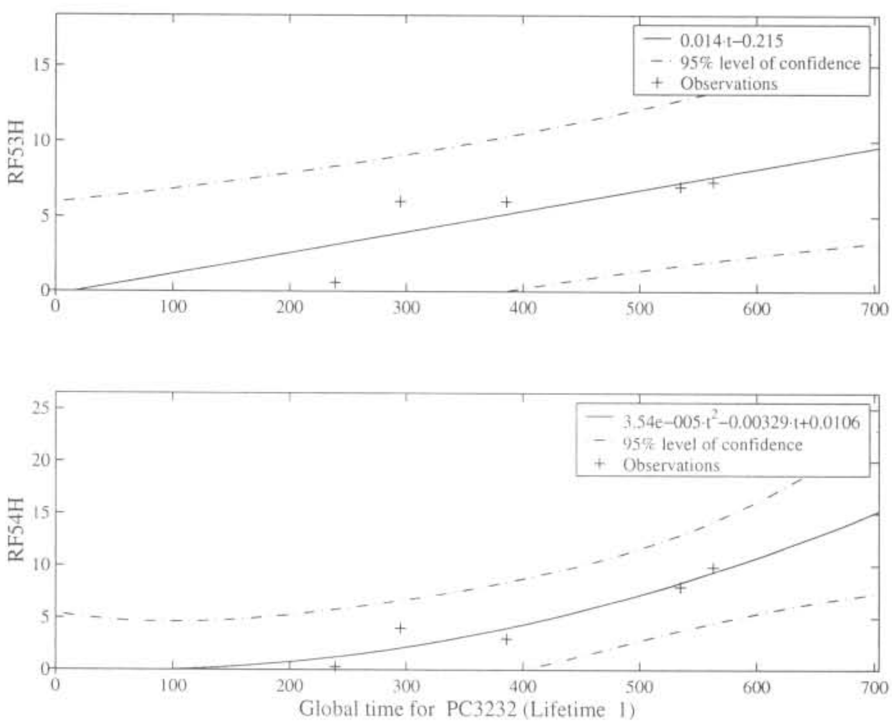

Figure E.25: Approximation of RF53H and RF54H measured on PC3232 during Lifetime 1

University of Pretoria etd – Vlok, P-J (2006)

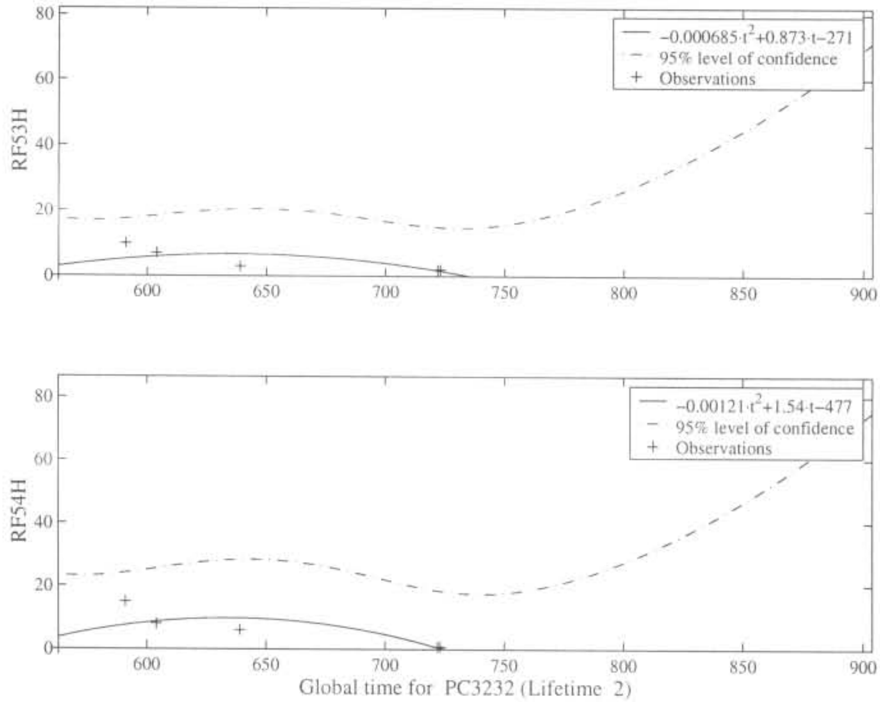

Figure E.26: Approximation of RF53H and RF54H measured on PC3232 during Lifetime 2

