UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Vibration Covariate Regression Analysis of Failure Time Data with the Proportional Hazards Model

by

Pieter-Jan Vlok

Submitted as partial fulfillment of the requirements for the degree

Master in Engineering
(Mechanical Engineering)

Department of Mechanical and Aeronautical Engineering
Faculty of Engineering
University of Pretoria

Supervisor: Mr. J.L. Coetzee
Co-supervisors: Prof. A.K.S. Jardine and Prof. J.L. van Niekerk

October, 1999

Acknowledgements

I would like to express my sincere gratitude towards,

- Mr. Jasper Coetzee, my study leader, for his support, guidance and encouragement throughout this study;
- Prof. Andrew Jardine of the University of Toronto for partially sponsoring and coordinating this research project;
- Dr. Dragan Banjevic of the CBM lab at the University of Toronto for giving me technical support and evaluating my results;
- Prof. Wikus van Niekerk for his co-supervision;
- SASOL, Secunda for supplying me with suitable data and permitting me to publish the results;
- My father and sister, for their continued interest in my work;
- Our heavenly Father, for giving me the strength and perseverance to complete this task which He laid upon me.

The Author
October, 1999

Summary

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Vibration Covariate Regression Analysis of Failure Time Data with the Proportional Hazards Model

by

Pieter-Jan Vlok

Supervisor: Mr. J.L. Coetzee
Co-supervisors: Prof. A.K.S. Jardine & Prof. J.L. van Niekerk

Master in Engineering
Department of Mechanical and Aeronautical Engineering
University of Pretoria

There is no doubt about the potential economical advantages of preventive renewals based on statistical failure analysis or vibration monitoring, if performed correctly. Despite their advantageous abilities to produce economical benefits, both techniques have shortcomings. Vibration monitoring strategies recommend renewal based on short term vibration information only (often by waterfall plots or other short term trending techniques) and no scientific technique exists with which long term vibration information can be included in recommendations. Conventional statistical failure analysis techniques, on the contrary, utilizes the statistical long term life cycle cost optimum to base renewal decisions on and do not take diagnostic information (like vibration information) into account. The mentioned techniques complement each other extremely well and in this dissertation a scientific method to integrate these techniques was searched for with the emphasis on practicality.

Regression models with the ability to handle covariates (explanatory variables) was found to be the most logical route to the dissertation objectives. After an extensive literature study, the Proportional Hazards Model (PHM) was selected as the most suitable model for this application because of its sound theoretical foundation, numerical tractability and previous successes. The PHM was thoroughly researched including the investigation of different model forms, numerical parameter estimation techniques, covariate behavior and goodness-of-fit tests. A decision model utilizing the PHM, with the ability to handle non-monotonic covariates (like vibration parameters) through Markovian chains, was also identified and studied.

Data obtained from a typical South African industrial concern was collected and modeled with the studied theory, with promising results.

Keywords: Proportional Hazards Model, Vibration Monitoring, Failure Analysis, Preventive Maintenance, Renewal

Opsomming

Vibrasie Verklarende Veranderlike Analise van Falingstyd Data met die Proporsionele Gevaarkoers Model

deur

Pieter-Jan Vlok

Leier: Mnr. J.L. Coetzee
Mede-Leiers: Prof. A.K.S. Jardine & Prof. J.L. van Niekerk

Magister in Ingenieurswese
Departement van Meganiese en Lugvaartkundige Ingenieurswese
Universiteit van Pretoria

Daar is geen twyfel oor die potensiële ekonomiese voordele van voorkomende hernuwing gebaseer op statistiese falingsanalise of vibrasie monitering nie, mits dit korrek bedryf word. Ten spyte van hul vermoëns wat groot finansiële voordele inhou, het beide tegnieke tekortkominge. Vibrasie monitering strategieë beveel hernuwing aan alleenlik gebaseer op korttermyn vibrasie inligting (tipies deur waterval grafieke of ander korttermyn tendens identifiseringstegnieke) en geen wetenskaplike tegniek bestaan waarmee langtermyn vibrasie inligting in besluite geïnkorporeer kan word nie. In teenstelling hiermee, neem konvensionele statistiese falingsanalises geen diagnostiese inligting (soos vibrasie inligting) in ag in hernuwings aanbevelings nie, maar grond besluite op die minimum statistiese langtermyn lewenssikluskoste. Die bogenoemde tegnieke komplimenteer mekaar uitstekend en in hierdie studie is daar gesoek na 'n wetenskaplike metode waarmee die twee tegnieke geïntegreer kan word met die klem op praktiese toepaslikheid.

Regressie modelle met die vermoë om verklarende veranderlikes te hanteer, het geblyk om die mees logiese roete na die studiedoelwitte te wees. Na 'n deeglike literatuurstudie is besluit dat die Proporsionele Gevaarkoersmodel (PGM) die mees geskikte model is vir die genoemde toepassing vanweë sy breë teoretiese fondasie, numeriese implementeerbaarheid, asook vorige suksesse wat behaal is in soortgelyke toepassings. Die PGM is in diepte bestudeer, onder andere verskillende modelvorms, numeriese parameter beramingstegnieke, verklarende veranderlike gedrag en pasgehalte toetse. 'n Besluitnemingsmodel wat op die PGM gebasseer is met die vermoë om nie-monotone verklarende veranderlikes (soos vibrasie parameters) te hanteer deur Markov kettings is ook geïdentifiseer en bestudeer.

Data is verkry uit 'n tipiese Suid-Afrikaanse industriële situasie en is gebruik om die bestudeerde teorie te toets, met belowende resultate.

Sleutelwoorde: Proporsionele Gevaarkoersmodel, Vibrasie Monitering, Falingsanalise, Voorkomende Instandhouding, Hernuwing

Table of Contents

Chapter 5: Closure 99

References 102

Notation

$\overline{z}$ / $\overline{z(t)}$	Covariate vector dependent / independent of time
$\lambda(\overline{z(t)})$	Functional term dependent on time and covariates
z_i / $z_i(t)$	i^{th} covariate dependent / independent of time
$\overline{\gamma}$	Vector of regression coefficients
$\hat{P}_{ij}$	Transition probability estimate
Δ	Inspection interval
β	Weibull shape parameter
η	Weibull scale parameter
λ_{ij}	Transition rate from state i to j
$[]^{T}$	Transposed vector
AFTM	Accelerated failure time model
CBM	Condition Based Maintenance
C_f	Cost of unexpected failure renewal
CM	Condition monitoring
C_p	Cost of planned preventive renewal
d	Threshold risk level
f	Unreliability function / Cumulative failure density function
F	Failure density function
$g(t)$	Warning level function
h	Hazard rate
H	Cumulative hazard rate
h_0	Baseline hazard rate
IID	Independent and identically distributed
KS	Kolmogorov-Smirnov
L	Likelihood
l	Log-likelihood
LCC	Long term life cycle cost
L_e	Expected life
MTTF	Mean time to failure
NHPP	Non Homogeneous Poison Process
$P\{\}$	Probability
PDF	Probability density function
PHM	Proportional hazards model / modelling
PL	Partial likelihood
POM	Proportional odds model
PWP	Prentice Williams Peterson model
$Q(d)$	Probability of failure renewal

R	Reliability / Survivor function
R_0	Baseline survivor function
ROCOF	Rate of occurrence of failure
t	Continuous time
T_i	i^{th} Observed renewal time
$\boldsymbol{T_i}$	i^{th} Random renewal time variable
TPM	Transition probability matrix
$W(d)$	Expected time to renewal

Chapter 1

Problem Statement

1 Introduction

There will always be a need for maintenance on mechanical systems and components because of friction and wear. It is however not always as simplistic as "if it breaks, repair it" and in recent years the need for scientific maintenance has increased tremendously. Reasons for this could be the increased sophistication of production equipment, the need for a high return on an investment and the high cost of maintenance.

The objective of an organization's maintenance activities should be to support the production process with maximal levels of availability, reliability, operability and safety at acceptable cost. If this objective is pursued the results will be clearly evident in the form of increased production capacity and thus company profit. As high profit is the reason for the existence of production concerns, maintenance should be regarded in a very serious way.

Various maintenance models exist to act as guidelines for an effective maintenance function. One such model called 'The Maintenance Cycle' is proposed by Coetzee[22] and act as a good overview of the total complicated maintenance function. See Figure 1.1. on the next page.

The maintenance cycle is divided into two orbits namely: (1) The outer cycle which represents the managerial processes in the maintenance organization; and (2) The inner cycle which involves the operational and technical processes. In this dissertation we will focus on maintenance strategy setting, one of the components in the inner cycle.