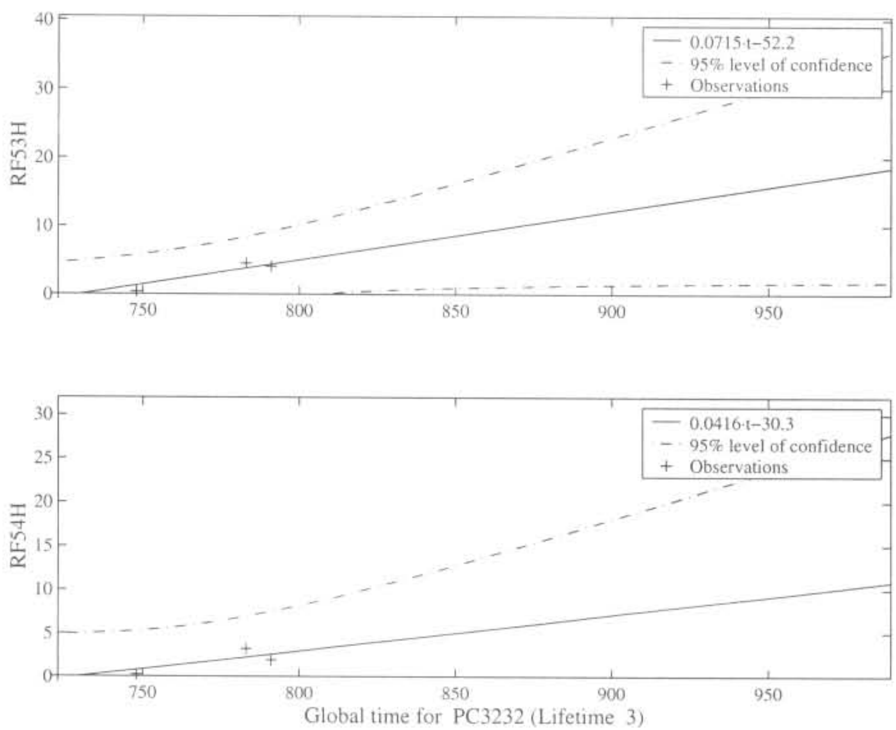

Figure E.27: Approximation of RF53H and RF54H measured on PC3232 during Lifetime 3

University of Pretoria etd – Vlok, P-J (2006)

Bibliography

Aalen, O. O. (1980). A model for non-parametric regression analysis of counting processes. *Lecture Notes in Statistics 2*, 1–25.

Aalen, O. O. (1989). A linear regression model for the analysis of life times. *Statistics in Medicine 8*, 907–925.

Andersen, P. K. (1985). Counting process models for life history data: A review. *Scandinavian Journal of Statistics 12*, 97–158.

Anderson, P. K. (1982). Testing goodness of fit of Cox's regression and life model. *Biometrics 38*, 67–77.

Anderson, P. K., O. Borgan, R. D. Gill, and N. Keiding (1993). *Statistical models based on counting processes.* Springer-Verlag.

Ansell, R. O. and J. Ansell (1987). Modelling the reliability of sodium sulphur cells. *Reliability Engineering 17*, 127–137.

Arnold, B. C. and H. Zahedi (1988). On multivariate mean remaining life functions. *Journal of Multivariate Analysis 25*, 1–9.

Ascher, H. E. (1983). Regression analysis of repairable systems reliability. Technical report, Naval Research Laboratory, Washington, D.C., 20375/USA.

Ascher, H. E. (1999, December). Numerous personal discussions with Ascher on the subject of reliability modeling.

Ascher, H. E. and H. Feingold (1984). *Repairable systems reliability: Modeling, inference, misconceptions and their causes.* Marcel Dekker.

Ascher, H. E. and C. K. Hansen (1998). Spurious exponentiality observed when incorrectly fitting a distribution to nonstationary data. *IEEE Transactions on Reliability 47*(4), 451–459.

Ascher, H. E., K. A. H. Kobbacy, and D. F. Percy (1997). Issues concerning effective modelling/analysis of PM scheduling. Technical report, Department of Computer and Mathematical Sciences, University of Salford, Salford M5 4WT.

Atkinson, K., R. Storb, R. L. Prentice, P. L. Weiden, R. P. Witherspoon, K. Sullivan, D. Noel, and E. D. Thomas (1979, April). Analysis of late infections in 89 long-term survivors of bone marrow transplantation. *Blood 53*(4), 720–731.

University of Pretoria etd – Vlok, P-J (2006)

Atwood, C. L. (1992). Parametric estimation of time-dependent failure rates for probabilistic risk assessment. *Reliability Engineering and System Safety 37*, 181–194.

Aven, T. and B. Bergman (1986). Optimal replacement times - a general set-up. *Journal of Applied Probability 23*, 432–442.

Aven, T. and R. Dekker (1997). A useful framework for optimal replacement models. *Reliability Engineering and System Safety 58*, 61–67.

Baganha, M. P., F. Geraldo, and D. F. Pyke (1999). The residual life of the renewal process: A simple algorithm. *Naval Research Logistics 46*, 435–443.

Bain, L. J., M. Engelhardt, and F. T. Wright (1985, June). Tests for an increasing trend in the intensity of a Poisson process: A power study. *Journal for the American Statistical Association 80*(390), 419–422.

Balakrishnan, N. (Ed.) (1995). *Recent advances in life-testing and reliability*, Chapter 32-33, pp. 589–617. CRC Press, Inc.

Banjevic, D. (2001). Numerous personal discussions with Banjevic on the subject of reliability modeling.

Banjevic, D., M. Ennis, D. Braticevic, V. Makis, and A. K. S. Jardine (1997). Analysis of the optimal policy and the structure of the software for condition-based maintenance. *COMADEM*.

Barlow, R. E. and L. Hunter (1960). Optimum preventive maintenance policies. *Operations Research 8*, 90–100.

Bartholomew, D. J. (1955). Discussion of Cox (1955). *Journal of the Royal Statistical Society 17*, 162–163.

Bartholomew, D. J. (1956a). A sequential test of randomness for events occurring in time or space. *Biometrika 43*, 64–78.

Bartholomew, D. J. (1956b). Tests for randomness in a series of events when the alternative is a trend. *Journal of the Royal Statistical Society 18*, 234–239.

Bates, G. E. (1955). Joint distributions of time intervals for the occurrence of successive accidents in a gereralized polya scheme. *Annals of Mathematical Statistics 26*, 705–720.

Bendell, A., M. Walley, D. W. Wrightman, and L. M. Wood (1986). Proportional hazards modelling in reliability analysis - an application to brake discs on high speed trains. *Quality and Reliability Engineering International 2*, 45–52.

Bennet, S. (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine 2*, 273–277.

Berman, M. and T. R. Turner (1992). Approximating point process likelihoods with GLIM. *Applied Statistics 41*(1), 31–38.

University of Pretoria etd – Vlok, P-J (2006)

Bhattacharjee, M. C. (1994). Aging influenced by repair, repair efficiency and realizability of a renewal related distribution. *International Journal of Quality, Reliability and Safety Engineering 12*, 147–159.

Billard, L. and M. R. Meshkani (1995). Estimation of a stationary Markov chain. *Journal of the American Statistical Association 90*, 307–315.

Blanchard, B. S. and W. J. Fabrycky (1990). *Systems engineering and analysis* (Second ed.). Prentice Hall International Series in Industrial and Systems Engineering.

Bodsberg, L. and P. Hokstad (1995). A system approach to reliability and life-cycle cost of process safety-systems. *IEEE Transactions on Reliability 442*, 179–186.

Booker, J., K. Campbell, A. G. Goldman, A. G. Johnson, and M. C. Bryson (1981). Application of Cox's proportional hazards model to light water reactor component failure data. Technical Report Report no. LA-8834-SR, Los Alamos Scientific Laboratory, Los Alamos, NM.

Boswell, M. T. (1966). Estimating and testing trend in a stochastic process of Poisson type. *Annals of Mathematical Statistics 37*, 1564–1573.

Boswell, M. T. and H. D. Brunk (1969). Distribution of likelihood ratio in testing against trend. *Annals of Mathematical Statistics 40*, 371–380.

Brillinger, D. R. (1978). Comparitive aspects of the study of ordinary time series and of point processes. In Krishnaiah (Ed.), *Developments in Statistics*, pp. 33–133. I.

Buchanan, J. L. and P. R. Turner (1992). *Numerical methods and analysis*. McGraw-Hill.

Burtschy, B., G. Albeanu, D. N. Boros, F. Popentiu, and V. Nicola (1997). Improving software reliability forecasting. *Microelectronics and Reliability 37*(6), 901–907.

Calabria, R., M. Guida, and G. Pulcini (1990). Point estimation of future failure times of a repairable system. *Reliability Engineering and System Safety 28*, 23–34.

Calabria, R. and G. Pulcini (2000). Inference and test in modeling the failure/repair process of repairable mechanical equipments. *Reliability Engineering and System Safety 67*, 41–53.

Chatfield, C. (1980). *The analysis of time series: An Introduction* (Second ed.). Chapman and Hall.