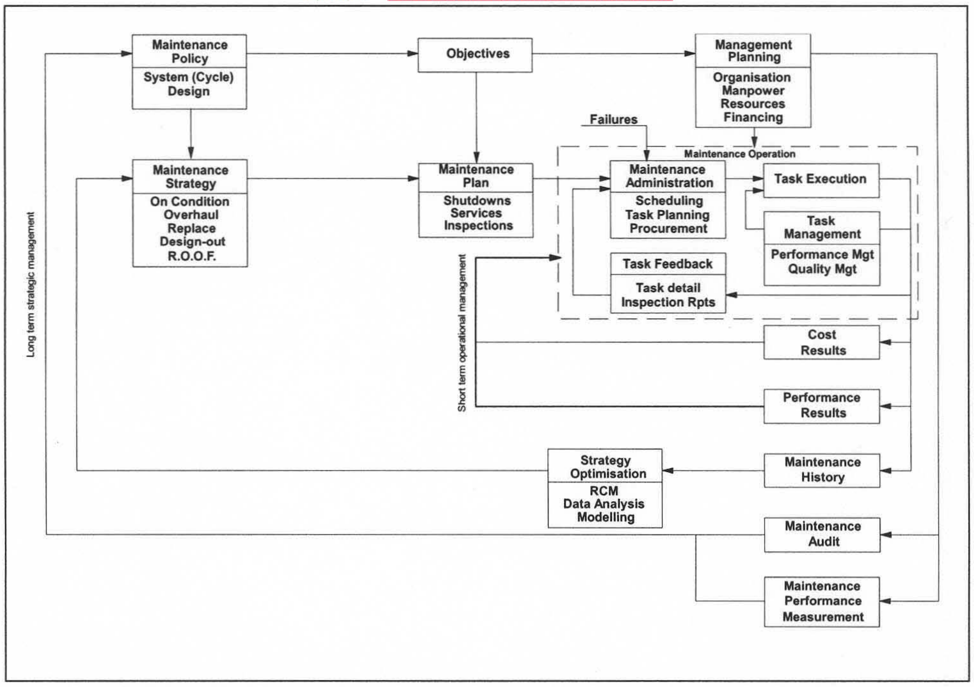

Figure 1.1.: *The Maintenance Cycle (taken from Coetzee[22])*

Several maintenance strategies are identified and are illustrated in the diagram in Figure 1.2. below. This dissertation is on an advanced type of preventive maintenance strategy which combines scheduled renewal on the use based maintenance branch and condition monitoring, specifically vibration monitoring, on the predictive maintenance branch. Scheduled renewal is defined as the complete renewal of a system or component due to replacement or complete overhaul. Only scenarios where complete renewal of a system or component takes place after failure are considered in this dissertation. Since this type of scenario is usually (not always) associated with components and not systems there will only be referred to components in the text when renewal situations are discussed.

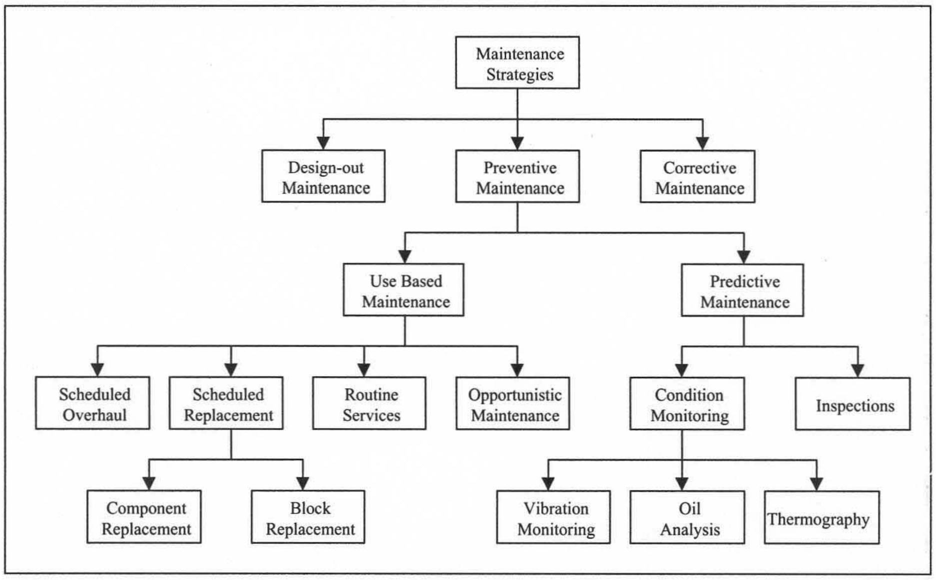

Figure 1.2: *Maintenance Strategies (adapted from Coetzee[22])*

Although the science of scheduled renewal, i.e. determining the optimal economic service life of a component before renewing it, and the technological marvel of vibration monitoring seem to be very close in the maintenance strategy tree, it is not nearly the case. A statistical approach is usually followed to calculate the optimal time for use based preventive renewal from time failure data while a methodology based on previous experience in the form of empirical rules and benchmarks are used to predict optimal renewal times with vibration monitoring.

Both mentioned strategies are well established in practice and have proven themselves in many different situations although some disadvantages and deficiencies are experienced. It is believed that if a single scientific preventive maintenance strategy setting technique can be found that is able to overcome the current shortcomings and incorporate both techniques' advantages, much improved maintenance renewal decisions will be the result. This is the main goal of this dissertation.

The goal of preventive maintenance strategy setting is always to minimize the total life cycle cost of a component, i.e. to determine the optimal economical renewal instant for a component and not necessarily to predict the time to failure of the component. This fact influences the approach of this research project immensely.

In chapter 1 the two strategies are introduced in enough detail to be able to understand the objectives for the research project. After this discussion it would also be possible to identify and describe the most promising and logical research route that will lead to the achievement of the objectives.

2 Use Based Preventive Renewal

Use based preventive renewal is all about determining the optimal economic instant to replace a component preventively. This optimal instant could be measured in any use parameter for instance running hours, tonnage handled or production throughput. Running hours (also called working age or time) is most often used. Failure data of identical components in the past is used to estimate optimal preventive renewal of future components. This strategy is only a feasible possibility if failure data of a specific component is available.

2.1 Failure Data

Since time is usually used to describe a component's age this discussion will be on *time failure data* but the same principles apply for any other use parameter. Different types of time failure data are identified, i.e. time failure data sets with different inherent characteristics. This fact is very important since different analysis techniques are used for every type of data set. Classification of data set types is a field on its own and details regarding the topic are beyond the scope of this dissertation, but logical steps that are taken in the process of classification are mentioned briefly. These steps are discussed with the aid of the flow diagram in Figure 2.1. on the next page.

STEP (i): *Order T_i's (times to failure) chronologically*

Before the classification process starts, the failure times should be sorted in its original chronological order of occurrence. This has to be done to discern trends in the data.

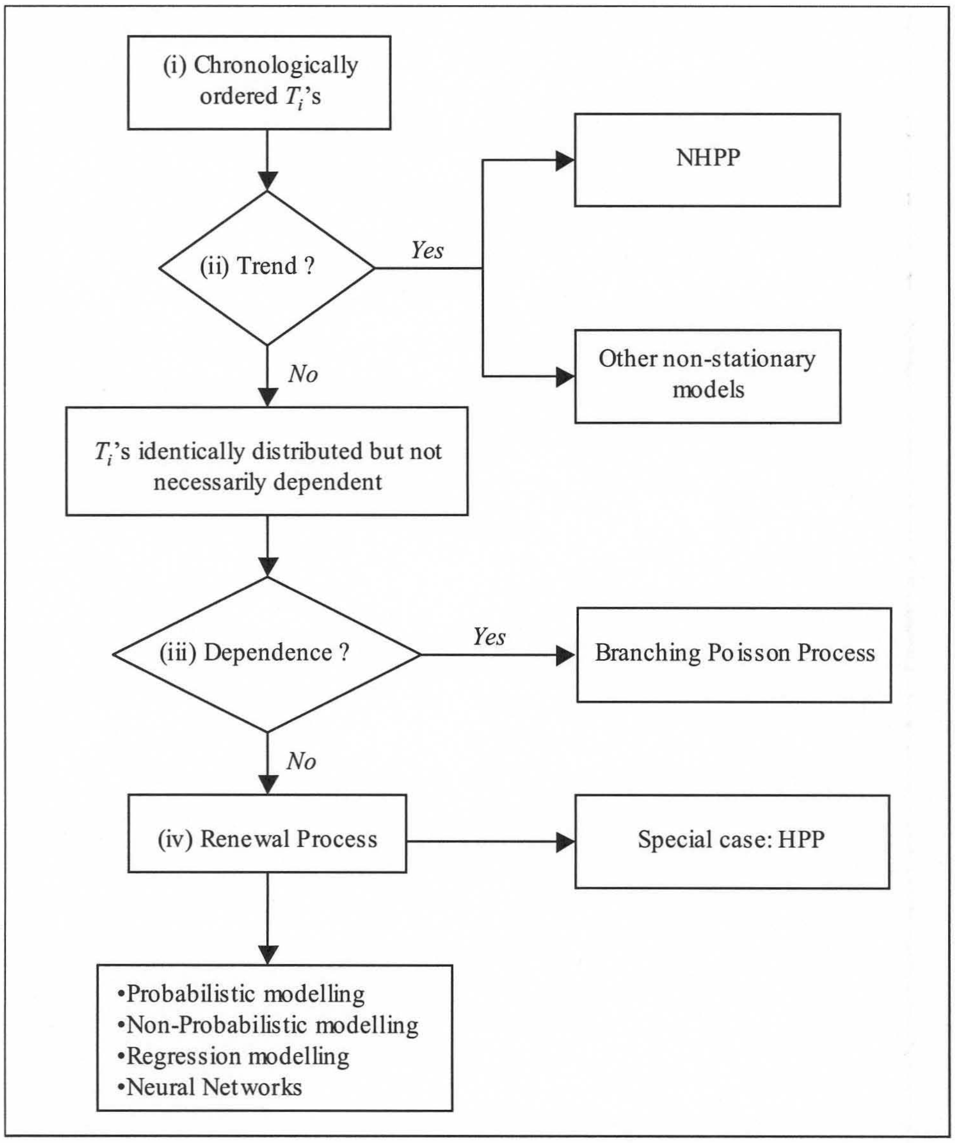

Figure 2.1: *Diagram to aid with classification of data (Adapted from Ascher and Feingold[15])*

STEP (ii): *Test for a trend in data*

A number of techniques exist to recognize trends in data. Graphical techniques include (a) plotting cumulative failures versus cumulative time on linear paper; (b) estimating average failure rate in successive time periods; and (c) Duane plots.

Mathematical techniques proposed by Laplace (1773), Bartholomew (1955), Bates (1955), Boswell (1966), Cox and Lewis (1966) and Boswell and Brunk (1969) are typically used to determine the existence of trends. Note that these techniques do not all apply for every possible scenario and that certain techniques are only applicable to certain situations.

If a trend is found in a data set, the data set can be modeled by non-stationary models like the Non-Homogeneous Poison Process or models based on a sequence of independent but not identically distributed random variables. If not, it implies that the data is identically distributed, but no conclusion can be made about the dependence of the data and therefore step (iii) has to be evaluated.