Chatterjee, S., R. B. Misra, and S. S. Alam (1997). Prediction of software reliability using an autoregressive process. *International Journal of System Science 28*(2), 211–216.

Cheng, R. C. H. and N. A. K. Amin (1983). Estimating the parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society B45*, 394–403.

Cheng, R. C. H. and T. C. Isles (1987). Corrected maximum likelihood in non-regular problems. *Journal of the Royal Statistical Society B45*, 95–101.

University of Pretoria etd – Vlok, P-J (2006)

Christer, A. H. and W. Wang (1995). A simple condition monitoring model for a direct monitoring process. *European Journal of Operational Research 82*, 258–269.

Christer, A. H., W. Wang, and J. M. Sharp (1997). A state space condition monitoring model for furnace erosion prediction and replacement. *European Journal of Operational Research 101*, 1–14.

Ciampi, A. and J. Etezadi-Amoli (1985). A general model for testing the proportional hazards and accelerated failure time hypothesis in the analysis of censored failure data with covariates. *Communication in Statistics - Theory and Methods 14*(3), 651–667.

Clifton, K. H. and J. J. Crowley (1978). Effects of radiation type and dose and the role of glucocorticoids, gonadectomy, and thyroidectomy in mammary tumor induction in mammotropin-secreting pituitary tumor grafted rats. *Cancer Research 38*, 1507–1513.

Cmiel, A. and H. Gurgul (2000). State-space form of dynamic input-output models with time-lags. *Systems Analysis Modelling Simulation 37*(1), 69–75.

Coetzee, J. L. (1997). *Maintenance*. Maintenance Publishers.

Collins, L. (1973). Estimating Markov transition probabilities from micro-unit data. *Applied Statistics 23*(3), 355–371.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society 17*, 129–164.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society B34*, 187–220.

Cox, D. R. (1975). Partial Likelihood. *Biometrika 62*(2), 269–276.

Cox, D. R. and P. A. Lewis (1966). *The statistical analysis of series of events*. Metheun, London.

Cozzolino, J. M. (1968). Probabilistic models of decreasing failure rate processes. *Naval Research Logistics Quarterly 17*, 361–374.

Crowder, M. (1991). The identifiability crisis in competing risks. Technical report, Royal Statistical Society Meeting on Dependency Modeling.

Crowder, M. J., A. C. Kimber, R. L. Smith, and T. J. Sweeting (1991). *Statistical analysis of reliability data*. Chapman and Hall.

Dale, C. J. (1985). Application of the Proportional Hazards Model in the reliability field. *Reliability Engineering 10*, 1–14.

De Laplace, P. S. (1773). See Bartholomew (1955).

Dekker, R. (1995). A general framework for optimization, priority setting, planning and combining of maintenance activities. *European of Operational Research 82*, 225–240.

Duane, J. T. (1964). Learning curve approach to reliability monitoring. *IEEE Transactions A-2*, 563–566.

University of Pretoria etd – Vlok, P-J (2006)

Ebrahimi, N. (1996, September). Engineering notion of mean residual life and hazard rate for finite populations with known distributions. *IEEE Transactions on Reliability 45*(3), 362–368.

Ellis, R. and D. Gulick (1990). *Calculus with analytic geometry* (Fourth ed.). Harcourt Brace Jovanovich College Publishers.

Etezadi-Amoli, J. and A. Ciampi (1987). Extended hazard regression for censored survival data with covariates: A spline approximation for the baseline hazard function. *Biometrics 43*, 181–192.

Fahrmeir, L. and G. Tutz (1994). *Multivariate Statistical Modeling based on Generalized Linear models.* Springer-Verlag.

Gail, M. H., T. J. Santner, and C. C. Brown (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics 36*, 255–266.

Ghai, G. L. and J. Mi (1999). Mean residual life and its association with failure rate. *IEEE Transactions on Reliability 48*(3), 262–266.

Gross, D. and P. Harris (1985). *Fundamentals of queuing theory* (Second ed.). Wiley.

Guess, F. and F. Prochan (1988). Mean residual life: Theory and applications. In P. Krishnaiah and C. R. Rao (Eds.), *Handbook of Statistcis*, pp. 215–224. Elsevier Science Publishers.

Guess, F. M. and D. H. Park (1988). Modeling discrete bathtub and upside-down bathtub mean residual life functions. *IEEE Transactions on Reliability 37*, 545–549.

Guess, F. M. and D. H. Park (1991). Non-parametric confidence bounds, using censored data, on the mean residual life. *IEEE Transactions on Reliability 40*, 78–80.