STEP (iii): *Test for dependence*

Although very important, this step is often neglected. Ascher and Feingold[15] suggest three reasons for this:

(a) The need for a large number of times to failure or T_i's.
(b) The complexities involved in implementing and interpreting these tests.
(c) An almost complete lack of understanding of the need to perform this type of test.

The most straightforward way of testing for dependency of successive T_i's, is by means of the sample correlation coefficient of lag j, called C_j. (Lag j refers to the correlation between T_i and T_{i+j}, $i=1,2,\ldots,m-j$, $1<i+j<m$ where m is the total number of observed T_i's).

If the data turns out to be dependent, a Branching Poison Process is suitable to represent it. This type of data set is encountered in situations where primary failures has a positive probability of triggering one or more subsidiary failures. An example will be where a primary failure causes one or more secondary failures which are not detected until after the system is put back into operation. Another example is where repair to a system is not done properly and the system fails again after repair because of the same undetected problem.

It might be argued that subsidiary failures are not true functional system failures and should not be dealt with in the same way as with true functional failures. In practice however, subsidiary failures will also cause a system to be out of operation and it must thus be seen as true functional failures.

STEP (iv): *Renewal process*

If the diagram leads to step (iv), the data is independent and identically distributed (i.i.d.) and is generated by a renewal process. Because it is independent and identically distributed we can infer that a previous failure of a system, sub-system or component do not influence the variation of the future life of the system, sub-system or component and consecutive failures arise from the same failure distribution.

A renewal process that generates i.i.d. data is defined as a non-terminating sequence of independent and identically distributed non-negative variables $T_1, T_2, \ldots, T_n$ which with probability one, are not all zero. In this case the typical variable is time to failure and it is usually the time to failure of sub-systems or components. Failure data of sub-systems or components is in general more likely to be part of a renewal process than the failure data of systems, although this is not a rule.

In practice a renewal process is also referred to as an "As-Good-As-New" scenario. If a component fails, it is replaced or completely overhauled and the component is as-good-as-new again. As stated earlier will this dissertation be limited to the event of complete renewal. This type of data is analyzed with methods referred to as renewal theory which comprises probabilistic, non-probabilistic or regression models.

2.2 Incomplete Observations or Censored Data

Up to this point it was assumed that time failure data was compiled from observed times to failure of a specific component. This is not always true and it could be that observations were incomplete. Incomplete observations are frequently encountered in data sets. An observation that is only known to have occurred before a certain time, within a certain interval of time or after a certain time is called a censored observation. These cases are called left-censored, interval censored or right-censored observations respectively. Censored observations contain valuable information and have to be included in failure time models.

Two distinct types of right-censoring are identified. The first type, Type I, occurs when n components are put on test for a fixed time c and at the end of time c a number of components are still operational. In this case it is known that for the failed items indexed by i, $T_i \leq c$ and for the components still operational indexed by j, $T_j > c$. A different type of right-censoring occurs where observation ceases after a pre-specified number of failures. This is called Type II-censoring and the right-censored time (or times in the case where all items were not started up on the same time) is (are) not known before observation starts.

Only Type I right-censored observations (also called suspensions) are considered in this research project. The terminology will remain the same for the remainder of the text as before, i.e. time failure data or failure times refers to both failures and suspensions.

2.3 Renewal Theory

Renewal theory comprises of estimating component reliability from recorded failure times of replaced components and calculating the renewal time which minimize the total average life cycle cost of future components. Four important reliability concepts are defined in renewal theory to form the base of reliability modelling approaches used on i.i.d. data:

(i) The *failure density*, *f*, which is the probability of component failure at a specific instant.

(ii) The *cumulative failure density*, *F*, or the *unreliability function* which gives the probability of component failure up to a certain instant.

(iii) The *reliability function*, *R*, which gives the probability of component success up to a certain instant.

(iv) The *instantaneous failure rate*, *force of mortality* or *hazard rate*, *h*. This function gives the probability of component failure in a short interval $(t, t+\Delta t]$ provided that the component is still operational at time t.

Three different mathematical approaches suitable to approximate the reliability functions are identified: (1) A probabilistic modelling approach; (2) A non-probabilistic modelling approach; or (3) A regression modelling approach. A fourth approach used to estimate the reliability of a component, part of a renewal process, is the utilization of neural networks. Neural networks do not strive to estimate the reliability functions but build dynamic models with observed data suitable to predict future events.

The four approaches are introduced below without concentrating too much on the underlying mathematics whereafter comments will be made about optimal renewal instant calculation.

2.3.1 Probabilistic Modelling Approach

Probabilistic modelling approaches are primarily based on the fact that failure times generated by a specific renewal process all arise from the same underlying distribution. Estimation of this underlying distribution or failure density function plays an important role in this type of approach. Another characteristic of probabilistic modelling is that only time (or any other use parameter) is used to represent the reliability functions and no concomitant information concerning failures is included. With these facts stated it is possible to define the formal reliability functions for this approach in terms of recorded failure times T_i and continuous time t.

2.3.1.1 Probability Density Function (PDF)

The probability density function, $f(t)$, denotes the positive probability of a failure T occurring within an interval dt. In probabilistic notation it can be expressed as $f(t)dt = P\{t \leq T \leq t + dt\}$ with $f(t) \geq 0$ for all t and the probability of all outcomes $\int_{t=0}^{\infty} f(t)dt = 1$.

2.3.1.2 Cumulative Probability Density Function (Unreliability function)

$F(t)$ is the cumulative probability of failures or the probability of component failure as a function of time, i.e. $F(t) = P\{T \leq t\}$ or $F(t) = \int_{t=0}^{t} f(t)dt$.

2.3.1.3 Reliability Function (Survivor function)

This function, $R(t)$, represents the probability of component success and is closely related to the unreliability function. In probabilistic terms is $R(t) = P\{T \geq t\}$ or $R(t) = 1 - F(t)$.

2.3.1.4 Hazard Rate (Instantaneous Rate of Failure)

The hazard rate, *h(t)*, is the most important and most valuable reliability function and because of its importance it is introduced in some detail. For this purpose we first define conditional probability. Suppose that one event, say *X*, is dependent on a second event, *Y*. We define the conditional probability of event *X*, given event *Y* as $P\{X|Y\}$. From the third axiom of probability is:

$$P\{X \cap Y\} = P\{X \mid Y\}P\{Y\} \tag{2.1.}$$

In equation (2.1.), $X \cap Y$ denotes both *X* and *Y* take place. This imply the probability that both *X* and *Y* occur is the probability that *Y* occurs multiplied by the conditional probability that *X* occurs, given the occurrence of *Y*. Equation (2.1.) can be written as the definition of conditional probability:

$$P\{X \mid Y\} = \frac{P\{X \cap Y\}}{P\{Y\}} \tag{2.2.}$$

For further discussion on conditional probability, see Lewis[17].

Let *h(t)dt* be the probability that the system will be in the failed state at some time $T < t + dt$, given that it has not yet failed at $T = t$. From the definition of conditional probability we have:

$$h(t)dt = P\{T < t + \Delta t \mid T > t\} = \frac{P\{(T > t) \cap (T < t + \Delta t)\}}{P\{T > t\}} \tag{2.3.}$$

The numerator on the right-hand side of (2.3.) is an alternative expression for the probability density function, i.e.:

$$P\{(T > t) \cap (T < t + dt)\} \equiv P\{t < T < t + dt\} = f(t)dt \tag{2.4.}$$

Combining equations (2.3.) and (2.4.) then yields:

$$h(t) = \frac{f(t)}{R(t)} \tag{2.5.}$$

For an increasing hazard rate a component has an increasing probability to fail and use based preventive renewal will be a definite option to consider, although costs will have the final say. Preventive renewal will only be used if the total cost of a failure is considerably higher than the total cost of preventive actions. If equation (2.5.) yields a constant hazard rate, the component is said to have a random shock failure pattern because the risk of failure of the component remains the same throughout the component's life. Corrective renewal will be the first option to consider for this case, i.e. repair only on failure. Corrective renewal will also most probably be used for a component with a decreasing hazard rate, since the probability of component failure becomes less as time progresses. It should be kept in mind however that condition based preventive renewal, like vibration monitoring, could be used for any shape of the hazard rate. Clearly the hazard rate has enormous value.

It is important to note that the reliability functions as defined above for the probabilistic approach are all related and if one of the functions is determined any other function can be derived.

2.3.1.5 Estimation Methods

There are several appropriate techniques that can be used to estimate the reliability functions. These techniques are discussed below.

- Parametric distributions. A continuous parametric distribution can be fitted to the data, usually to represent the failure density function. The Weibull distribution is often used for this purpose but exponential, normal, hyper-exponential or log-normal distributions are used in certain cases as well. See Coetzee[22].
- Techniques based on partial distributional knowledge. Certain properties of a data set can often be assumed before analysis, for instance an increasing hazard rate, because of observed physical properties of the component which generated the data set. This allows for slightly simplified techniques to fit parametric distributions on a data set. Ascher and Feingold[15] describe these simplifications briefly with references.
- Hazard plotting. This graphical procedure can be used to fit an appropriate hazard rate to data without any analytical techniques. It is performed on distribution specific paper, usually Weibull probability paper, is fast and easy to use and is often taught to semi-skilled analysts in the industry. See Nelson[16].
- Bayes's methods or Bayesion theory. Bayesian statistical inference is based on a subjective viewpoint of probability. This subjective viewpoint is often referred to as the "degree of believe" in the

behavior of a certain parameter, which is considered as a random variable. The subjective viewpoint is captured by a specified *prior distribution* based on prior knowledge about a parameter. This prior distribution is then updated with the aid of Bayes's theorem to a *posterior distribution* after new observation of the parameter of interest. See Hines and Montgomery[26].

- Multivariate models. This type of modelling is used more at a system level than at a simple component level. It is assumed that the components in the system as well as the system itself behave according to a renewal process and that more than one type of failure is associated with the system. Examples are a system of components in parallel where a number of components have to fail before system failure or components in series where failure of any component causes system failure. The reliability functions for this type of modelling are similar to those outlined above except that it is expressed in terms of the joint probabilities of different occurrences of failure. See Crowder et. al.[27].

All of the above mentioned techniques have the ability to accommodate incomplete observations or suspensions by allowing for it in parameter estimation procedures like the maximum likelihood method.

2.3.2 Non-probabilistic Modelling Approach

Non-probabilistic models or natural reliability estimators are the direct statistical analog of the probabilistic models described in the previous section. Natural reliability estimators represent the reliability functions in a discrete manner based on observed failure times and do not recognize the existence of an underlying distribution as a primary assumption. As is the case with the probabilistic approach, the non-probabilistic approach only models the primary use parameter, thereby excluding concomitant information on failures. A brief introduction to the reliability functions in their non-probabilistic forms is given below.

2.3.2.1 Failure Density Function

To estimate the failure density function in a discrete manner is not non-probabilistic in the true sense of the word since we are estimating the underlying distribution of a data set. It is listed here however because no probabilities are predicted with the discrete function but probabilities are simply reproduced in retrospect.

$$\begin{aligned} f(t) &= \frac{\frac{\Delta n}{N}}{\Delta t} \\ &= \frac{1}{N}\frac{\Delta n}{\Delta t} \end{aligned} \tag{2.6.}$$

In (2.6.) Δn denotes the number of failures in the interval $[t,t+\Delta t]$, N the total number of failure observations and Δt an appropriate time length.

2.3.2.2 Reliability and Unreliability Functions

If no incomplete observations are present in the data, the reliability or unreliability function can be estimated with:

$$R(t) = 1 - F(t) = 1 - \frac{\text{Number of failures up to time } t}{N} \tag{2.7.}$$

This function was generalized by Kaplan and Meier[25] to handle censored observations.

2.3.2.3 Hazard Rate

Without making an assumption on the underlying distribution in the data, the hazard rate can be obtained by:

$$h(t) = \frac{1}{n(t)}\frac{\Delta n}{\Delta t} \tag{2.8.}$$

$n(t)$ is the population surviving up to time t.

Non-probabilistic estimators act as very useful guidelines on the form of the reliability functions before having to decide on a continuous distribution to fit to the data if a probabilistic approach is going to be used.

2.3.3 Regression Modelling Approach

Regression modelling of failure data can be seen as a hybrid between the previous two approaches but it has enough unique features to be recognized as a solitary third approach to renewal data modelling. Two clear properties define the regression modelling approach:

(a) Regression models do not use the existence of an underlying failure distribution as primary assumption but immediately recognize the being of the survivor function or hazard rate, similar to non-probabilistic models.
(b) Not only is the primary use parameter modeled by regression models but also concomitant information surrounding failures or covariates.

Regression models found in the literature are introduced below according to the date of first introduction without mathematical details.

2.3.3.1 Accelerated Failure Time Models (AFTM) - 1966

Accelerated failure time models or accelerated life models strive to estimate the survivor function of a component as a function of its *accelerated* life. The accelerated life is a modified use parameter for a component that is determined by the influence of covariates on the original use parameter. Covariates accelerate (or decelerate) the predicted arrival of failures, thereby allowing for the effects of circumstantial influences surrounding failures.

2.3.3.2 Proportional Hazards Models (PHM) - 1972

For proportional hazards models, the hazard rate is of primary concern. The hazard rate is determined by a baseline hazard rate which is a function of time only and a functional term dependent on time and covariates which acts multiplicatively on the baseline hazard rate. The multiplicative effect of the covariates on the baseline hazard rate implies that the ratio of the hazard rates of any two items observed at any time t associated with two different covariate sets will be a constant with respect to time and proportional to each other.

2.3.3.3 Prentice Williams Peterson Model (PWP model) - 1981

This model is a major elaboration of the PHM and although very little research has been done on this model it is considered to have enormous potential. It is suitable to model data sets generated by both renewal processes and repairable systems and allows for the effects of covariates on the failure process. It also takes scenarios into account where more than one failure have occurred on a specific unit and it is possible to stratify data, i.e. group data based on influential differences.

2.3.3.4 Proportional Odds Models (POM) - 1983

The odds of a failure occurring is defined as the ratio between the unreliability function and the reliability function. For this model, a value φ is introduced as the ratio between the odds of a failure occurring estimated when considering the influence of covariates and the odds of a failure occurring estimated without considering covariates. The model assumes that the covariates has a diminishing effect on a component as time increases, i.e. $\varphi \to 1|_{t\to\infty}$.

2.3.3.5 Additive Hazards Models (AHM) - 1990

For additive hazards models the hazard rate is also of primary interest as in the case of PHM. This time the hazard rate is constructed as the sum of a baseline hazard rate which is a function of time only and a functional term dependent on time and covariates. The direct analogue of this model could also be used in repairable systems to model *failure rate* in an additive manner.

2.3.4 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 modelling 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 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:

1. Processing units, where each processing unit has a certain activation level at any point in time.
2. Weighted interconnections between the various processing units which determine how the activation of one unit leads to input for another unit.
3. An activation rule which acts on the set of input signals at a unit to

produce a new output signal, or activation.

4. Optionally, a learning rule that specifies how to adjust the weights for a given input/output pair.

Recently attempts were made to apply neural networks in the reliability modelling field[28,29,30]. Failure data with covariates were used as processing units to estimate or teach the neural network and additional data was then used as inputs to predict future outputs. The results were compared to the predictions of proportional hazards models and accelerated failure models and proved to be very promising.

Neural networks are not considered to be an "official" renewal modelling approach by failure data analysts because very little research has been done on its application in reliability modelling up to date. It is listed here however because of its enormous future potential. No further reference to neural networks will be made in this dissertation.

2.4 Optimal Use Based Preventive Renewal Decision Models

As stated in the introduction the aim for optimal decision making in renewal theory is not to predict the exact time to failure of a component but to minimize the total life cycle cost of a component. This is done with the aid of the described reliability functions estimated by single variable techniques or regression models.

Two very important quantities in optimal decision making are the cost of unexpected replacement or failure of a component, C_f, and the cost of preventive replacement C_p. It is normally much more expensive to deal with an unexpected failure than it is to renew preventively. A balance has to be obtained between the risk of having to spend C_f and the advantage in the cost difference between C_f and C_p without wasting useful remaining life of a component. The optimum economic preventive renewal time will be at this balance point.

A similar argument to the one above can be followed if availability of a component is more important than cost. Only this time the downtime due to unexpected failure, T_f, and the downtime due to preventive replacement, T_p, are weighed against each other to determine an optimum.

2.5 Concluding Remarks

Use based preventive renewal has established itself as a maintenance strategy with the ability to bring huge cost savings about if implemented correctly. This statement is supported by countless instances in the industry where use based preventive renewal is practised successfully.

When considering this preventive strategy it is important to note that it is extremely dependent on accurate failure data, which is often not easy to find in practice because of negligence in failure data recording processes. This requirement also implies that use based preventive renewal can only be applied after several failures of a component have occurred.

Renewal theory as outlined above is familiar to most reliability engineers, especially the probabilistic and non-probabilistic approaches. Regression modelling in renewal theory is still in its infant stage in the reliability world although there are no doubts about is potential. The probabilistic and non-probabilistic approaches have one prominent disadvantage in the fact that concomitant information is not included in models but this problem is being addressed in regression models.

3 Preventive Renewal Based on Vibration Monitoring Predictions

By analyzing the vibrational behavior of a component, an enormous amount of information about the component's condition can be learned. This fact has been proven over and over in the past and it has driven development on theoretical vibration analysis techniques up to a very high technological level. Neural networks, self organizing maps, fuzzy logic, time series analysis, coherence, frequency band energy methods, trending, correlation and many other techniques were developed for decision making in vibration monitoring or were applied to vibration monitoring as a result. Very little of this high level vibration technology is found in the industry however and often only absolute basic vibration techniques are used in condition monitoring programs.

An overview of techniques often used in preventive renewal based on vibration monitoring is presented in this section. The term "vibration monitoring" here, refers to the typical vibration monitoring practices found in the industry and not to the total complex field.

3.1 Methodology

Vibration monitoring is a condition monitoring maintenance strategy which relies almost entirely on the current condition of the component as determined by vibration parameters to make renewal decisions. Benchmarks or envelopes are specified for every measured vibration parameter and if a parameter or certain parameters exceed these specified levels, the component is renewed. Reliable benchmark levels are usually specified by the component's manufacturer although experience could optimize these levels to a certain extent. Graphical aids like vibration severity charts and waterfall plots are often used to assist the vibration monitoring person in making renewal decisions.

3.2 Vibration Parameters

In most cases only time domain vibration parameters are measured to evaluate the condition of a component. These include peak signal values, RMS values, Crest factors and Kurtosis. Frequency domain analyses commonly found are power spectral density analysis, cepstrum and high-frequency resonance techniques.

3.3 Shortcomings of Preventive Renewal Based on Vibration Monitoring

A number of shortcomings that have an influence of this research project are outlined below.

3.3.1 Lack of Comparative Means Between Current Vibration Condition and Past Vibration Behavior

Very often in practice only short term changes in vibration levels are considered to estimate component reliability, i.e. only the vibrations measured during a specific component's life time are used to predict useful remaining life. This is usually done with the aid of waterfall plots where different vibration levels are presented in a user-friendly, graphical manner such that it is easy to recognize trends in vibration behavior.

No verified or established means exist to consider long term vibration behavior in reliability estimations. Long term vibration behavior refers to vibration histories recorded from similar components that have failed under equivalent conditions in the past. Long term vibration behavior of components certainly holds extremely valuable information in terms of current component reliability since vibration conditions during a component's life tend to repeat itself in subsequent components.