Harvey, A. C. (1981). *Time series models.* Philip Allan, Oxford.

Hasser, T. F., D. L. Dietrich, and F. Szidarovszky (1995). Time-varying failure rates in the availability and reliability analysis of repairable systems. *IEEE Transactions on Reliability 44*(1), 155–160.

Hines, W. W. and D. C. Montgomery (1980). *Probability and statistics in engineering and management science* (Second ed.). John Wiley and Sons.

Hokstad, P. (1997). The failure intensity process and the formulation of reliability and maintenance models. *Reliability Engineering ans System Safety 58*, 69–82.

Hosmer, D. W. and S. Lemeshow (1999). *Applied Survival Analysis.* Wiley and Sons, New York.

Hsu, L. (1999). Simultaneous determination of preventive maintenance and replacement policies in a queue-like production system with minimal repair. *Reliability Engineering and System Safety 63*, 161–167.

University of Pretoria etd – Vlok, P-J (2006)

Huang, J., C. R. Miller, and O. G. Okogbaa (1995, January). Optimal preventive-replacement intervals for the Weibull life distribution: solutions and applications. In *Proceedings of the Annual Reliability and Maintenance Symphosium*, pp. 370–377.

Jardine, A. K. S. (1973). *Maintenance, Replacement and Reliability.* Pitman Publishing.

Jardine, A. K. S. and M. Anderson (1988). Use of concomitant variables for reliability estimation. *Maintenance Management International 5*, 135–140.

Jardine, A. K. S., D. Banjevic, and V. Makis (1997). Optimal replacement policy and the structure of software for condition-based maintenance. *Journal of Quality in Maintenance Engineering 3*(2), 109–119.

Jardine, A. K. S., P. Ralston, N. Reid, and J. Stafford (1989). Proportional hazards analysis of diesel engine failure data. *Quality and Reliability Engineering International 5*, 207–216.

Jensen, U. (1990). Stochastic models of reliability and maintenance: An overview. Technical report, Institute of Stochastics, University of Ulm, Germany.

Kalbfleisch, J. D. and R. L. Prentice (1980). *The Statistical Analysis of Failure Time Data.* New York: John Wiley & Sons.

Karpinski, J. (1988, December). Distribution of residual system-life after partial failures. *IEEE Transactions on Reliability 37*(5), 539–544.

Kay, R. (1984). Goodness of fit methods for the Proportional Hazards regression model: A review. *Rev. Epidémiol. Et Santé Publique 32*, 185–198.

Klein, J. P. and M. L. Moeschberger (1990). *Survival Analysis: Techniques for censored and truncated data.* Springer.

Kobbacy, K. A. H., D. F. Percy, and B. B. Fawzi (1994). Sensitiwity analyses for preventive maintenance models. Technical report, Department of Mathematics and Computer Science, University of Salford, Salford M5 4WT.

Kulkarni, H. V. and R. N. Rattihalli (1996, June). Characterization of bivariate mean residual-life function. *IEEE Transactions on Reliability 45*(2), 249–253.

Kumar, D. (1996). Proportional hazards modelling of repairable systems. Paper published in Kumar's PhD.

Kumar, D. and B. Klefsjo (1993). Proportional hazards model: a review. *Reliability Engineering and System Safety 44*, 177–188.

Kumar, D. and U. Westberg (1996a). Maintenance scheduling under age replacement policy using PHM and TTT-plotting. *European Journal of Operational Research.*

Kumar, D. and U. Westberg (1996b). Some reliability models for analyzing the effect of operating conditions. Paper published in Kumar's PhD.

University of Pretoria etd – Vlok, P-J (2006)

Lagakos, S. W., C. J. Sommer, and M. Zelen (1978). Semi-Markov models for partially censored data. *Biometrika 65*(2), 311–317.

Lakey, P. B. (1993). Reliability modelling and design criteria for the backpropagation artificial neural networks. *Proceedings Annual Reliability and Maintainability Symposium*, 114–119.

Lam, C. T. and R. H. Yeh (1994). Optimal replacement policies for multisatae deteriorating systems. *Naval Research Logistics 41*, 303–315.

Lancaster, T. (1990). *The econometric analysis of transition data.* Cambridge University Press.

Lawless, J. F. (1987, September). Regression methods for Poisson Process data. *Journal of the American Statistical Society 82*(399), 808–815.

Lawless, J. F. and C. Nadeau (1995, May). Some simple robust methods for the analysis of recurrent events. *Technometrics 37*(2), 158–168.