The lack of a scientific technique with which long term vibration information can be incorporated in present component reliability estimation is considered to be a shortcoming of preventive renewal based on vibration monitoring.

3.3.2 Significance of Vibration Parameters

Numerous vibration parameters are usually measured and evaluated when monitoring the condition of a component as discussed in (3.2.). In very few cases all of the measured parameters are significant in the failure process and often renewal decisions are made based on the level of a parameter totally insignificant in the failure process.

Up to date, it is impossible to identify vital parameters in the failure process if only the current vibration behavior is considered. This is a second major deficiency of vibration monitoring.

3.3.3 Calculation of Optimal Renewal Instant

Vibration monitoring is definitely not perfect as a predictive preventive maintenance strategy. A perfect predictive preventive maintenance strategy would be able to determine the exact length of a component's remaining life. No such method exists. Unexpected failures of components still do occur regardless of the fact that the vibration levels are monitored and unexpected failures are normally very expensive relative to preventive replacements. Thus, renewal decisions based on vibration monitoring do not bring into account the risk of an expensive unexpected failure or the possibility of loss of useful remaining life due to premature renewal.

3.3.4 Lack of Commitment towards Vibration Monitoring

In general there was found during this research project that there is a lack of commitment towards vibration monitoring in the industry. In many cases expensive vibration monitoring equipment is used as the flagship of the maintenance department although inspections are done very irregularly and not recorded properly. Often the information supplied by vibration monitoring is totally disregarded when a decision has to be made and experience or intuition is relied on. Even if vibration information is considered, the final decision is frequently left to the discretion of the vibration technician.

It does not matter how technologically advanced vibration monitoring is, if it is not practiced correctly meaningful results are impossible to obtain. This is a maintenance management issue and will not be addressed in the dissertation but this shortcoming could obviously have a huge effect on results from this dissertation.

4 Integrating Use Based Preventive Renewal and Vibration Monitoring

There is very little doubt as to the enormous economical advantages that preventive renewal have in maintenance engineering, whether the renewal instant is determined by use based statistics from the past or present condition monitoring (including vibration) information. The discussion above supports this statement but also identifies shortcomings in current practices where these approaches are used. A method with the ability to integrate the principles of use based preventive renewal and renewal based on vibration monitoring could potentially overcome the mentioned shortcomings while encompassing all the advantages of both approaches. Successful identification and implementation of such a method is the objective for this dissertation. The formal research objectives are:

(i) Combining used based preventive renewal principles and preventive renewal based on vibration monitoring to make more appropriate renewal decisions than with either one of the mentioned techniques alone.

(ii) Verifying the theory used to achieve (i) with data obtained from the industry.

The route to the objective certainly runs through use based regression models with measured vibration parameters as covariates. From the discussion above it should be evident that no other logical route exist to approach the problem since the probabilistic and non-probabilistic approaches are lacking to handle covariates and almost no research has been done on neural networks in reliability. The research area will thus immediately be narrowed down to the above-mentioned route. The strategy to be followed to the objectives is outlined below.

4.1 Literature Study

A thorough literature study has to be done on regression models suitable to model failure data generated by a renewal process to become acquainted with existing techniques and models. The aim with the literature study is not to go into the details of the various models but rather to become aware of the abilities of the different techniques.

4.2 Identification of Most Suitable Model

After an appropriate literature survey it would be possible to identify the model most suitable to integrate used based preventive renewal and preventive renewal based on vibration monitoring. This means that all the advantages and disadvantages of the various models considered in the literature study will have to be balanced to determine the best model.

4.3 Thorough Study on Chosen Model

To be able to implement the chosen model successfully an in-depth study on the mathematics of the chosen model will be done. This study will range from the original proposal of the model to optimal decision making, using the model.

4.4 Practical Evaluation of the Model

For this research project to have worth in the reliability modelling world, the theoretical results will have to be evaluated in practice. Data recorded in industrial situation will thus be collected and modeled to prove the success of this dissertation.

Chapter 2

Regression Models in Renewal Theory

1 Introduction

In Chapter 1, five regression models in renewal theory were identified that have the potential to lead to the dissertation objectives. In this chapter the results of a thorough literature study on these models are discussed.

It was discovered that all the models have the same broad structure. First, a baseline function that describes the component's reliability as a function of time (usually the survivor function or hazard rate function) is estimated with either parametric or non-parametric techniques. Secondly, a functional term dependent on time and covariates is allowed to influence the baseline function (usually by multiplication) to estimate the total reliability of the component.

Throughout this chapter, the functional term is referred to as λ, which is a function of time and covariates, i.e. $\lambda(\overline{z(t)})$. Let $\overline{z(t)}$ denote a column vector of m measured covariates or $\overline{z(t)} = [z_1(t)\, z_2(t) \ldots.\, z_m(t)]^T$. For the sake of generality, covariates are assumed to be functions of time, although covariates may be time-independent. Further, suppose that $\overline{\gamma}$ is a row vector of regression coefficients associated with a specific model's covariate vectors i.e., $\overline{\gamma} = [\gamma_1\, \gamma_2 \ldots..\, \gamma_m]$, estimated during model fitting procedures. The terminology is followed closely, except for the PWP model where some additional variables are needed to describe the model.

Each model is first introduced in mathematical terms and then some comments are made about the model's abilities, deficiencies and applicability. After the discussion of the different models, the model most suitable for this research project is selected with a proper motivation for the selection.

2 Accelerated Failure Time Model

The accelerated failure time model was introduced by Pike[31] in 1966 and is considered as the second most popular regression model used in renewal theory today. It is a fully parametric type of model and strives to estimate the survivor function. The model allows covariates to influence expected life time of a component directly, in a multiplicative manner.

2.1 Mathematical Model

Pike[31] presents the model as follows:

$$R(t,\overline{z(t)}) = R_0(\lambda(\overline{z(t)}) \cdot t) \tag{2.1.}$$

In probabilistic terms the model can be written as:

$$R(t,\overline{z(t)}) = P\{T \geq t \mid \overline{z(t)}\} \tag{2.2.}$$

$R_0(t)$ is a parametric baseline survivor function estimated without considering covariates. The AFTM is then constructed by allowing covariates to influence life time by the functional term, $\lambda(\overline{z(t)})$. In (2.1.), it is required that $\lambda(\overline{0}) \equiv 1$ for the case where covariates do not have an effect on life time and $\lambda(\overline{z(t)}) > 0$ where covariates do play a role.

A popular form of the functional term is the log-linear function, i.e. $\lambda(\overline{z(t)}) = \exp(\overline{\gamma} \cdot \overline{z(t)})$, where $\overline{\gamma}$ is a vector of regression coefficients. In this case, $\lambda(\overline{z(t)}) > 1$ accelerates and $\lambda(\overline{z(t)}) < 1$ decelerates the rate at which a component moves through time with respect to the baseline survivor function.

Leemis[32] derived the general hazard rate function for the AFTM as:

$$h(t,\overline{z(t)}) = \lambda(\overline{z(t)}) \cdot h_0(\lambda(\overline{z(t)}) \cdot t) \tag{2.3.}$$

Newby[33] suggests the maximum likelihood method to estimate the model parameters although the method of moments have also been used successfully.

2.2 Comments

The theory of AFTMs has been developed in detail over the 33 years of the model's existence. Numerous theoretical publications on model estimation techniques, goodness-of-fit tests and extensions for the model to suit repairable systems reliability are found in the literature. A very good example of such a publication is Lin et. al.[35].

Not only has the AFTM a sound theoretical base, but it has also been applied widely on failure time data, especially in biomedical applications and more recently in reliability situations. Four relevant publications proving the abilities of the AFTM are:

1. Martorell et. al.[36]. In this paper the AFTM is used successfully to estimate the useful remaining life of nuclear power plants. Results are compared to methods not incorporating covariates. The authors conclude that this model is a useful maintenance management support tool.
2. Addison et. al.[37] used the AFTM to model unemployment duration data with attributes like employee age and profession. The results are compared to Cox's proportional hazards model (considered later in this chapter).
3. Shyur and Luxhoj[38] use Cox's PHM, AFTM and neural networks to model data obtained from ageing aircraft with success.
4. Publications where fatigue crack growth is modeled by the AFTM are encountered frequently in the literature. Principles of fracture mechanics as applied in fatigue crack growth are very suitable for the parametric approach of AFTMs. See [27] and [33].

Solomon[34] identified several cases where the AFTM was specified inappropriately because of many illustrations where accelerated failure time models seemed to arise naturally in practice. Newby[33] reports some of these misspecifications as well. Crowder[27] gives guidelines as to when the AFTM is appropriate.

The AFTM is established in regression type failure analyses although it has certain limitations. Newby[33] describes this model to be "an effective alternative to the proportional hazards model in appropriate cases" after a thorough study on both PHMs and AFTMs in 1988, thereby suggesting that the PHM is superior to the AFTM.

3 Proportional Hazards Model

Failure time data analysis underwent a total revolution after Cox[2] proposed this model in 1972. It was first intended for use in biomedicine but was soon modified to be suitable for the field of reliability. PHMs model the hazard rate of a component as the product of a baseline hazard rate dependent of time only and a functional term dependent on time and covariates.

The PHM was originally proposed as a semi-parametric model and regression parameters can be determined independently of the estimation of the baseline hazard rate although this only yields relative risks. For an absolute hazard rate, the baseline hazard rate has to be estimated. In general the PHM is used in its parameterized form to overcome numerical difficulties.

Extensions made to the original PHM by Prentice, Williams and Peterson[11] lead to the famous PWP model and extensions made by Pijnenburg[12] resulted in the additive hazards model, both considered later in this chapter.

3.1 Mathematical Model

The model is proposed by Cox[2] as:

$$h(t,\overline{z(t)}) = h_0(t)\cdot\lambda(\overline{z(t)}) \qquad (3.1.)$$

Analogous to AFTMs, the PHM consists of a baseline hazard rate, $h_0(t)$, which is influenced multiplicatively by a functional term $\lambda(\overline{z(t)})\geq 0$, thereby including the effects of covariates. Inspection shows that the total hazard rate is identical and equal to the baseline hazard rate when the covariates have no influence on the component's risk to fail.

The assumption of the multiplicative effect of the covariates on the baseline hazard rate implies that the ratio of the hazard rates of any two items observed at any time t associated with covariate sets $\overline{z_1}$ and $\overline{z_2}$, respectively, will be a constant with respect to time and proportional to each other, i.e. $h(\overline{\gamma},\overline{z_1})\ \alpha\ h(\overline{\gamma},\overline{z_2})$. This property is referred to as the proportional hazards property of the model.

There are several possible forms for the functional term $\lambda(\overline{z(t)})$. Some are: the exponential form, $\exp(\overline{\gamma}\cdot\overline{z(t)})$; the logarithmic form, $\log(1+\exp(\overline{\gamma}\cdot\overline{z(t)}))$; the inverse linear form, $1/(1+\overline{\gamma}\cdot\overline{z(t)})$; or the linear form, $1+\overline{\gamma}\cdot\overline{z(t)}$. The exponential form of the functional term is used most widely and then equation (3.1.) becomes:

$$h(t;\overline{z(t)}) = h_0(t)\cdot\exp(\overline{\gamma}\cdot\overline{z(t)}) \qquad (3.2.)$$

where the regression vector $\overline{\gamma}$ and the baseline hazard function $h_0(t)$ needs to be determined. Methods to estimate $h_0(t)$ for the semi-parametric model in (3.2.) involve maximum likelihood theory and can be found in [2],[3] and [6].

The PHM is often used in its fully parameterized form to increase numerical practicability with the aid of the Weibull distribution, which is very suitable to model failure time data. Parameterization is done by approximating $h_0(t)$ with the Weibull representation of the hazard rate, i.e.:

$$h(t;\overline{z(t)}) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \cdot \exp(\overline{\gamma} \cdot \overline{z(t)}) \qquad (3.3.)$$

where β and η are the shape and scale parameters of the Weibull distribution respectively. Parameters of (3.3.) are generally determined by maximum likelihood methods, see [5] for example.

3.2 Comments

Proportional hazards modelling was probably the greatest contribution to time failure data analysis up to date and it is still the most popular model of its kind. The model was used extensively in biomedicine after its introduction and since the mid 1980's the model has been accepted more and more in the reliability modelling world.

The theoretical foundation of proportional hazards modelling and its successes are well established in the literature. Since its introduction in 1972, countless papers have been published on more efficient estimation methods, goodness-of-fit tests and minor extensions to the model.

Application of the model to practical failure time data is predominantly found in the field of biomedicine until the mid 1980's, especially the semi-parametric PHM. The reason for the popularity of the semi-parametric PHM is because no assumption needs to be made about the baseline hazard rate. Younes and Lachin[39] sum this up by stating: "In biomedical applications, little is generally known *a priori* about the shape of baseline functions, and models that assume specific parametric shapes (such as the accelerated failure time model) can be difficult to justify."

Bendell[1] expressed his disappointment in failure data analysts in 1985 with: "Why is it, however, that recorded applications of the PHM to date have until recently been almost entirely associated with medical data?", thereby criticizing them for overlooking this very logical approach to life data analysis. During this time many more publications on the PHM in reliability applications were made with data obtained from, amongst others, majorettes[7], marine gas turbines and ships' sonar[8], valves in light water reactor nuclear generating plants[9] and aircraft engines[10]. More recently the model was used with great success to model the reliability of locomotive diesel engines in Canada by using types and quantities of foreign particles in the engine oil[5] as covariates with the failure time data. It was also used on aircraft and marine engine failure data[4].

The PHM has certainly made good ground in the field of reliability modelling and has proved itself to be an excellent support tool for maintenance renewal decisions.

4 Prentice Williams Peterson Model

Prentice, Williams and Peterson[11] made a major extension to the original Cox PHM in 1981, with a model which will be referred to as the PWP model. The PWP model extends the PHM to handle situations where a specific item (component or system) experiences multiple failures during its life time, by allowing for stratification of the data in the model. The model is defined in such a way that it is suitable to model data generated by repairable systems and data from renewal situations (situations where the hazard rate is of primary importance). This model has three dimensions (compared to the two of the PHM) namely age, covariates and stratum which makes the PWP model extremely powerful.

4.1 Mathematical Model

Because of its complexity, the PWP model will only be introduced in very general terms. Suppose a specific item can experience more than one failure during its life time. For the moment it is not important whether the item is repaired to the as-good-as-new or as-bad-as-old condition. Define u as the long term life variable, where $u=0$ at the item's initial startup and let $N(u)$ be the counting process of the multiple item failures. Every short term item life, i.e. operational period in-between failures, is referred to as a stratum, S, where $S=N(u)+1$ at any instant u. Thus, the item enters the i^{th} stratum at occurrence of the $(i-1)^{\text{th}}$ failure, where $i=2,3,...$ and enters stratum 1 at $u=0$ when $N(0)=0$. Also define t, the time from the most recent failure to the current time u. Let $Z(u)$ be the covariate process observed from $u=0$.

From the above it is possible to define a concept used in repairable systems theory, that of *rate of occurrence of failure* (ROCOF):

$$v(u) = \frac{\mathrm{d}}{\mathrm{d}u} N(u) \tag{4.1.}$$

Prentice, Williams and Peterson tried to estimate the ROCOF of items experiencing multiple failures with their model. They suggest two possibilities:

PWP model 1: v_1

Let the baseline function be a combination of the item's ROCOF and hazard rate by defining the function to be stratum-specific but dependent on long term time u. This leads to:

$$v_1[u \mid \{N(u), u \geq 0\}, \{Z(u), u \geq 0\}] = v_{0_s}(u) \cdot \exp(\overline{\gamma_s} \cdot \overline{z(u)}) \qquad (4.2.)$$

where $\overline{\gamma_s}$ is a vector of stratum-specific regression coefficients.

PWP model 2: v_2

For model 2 the baseline function is the stratum-specific hazard rate (as a function of t) of the S^{th} stratum:

$$v_2[u \mid \{N(u), u \geq 0\}, \{Z(u), u \geq 0\}] = h_{0_s}(t) \cdot \exp(\overline{\gamma_s} \cdot \overline{z(u)}) \qquad (4.3.)$$

Cox's PHM is the general case of v_2, where $h_{0_s}(\cdot) = h_0(\cdot)$ for all strata.

Prentice, Williams and Peterson used partial likelihood concepts similar to those used by Cox[2] to estimate the regression coefficients.

4.2 Comments

From the brief description above it should be clear that this model is extremely powerful. Surprisingly enough very little research has been done on this model up to date. Except for the original proposal of the model and an unpublished Ph.D. dissertation of Williams[40], only a few attempts have been made to utilize the endless advantages of this model. Two examples are Ascher[41], who investigated gas turbine engine reliability in 1982 with the PWP model and Dale[42] who has illustrated in 1983 how the model can be used for repairable systems reliability.

Authors who support the statement about the PWP model's extreme potential are, amongst others, Pijnenburg[12] and Ascher and Feingold[15]. Ascher and Feingold are of opinion that "the importance of the PWP model can scarcely be overemphasized".

5 Proportional Odds Model

The proportional odds model originated from epidemiological studies and was introduced by Bennett[43] in 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, does the effect of covariates in the POM model diminish as time approaches infinity. This diminishing property of the covariates means that the model is suitable for situations where a component adjusts to factors imposed on it or the factors only operate in early stages.

5.1 Mathematical model

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

$$\frac{F(\cdot)}{R(\cdot)} = \frac{1 - R(\cdot)}{R(\cdot)} \tag{5.1.}$$

This definition of 'odds' is used to introduce the POM:

$$\frac{1 - R(t,\overline{z(t)})}{R(t,\overline{z(t)})} = \varphi \cdot \frac{1 - R_0(t)}{R_0(t)} \tag{5.2.}$$

Equation (5.2.) states that the odds for a failure to happen under the influence of covariates are φ times higher than the odds of a failure without the effects of covariates. If φ increases, so does the probability of a shorter life time. Differentiation of (5.2.) with respect to time leads to:

$$\frac{h(t,\overline{z(t)})}{R(t,\overline{z(t)})} = \varphi \cdot \frac{h_0(t)}{R_0(t)} \tag{5.3.}$$

after using the coefficient rule. By rearranging the terms in (5.3.) and re-using (5.2.), a hazard ratio can be obtained:

$$\frac{h(t,\overline{z(t)})}{h_0(t)} = \varphi \cdot \frac{R(t,\overline{z(t)})}{R_0(t)} = \frac{1 - R(t,\overline{z(t)})}{1 - R_0(t)} \tag{5.4.}$$

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

Bennett derives the full unconditional likelihood for the model in his original paper to estimate the model parameters. Research done by Shen[44] provides more efficient estimation methods and methods to enable the model to handle suspended observations.

5.2 Comments

The POM has not been used very often in reliability modelling up to date although it has been fairly popular in biomedical data analyses since first publication. Its diminishing covariate effect property is probably the primary reason for its unpopularity in the reliability modelling field. It is argued that the effect of covariates describing components' reliability will seldom taper down close to failure or suspension.

6 Additive Hazards Model

Pijnenburg[12] proposed the additive hazards model in 1991. For this model a time dependent hazard rate is used as a baseline function and a functional term is added to the baseline function. Because of the addition, the functional term need not be positive as in the case of the PHM which immediately gives the model more flexibility. David and Moeschberger[45], Aranda-Ordaz[46] and Elandt-Johnson[47] are all of the opinion that this is a very valuable advantage. The AHM is also suitable for repairable systems when the ROCOF is modeled in an additive manner.