Lawless, J. F. and K. Thiagarajah (1996, May). A point-process model incorporating renewals and time trends, with application to repairable systems. *Technometrics 38*(2), 131–138.

Lawrence, J. J. (1999). Use mathematical modeling to give your TPM implementation effort an extra boost. *Journal of Quality in Maintenance Engineering 5*(1), 62–69.

Ledoux, J. and G. Rubino (1977). Simple formulae for counting processes in reliability models. *Advanced Applied Probability 29*, 1018–1038.

Lee, B. S., S. H. Chung, K. Kim, F. P. Ford, and P. L. Andersen (1999). Remaining life prediction methods using operating data and knowledge on mechanisms. *Nuclear Engineering and Design 191*, 157–165.

Leitao, A. L. F. and D. W. Newton (1989). Proportional hazards modelling of aircraft cargo door complaints. *Quality and Reliability Engineering International 5*, 229–238.

Lewis, E. E. (1987). *Introduction to reliability engineering.* Wiley, New York.

Lim, J. and D. H. Park (1995). Trend change in mean residual life. *IEEE Transactions on Reliability 44*, 291–296.

Lim, J. and D. H. Park (1999). Evaluation of average maintenance cost for imperfect-repair model. *IEEE Transactions on Reliability 48*(2), 199–204.

Lindqvist, B., E. Molnes, and M. Rausand (1988). Analysis of SCCV performance data. *Reliability Engineering and System Safety 20*, 3–17.

Littlewood, B. and J. L. Verrall (1973). A Bayesian reliability growth model for computer software. *Applied Statistics 22*, 332–346.

Lorden, G. and I. Eisenberger (1973). Detection of failure rate increases. *Technometrics 15*, 167–175.

University of Pretoria etd – Vlok, P-J (2006)

Love, C. E. and R. Guo (1991a). Application of weibull proportional hazards modelling to bad-as-old failure data. *Quaility and Reliability Engineering International 7*, 149–157.

Love, C. E. and R. Guo (1991b). Using proportional hazards modelling in plant maintenance. *Quality and Reliability Engineering International 7*, 7–17.

Makis, V. and A. K. S. Jardine (1991). Optimal replacement in the proportional hazards model. *INFOR 30*, 172–183.

Makis, V. and A. K. S. Jardine (1992). Computation of optimal policies in replacement models. *IMA Journal of of Mathematics Applied in Business Industry 3*, 169–175.

Malik, M. A. K. (1979). Reliable preventive maintenance scheduling. *AIIE Transaction 11*, 221–228.

Martorell, S., A. Munoz, and V. S. Serradell (1996). Age-dependent models for evaluating risks and costs of surveillance and maintenance of components. *IEEE Transactions on Reliability 45*(3), 433–441.

Martorell, S., A. Sanchez, and V. Serradell (1999). Age-dependent reliability model considering effects of maintenance and working conditions. *Reliability Engineering and System Safety 64*, 19–31.

Martz, H. F. and P. H. Kvam (1996). Detecting trends and patterns in reliability data over time using exponentially weighted moving-averages. *Reliability Engineering and System Safety 51*, 201–207.

Mazzuchi, T. A. and R. Soyer (1989). Assessment of machine tool reliability using a proportional hazards model. *Naval Research Logistics 36*, 765–777.

McWilliams, T. P. (1979). Properties of isotonic estimators of mean lifetime in a simple prototype development model. Technical Report 194, Department of Operations Research, Stanford University.

Meeker, W. Q. and L. A. Escobar (1998). *Statistical methods for reliability data*. Chapter 4, John Wiley and Sons.

Mi, J. (1995, September). Bathtub failure rate and upside-down bathtub mean residual life. *IEEE Transactions on Reliability 44*(3), 388–391.

Moreau, T., J. O'Quigly, and M. Mesbah (1985). A global goodness-of-fit statistic for the Proportional Hazards Model. *Applied Statistics 34*, 212–218.

Morimura (1970). On some preventive maintenance policies for IFR. *Journal for the Operations Research Society of Japan 12*, 94–125.

Muth, E. J. (1977). *Reliability models with positive memory derived from the mean residual life function*, Volume 2, pp. 401–425. Academic Press.

Nair, K. R. M. and N. U. Nair (1989). Bivariate mean residual life. *IEEE Transactions on Reliability 38*(3), 362–364.

University of Pretoria etd – Vlok, P-J (2006)

Nelson, W. (1982). *Applied life data analysis.* Wiley series in probability and mathematical statistics. John Wiley & Sons.