6.1 Mathematical Model

As before, suppose the time dependent hazard rate is denoted by $h_0(t)$ and the functional term which incorporates covariates is represented by $\lambda(\overline{z(t)})$. The additive hazard model is then:

$$h(t,\overline{z(t)}) = h_0(t) + \lambda(\overline{z(t)}) \qquad (6.1.)$$

There are many possibilities for the functional term. The most attractive form of the functional term is a polynomial. In practice the 1st order polynomial or straight line is used most often because of a lack of data availability, i.e. $\lambda(\overline{z(t)}) = \overline{\gamma} \cdot \overline{z(t)}$.

The behavior of the functional form gives the model its flexibility. If the measured covariates cause $\lambda(\overline{z(t)}) \geq 0$, it implies that the covariates have an accelerating effect on the wear out process of the component. A negative $\lambda(\overline{z(t)})$ would mean that the covariates are such that the expected wear out process is decelerated. If $\lambda(\overline{z(t)}) = 0$, the covariates have no additional effect on the wear out process.

Maximum likelihood can be used to estimate model parameters. Pijnenburg also provides a technique with which the additive assumption of the model can be evaluated.

6.2 Comments

Although Pijnenburg has shown the potential of the AHM with Davis'[47] bus engine data and Proschan's[48] aircraft air-conditioning system data, the model is not found very often in the literature. Pijnenburg is of opinion that "the AHM seems to be preferable to the PHM" for the mentioned data sets.

The model has not really been evaluated by the reliability modelling world and it is difficult to judge the model's potential based on publications.

7 Selection of Most Suitable Model

To be able to make an educated decision regarding the most suitable model for this research project, the following evaluation criteria were identified with which the different models could be compared (in order of importance):

(i) Theoretical foundation
(ii) Previous practical successes in reliability modelling
(iii) Potential to lead to the dissertation objectives
(iv) Achievability of numerical implementation
(v) Future potential in reliability modelling

With these criteria, a decision matrix can be constructed where different weights are allocated to the criteria and each model is evaluated with a mark out of 5 according to each criterion. The decision matrix is given on the next page in Table 7.1.

		Regression model				
Criterion	**Weight**	**AFTM**	**PHM**	**PWP**	**POM**	**AHM**
(i) Theoretical foundation	2^5	4	5	3	3	3
(ii) Previous practical successes in reliability modelling	2^4	5	5	1	1	3
(iii) Potential to lead to the dissertation objectives	2^3	4	5	4	1	4
(iv) Achievability of numerical implementation	2^2	4	4	2	4	4
(v) Future potential in reliability modelling	2^1	3	3	5	1	3
	Weighted total:	262	302	162	138	198

Table 7.1.: *Decision matrix*

The decision matrix shows that the proportional hazards model is the most suitable for this project and all research efforts will be focused on this model for the remainder of this dissertation.

Chapter 3

Proportional Hazards Modelling

1 Introduction

In the preceding chapters, the problem that motivated this research was described, possible solutions to the problem were considered and the Proportional Hazards Model (PHM) was selected as the most logical route to the solution of the problem. In Chapter 3, the PHM is considered in detail.

Cox proposed the original PHM in 1972, initially intended for biomedical applications[2]. The model was immediately considered to be a revolution in life data analysis and it is still applied on a wide variety of survival data today. It only started to become popular amongst reliability modelers in the early 1980's, especially because of the model's ability to model the hazard rate without making assumptions about its functional form (if used in its semi-parametric form). The PHM have ever since the 1980's been applied in diverse reliability applications, for example, component failures in a light water reactor plant[61], marine gas turbine and ship sonar[8], motorrettes[7], aircraft engines[50], high speed train brake discs[51], sodium sulfur cells[52], surface controlled subsurface safety valves[53], machine tools[54], diesel engines[5], aircraft cargo doors[55], rolling mills[56,57], power transmission cables[58,59] and components of a mine loader[60].

It is impossible to include all the theory, which was developed over the years of the model's existence in this dissertation. For this reason, the discussion is limited to the Proportional Hazards theory required to apply the model in practical situations although some attention is given to the original PHM for the sake of completeness. An optimal renewal decision making technique developed specifically for use with the model is considered as well.

Numerical methods required to implement the model in real life situations are also described in a fair amount of detail.

2 The Proportional Hazards Model

Before introducing the PHM, the probabilistic hazard rate $h(t)$ as derived in Chapter 1 is repeated here for convenience as equation (2.1):

$$h(t) = \frac{f(t)}{R(t)} \tag{2.1}$$

The probabilistic hazard rate is a function of time only, a property which seriously limits the function's abilities in reliability modelling as discussed previously. Cox addressed this problem in the PHM by assuming that the hazard rate of a component can be determined by the product of an arbitrary and unspecified baseline hazard rate, $h_0(t)$ and a functional term $\lambda(\overline{z(t)})$. The baseline hazard rate is a function of time only and the functional term is a function of time and covariates (concomitant or explanatory variables). (If the covariates are independent of time, the functional term is only a function of the covariates, i.e. $\lambda(\overline{z})$)[1].

$$h(t,\overline{z}) = h_0(t) \cdot \lambda(\overline{z(t)}), \tag{2.2}$$

where $\overline{\gamma}$ is a regression vector estimated during model fitting procedures.

There are several possible forms for the functional term $\lambda(\overline{z(t)})$. Some are: the exponential form, $\exp(\overline{\gamma} \cdot \overline{z(t)})$; the logarithmic form, $\log(1+\exp(\overline{\gamma} \cdot \overline{z(t)}))$; the inverse linear form, $1/(1+\overline{\gamma} \cdot \overline{z(t)})$; or the linear form, $1+\overline{\gamma} \cdot \overline{z(t)}$. The exponential form of the functional term has been used most often in reliability applications (and is the only form considered in this dissertation) and then equation (2.2) becomes:

$$h(t,\overline{z(t)}) = h_0(t) \cdot \exp(\overline{\gamma} \cdot \overline{z(t)}) \tag{2.3}$$

All theory on PHM presented in this dissertation will allow for time-dependent covariates, i.e. $\overline{z(t)}$, for the sake of generality.

This discussion deals with two forms of the PHM, namely (a) the semi-parametric PHM; and (b) the fully parametric PHM. When using the semi-parametric form, no assumption needs to be made about the shape of the baseline hazard rate when estimating the

[1] Although time, t, is used throughout as the unit of measure in this dissertation, any other suitable use parameter could be used instead, such as mileage or tons processed.

regression coefficients, although this only yields relative risks. To determine absolute risks, the baseline hazard rate has to be estimated first. This feature is considered to be a huge advantage even though some numerical difficulties are often encountered. The fully parameterized model makes use of a continuous distribution, most often the Weibull distribution because of its flexibility, for the baseline hazard rate, which makes it much more numerically tractable.

2.1 Assumptions of the PHM

The assumptions on which the PHM is postulated are best illustrated for the model with time-independent covariates:

a) Renewal times are independent and identically distributed.
b) All influential covariates are included in the model.
c) The ratio of any two hazard rates as determined by any two sets of covariates $\overline{z_1}$ and $\overline{z_2}$ associated with a particular component has to be constant with respect to time, i.e. $h(t,\overline{z_1}) \alpha\, h(t,\overline{z_2})$. (This assumption implies that the covariates acts multiplicatively on the hazard rate of the component).

Assumption (c) is illustrated graphically in Figure 2.1 below:

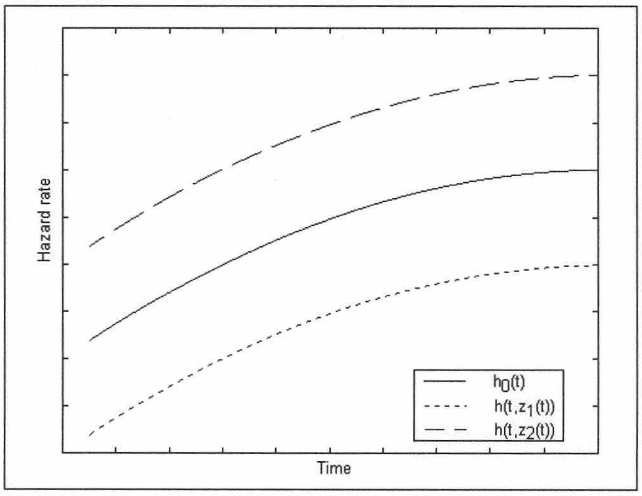

Figure 2.1: *Graphical illustration of PHM assumption (c)*

Figure 2.1 shows that the two hazard rates, $h(t,\overline{z_1})$ and $h(t,\overline{z_2})$, calculated from different covariate vectors associated with a specific component are proportional to each other with respect to time.

2.2 The Semi-Parametric PHM

The semi-parametric PHM has a very valuable attribute in that no assumption needs to be made about the baseline hazard rate of the model when estimating the regression coefficients. This means that relative risks of the item under consideration can be estimated without any knowledge about the time dependent failure behavior of the item, as is contained in the baseline hazard rate. The reason for the existence of this advantageous property becomes clear in the explanation of the estimation technique proposed by Cox with the proposal of the PHM, called the *method of partial likelihood.*

As an introduction to the method of partial likelihood, the *method of maximum likelihood* is discussed in general terms.

2.2.1 The Method of Maximum Likelihood

The method of maximum likelihood is a well known method for, amongst other uses, estimating regression coefficients and is used widely in the literature. It is important to describe this method before introducing partial likelihood since these methods are closely related and the method of maximum likelihood is used to determine the fully parametric PHM coefficients.

Likelihood refers to the hypothetical probability that an event, which has already occurred, would yield a specific outcome. This concept differs from that of probability in that probability refers to the occurrence of future events, while a likelihood refers to past events with known outcomes.