Newby, M. (1988). Accelerated failure time models for reliability data. *Reliability Engineering and System Safety 20*, 187–197.

Newby, M. (1993). A critical look at some point-process models for repairable systems. *IMA Journal of Mathematics in Business and Industry 4*, 375–394.

Ng, M. K. (1999). A note on policy algorithms for discounted Markov decision problems. *Operations Research Letters 25*, 195–197.

Nolan, F. A. and H. Heap (1978). *Reliability Centred Maintenance.* National Technical Information Service, US Department of Commerce.

Park, K. S. (1979). Optimal number of minimal repairs before replacement. *IEEE Transactions R-28*, 137–140.

Pepe, M. S. and J. Cai (1993, September). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association 88*(423), 811–820.

Percy, D. F., K. A. H. Kobbacy, and H. E. Ascher (1998). Using proportional-intensities models to schedule preventive maintenance intervals. *IMA Journal of Mathematics Applied in Business and Industry 9*, 289–302.

Pijnenburg, M. (1991). Additive hazard models in repairable systems. *Reliability Engineering and System Safety 31*, 369–390.

Pike, M. C. (1966). A method of analysis of a certain class of experiments in carcinogenesis. *Biometrics 22*, 142–161.

Prentice, R. L., B. J. Williams, and A. V. Peterson (1981). On the regression analysis of multivariate failure time data. *Biometrika 68*(2), 373–379.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1993). *Numerical Recipes in C : The Art of Scientific Computing* (Second ed.). Cambridge University Press.

Prochan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics 5*, 375–383.

Pulkkinen, U. (1991). A stochastic model for wear prediction through condition monitoring. *Operations, Reliability and Systematic Maintenance*, 223–243.

Ran, A. and S. I. Rosenlund (1976). Age replacement with discounting for a continuous maintenance cost model. *Technometrics 18*, 45–465.

Rao, C. R. (1980). *Handbook of Statistics*, Volume 9. North-Holland Publishers.

Rao, S. S. (1995). *Mechanical Vibrations* (3rd ed.). Addison-Wesley Publishing Company.

University of Pretoria etd – Vlok, P-J (2006)

Rausand, M. and R. Reinertsen (1996). Failure mechanisms and life models. *International Journal of Reliability, Quality and Safety Engineering 3*.

Rawicz, A. H. and D. R. Girling (1994). Neural network enhancement for a reliability expert-system. *Proceedings Annual Reliability and Maintainability Symposium*, 468–474.

Reinertsen, R. (1996). Residual life of technical systems; diagnosis, prediction and life extension. *Reliability Engineering and System Safety 54*, 23–34.

Rhodes, P. H., M. E. Halloran, and I. M. Longini (1996). Counting process models for infectious disease data: Distinguishing exposure to infection from susceptibility. *Journal of the Royal Statistical Society B58*, 751–762.

Ridder, G. (1990). The non-parameteric identification of generalized accelerated failure-time models. *Review of Economic Statistics 57*, 235–254.

Ross, M. S. (1990). *Introduction to probability models* (Fith ed.). Academic Press.

Ross, S. M. (1969). Optimal dispatching of a Poisson process. *Journal of Applied Probability 6*, 692–699.

Ruiz, J. M. and J. Navarro (1994, December). Characterization of distributions by relationships between failure rate and mean residual life. *IEEE Transactions on Reliability 43*, 640–644.

Sakamoto, Y., M. Ishiguro, and G. Kitagawa (1983). *AKAKIE Information Criterion Statistics*. Reidel: Boston.

Saw, J. G. (1975). Tests on the intensity of a Poisson process. *Communication in Statistics 4*, 777–782.

Schäbe, H. (1995). A new approach to optimal replacement times for complex systems. *Miroelectronics and Reliability 35*(8), 1125–1130.

Schafer, R. E., R. B. Sallee, and J. D. Torrez (1975). Reliability growth study. Technical Report RADC-TR-75-253, Hughes Aircraft Company.

Schoenfield, D. (1990). Partial residuals for the proportional hazards regression model. *Biometrika 67*, 134–53.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*, 461–464.

Shen, X. (1998). Proportional odds regression and sieve maximum likelihood estimation. *Biometrica 85*(1), 165–177.

Sherif, Y. S. and M. L. Smith (1981). Optimal maintenance models for systems subject to failure - a review. *Naval Research Logistics Quarterly 28*, 47–74.

Sherwin, D. J. (1999). Age-based opportunity maintenance. *Journal of Quality in Maintenance Engineering 5*, 221–235.

University of Pretoria etd – Vlok, P-J (2006)

Sheu, S. (1999). Extended optimal replacement model for deteriorating systems. *European Journal of Operational Research 112*, 503–516.

Shimizu, S., Y. Ando, T. Morioka, and N. Okuzumi (1991). Maintenance management support systems for components aging estimation at nuclear power plants. *Nuclear Science and Technology 28*, 1041–1057.

Shin, I., T. Lim, and C. H. Lie (1996). Estimating parameters of intensity function and maintenance effect for repairable units. *Reliability Engineering and System Safety 54*, 1–10.

Shyur, H. and J. T. Luxhoj (1995). Reliability curve fitting for aging helicopter components. *Reliability Engineering and System Safety 48*, 229–234.

Singpurwalla, N. D. (1978). Estimating reliability growth (or deterioration) using time series analysis. *Naval Research Logistics Quarterly 25*, 1–14.

Smith, R. L. and J. C. Naylor (1987). A comparison of maximum likelihood and Baysian estimators for the three parameter Weibull distribution. *Applied Statistics 36*, 358–369.

Snyman, J. A. (1982). A new and dynamic method for unconstrained minimization. *Applied Mathematical Modeling 6*, 449–462.

Snyman, J. A. (1983). An improved version of the original leap-frog dynamic method for unconstrained minimization LFOP1(b). *Applied Mathematical Modeling 7*, 216–218.

Solomon, P. (1984). Effect of misspecification of regression models in the analysis of survival data. *Biometrika 71*, 291–298.

Stadje, W. and D. Zuckerman (1991). Optimal maintenance strategies for repairable systems with general degree of repair. *Journal of Applied Probability 28*, 384–396.

Tang, L. C., Y. Lu, and E. P. Chew (1999, March). Mean residual life of lifetime distributions. *IEEE Transactions on Reliability 48*(1), 73–78.

Tijms, H. (1976). *Analysis of (s,S) inventory policies* (Second ed.). Mathematical Centre Tracts 40, Mathematical Centrum, Amsterdam.

Van Noortwijk, J. M. (2000). Optimal maintenance decisions on the basis of uncertain failure probabilities. *Journal of Quality in Maintenance Engineering 6*(2), 113–122.

Vaupel, J. W., K. G. Manton, and E. Stallard (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography 16*, 439–454.

Vaurio, J. K. (1999). Identification of process and distribution characteristics by testing monotonic and non-monotonic trends in failure intensities and hazard rates. *Reliability Engineering and System Safety 64*, 345–357.

Vineyard, M., K. Amoako-Gyampah, and J. R. Meredith (1999). Failure rate distribution for flexible manufacturing system: An empirical study. *European Journal of Operational Research 116*, 139–155.

University of Pretoria etd – Vlok, P-J (2006)

Vlok, P. J. (1999). Vibration covariate regression analysis of failure time data with the Proportional Hazards Model. Master's thesis, University of Pretoria.

Vlok, P. J., J. L. Coetzee, D. Banjevic, A. K. S. Jardine, and V. Makis (2001). Optimal component replacement decisions using vibration monitoring and the Proportional Hazards Model. *Accepted for publication in the Journal for the Operational Research Society*.

Wals, L. A. and A. Bendell (1987). Time series models in reliability. *Reliability Engineering*, 239–265.

Wang, G., X. Wang, and G. Mao (1999). The state-space model in robust design. *Control Theory and Applications 16*(2), 252–254.

Wei, L. J., D. Y. Lin, and L. Weissfeld (1989, December). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association 84*(408), 1065–1073.

Williams, B. J. (1981). Proportional intensity models for multiple event times. Unpublished manuscript [Presented at August, 1981 American Statistical Association Meetings, Detroit].

Wismer, D. A. and R. Chattergy (1978). *Introduction to Nonlinear Optimization: A Problem Solving Approach*. North-Holland Publishers, New York.

Yeh, L. (1991). An optimal repairable replacement model for deteriorating systems. *Journal of Applied Probability 28*, 843–851.

Zahedi, H. (1985). Some new classes of multivariate survival distribution functions. *Journal of Statistical Planning and Inference 11*, 171–188.

Zahedi, H. (1991). Proportional mean remaining life model. *Journal of Statistical Planning and Inference 29*, 221–228.

Zhang, Z. G. and C. E. Love (2000). A simple recursive Markov chain model to determine the optimal replacement policies under general repairs. *Computers & Operations Research 27*, 321–333.

Zijlstra, M. (1981). Renewal replacement policies for one unit systems. *European Journal of Operational Research 8*, 289–293.