The method of maximum likelihood starts off by specifying the joint probability distribution function for events, in this case, failures. (For the moment the possibility of suspensions is omitted). We assume a sample of size n drawn from some probability distribution. Let $\overline{\boldsymbol{T}} = [\boldsymbol{T}_1, \boldsymbol{T}_2, ..., \boldsymbol{T}_n]$ denote failure times as random variables. The probability that $\boldsymbol{T}_i = T_i$ can be expressed informally as the probability density at T_i, i.e. $P\{\boldsymbol{T}_i = T_i\} = f(\boldsymbol{T}_i, \overline{\gamma}, \overline{z_i(t)})$. By using the law of total probability, the joint probability of all the random variables is given by:

$$P\{\boldsymbol{T}_1 = T_1, \boldsymbol{T}_2 = T_2, ..., \boldsymbol{T}_n = T_n\} = \prod_{i=1}^{n} f(\boldsymbol{T}_i, \overline{\gamma}, \overline{z_i(t)}) \qquad (2.4.)$$

Once the random variables have been observed, (2.4.) can be used to calculate the likelihood of events with all observed values fixed and $\overline{\gamma}$ the only unknown as shown in (2.5.) below:

$$L(\overline{\gamma}) = \prod_{i=1}^{n} f(T_i, \overline{\gamma}, \overline{z_i(t)}) \qquad (2.5.)$$

Equation (2.5.) can be extended to include suspended observations (see reference [3]):

$$L(\overline{\gamma}) = \prod_{i=1}^{n} f(T_i, \overline{\gamma}, \overline{z_i(t)})^{c_i} \cdot R(T_i, \overline{\gamma}, \overline{z_i(t)})^{1-c_i} , \qquad (2.6.)$$

for a data set with n renewals each at time T_i with $c_i = 1$ in case of failure and $c_i = 0$ for suspended observations. The value of $\overline{\gamma}$ that maximizes L is the most appropriate regression vector for the model since it maximizes the probability of occurrence of the observed data set. Numerical methods suitable to estimate the regression vector are discussed later in Chapter 3.

Maximum likelihood estimation is suitable for estimating both the semi-parametric and the parametric PHM although numerical difficulties are often experienced with this technique and the semi-parametric model.

2.2.2 Method of Partial Likelihood

In his original paper[2], Cox suggested parameter estimation for the semi-parametric PHM by maximizing an expression which he called *conditional likelihood*. This term gave rise to much discussion in the literature where the validity of referring to the term as conditional likelihood was argued by critics. For the estimation technique of Cox to be a conditional likelihood function it had to be a likelihood function based on the conditional distribution of data, given some statistic, which it was certainly not[62]. It was neither a marginal likelihood function because then it had to be based on the marginal distribution of some reduction of the data. In reply[63] Cox showed (somewhat informally[3]) that his technique was accurate and consistent with considerably less numerical difficulties. He also eliminated all confusion in terminology by renaming the technique to *partial likelihood* because of the fact that it was not a likelihood function in the usual sense.

Suppose a vector of random variables denoting failure times (for the moment suspensions are not considered) is observed, $\overline{\boldsymbol{T}} = [\boldsymbol{T}_1, \boldsymbol{T}_2, ..., \boldsymbol{T}_n]$, which comes from a probability density $f(\overline{T}, \overline{z_i(t)}, \overline{\gamma}, h_0(t))$. The baseline hazard function $h_0(t)$ is considered to be a nuisance function in this case. If a one on one transformation on the data in $\overline{T_i}$ is performed with auxiliary variables $A_1, B_1, ..., A_n, B_n$ such that $A^{(k)} = [A_1, ..., A_k]$ and $B^{(k)} = [B_1, ..., B_k]$ the likelihood of $A^{(k)}$ and $B^{(k)}$ is:

$$\prod_{k=1}^{m} f(b_k \mid b^{(k-1)}, a^{(k-1)}, \overline{\gamma}, h_0(t)) \cdot \prod_{k=1}^{m} f(a_k \mid b^{(k)}, a^{(k-1)}, \overline{\gamma}) \qquad (2.7.)$$

Cox[63] defined the second product on the right hand side of (2.7.) as the partial likelihood function since only a part of the joint probability density function is considered and the nuisance function $h_0(t)$ is eliminated. The mathematical proof of the validity of partial likelihood is discussed in reference [63]. See also reference [64] for a thorough discussion on the topic of partial likelihood.

For the PHM, the partial likelihood can be constructed as follows: as before, suppose a certain number of similar items have been renewed on k occasions of which n were failures at $T_1, T_2, ..., T_n$ with corresponding covariate vectors $\overline{z_1(t)}, \overline{z_2(t)}, ..., \overline{z_n(t)}$. Define the order statistic to be $O(t) = [T_{(1)}, T_{(2)}, ..., T_{(n)}]$ and the rank statistic to be $\overline{r(t)} = [(1), (2), ..., (n)]$. The order statistic refers to the $T_{(i)}$'s ordered from smallest to largest and the notation (i) in the rank statistic refers to the label attached to the i^{th} order statistic. Consider a set $R(t_{(i)})$ of items at risk at time $T_{(i)}^{-}$. The partial probability that item (i) fails at $T_{(i)}$ given that the items $R(t_{(i)})$ are at risk and that exactly one failure occurs at $T_{(i)}$ is:

$$\frac{h(T_{(i)}, \overline{z_{(i)}(t)})}{\sum_{l \in R(T_{(i)})} h(T_{(i)}, \overline{z_l(t)})} = \frac{\exp(\overline{\gamma} \cdot \overline{z_{(i)}(t)})}{\sum_{l \in R(T_{(i)})} \exp(\overline{\gamma} \cdot \overline{z_l(t)})}, \qquad (2.8.)$$

where $i = 1, 2, ..., k$. Equation (2.8.) shows that the baseline hazard rate has no effect on the joint probability and hence no effect on the estimated values of $\overline{\gamma}$. The partial likelihood can now be calculated by taking the product over all the failure points:

$$PL(\overline{\gamma}) = \prod_{i=1}^{k} \frac{\exp(\overline{\gamma} \cdot \overline{z_{(i)}(t)})}{\sum_{l \in R(T_{(i)})} \exp(\overline{\gamma} \cdot \overline{z_l(t)})} \qquad (2.9.)$$

To account for the possibility of ties at a specific $T_{(i)}$, i.e. the occurrence of more than one failure at a specific $T_{(i)}$, Breslow[65] has derived the following approximation of the partial likelihood (for numerical tractability):

$$PL(\overline{\gamma}) = \prod_{i=1}^{k} \frac{\exp(\overline{\gamma} \cdot \overline{s_i(t)})}{\left[\sum_{l \in R(T_{(i)})} \exp(\overline{\gamma} \cdot \overline{z_l(t)}) \right]^{d_i}}, \tag{2.10.}$$

where $s_i = \sum_j \overline{z_{ij}(t)}$ is the sum of the covariates of the d_i items observed to fail exactly at $T_{(i)}$.

The value of $\overline{\gamma}$ that maximizes $PL(\overline{\gamma})$ is the most appropriate for the semi-parametric PHM. At this optimal point, the partial derivatives of $PL(\overline{\gamma})$ with respect to all the m measured covariates should be zero, i.e.:

$$U_j(\overline{\gamma}) = \frac{\partial \log PL(\overline{\gamma})}{\partial \gamma_j} = \sum_{i=1}^{k} [s_{ji} - d_i A_{ji}(\overline{\gamma})] = 0 \quad (j = 1,2,\ldots,m), \tag{2.11.}$$

where s_{ji} is the j^{th} element in the vector s_i and

$$A_{ji}(\overline{\gamma}) = \frac{\sum_{l \in R(T_{(i)})} z_{jl} \exp(\overline{\gamma} \cdot \overline{z_l(t)})}{\sum_{l \in R(T_{(i)})} \exp(\overline{\gamma} \cdot \overline{z_l(t)})} \tag{2.12.}$$

A suitable optimization technique is required to perform the maximization such as Newton-Raphson iteration. This technique and others are described later in Chapter 3.

2.2.3 Efficiency of Partial Likelihood Estimation

After the introduction of partial likelihood, its efficiency was measured and compared by many researchers, for example Kalbsfleich and Prentice[3], and Effron[68].

Kalbsfleich and Prentice investigated the efficiency of the partial likelihood with the following two questions in mind: (1) Can the estimated $\overline{\gamma}$ be improved at all for the case where $h_0(t)$ is unspecified?; and (2) How does the partial likelihood estimate of $\overline{\gamma}$ compare to a maximum likelihood estimate? Two separate investigations were done, one for time-independent covariates and one for time-dependent covariates.

It was found for time-independent covariates that the partial likelihood estimation was 'reasonably' efficient. The details concerning the investigation are not discussed, see reference [3] pp. 103-113 for a comprehensive explanation. Cooper and Darch[69] agree with this statement after doing a study on armoured vehicles. For time-dependent covariates it is difficult to predict the partial likelihood's efficiency and it could in some cases even be very low. See reference [3], pp. 140-141.

2.2.4 Estimation of the Baseline Hazard Rate

The baseline hazard rate function represents the hazard rate that an item would experience if covariates had no effect on the item. No assumption needs to be made about its functional form for the semi-parametric PHM provided that the regression vector $\bar{\gamma}$ is known. Several authors have developed techniques with which the baseline hazard rate can be estimated, including Cox[2], Kalbsfleish and Prentice[3], Breslow[65,66] and Link[67]. The method of Breslow is presented here.

Suppose that $h_0(t)$ is a step function which jumps just before the occurrence of a failure and is constant between times to failure, i.e.:

$$h_0(t) = h_{0_i}, \quad T_{(i-1)} < t \le T_{(i)}, \quad i = 1,2,....,n \tag{2.13.}$$

With $h_0(t)$ defined as in (2.13.), an expression for the joint distribution can be derived and reduced to Cox's partial likelihood, exactly as explained in 2.2.2 which results in:

$$h_{0_i} = \frac{d_i}{(T_{(i)} - T_{(i-1)}) \sum_{l \in R(T_{(i)})} \exp(\bar{\gamma} \cdot \overline{z_l(t)})}, \tag{2.14.}$$

with h_{0_i} completely distribution-free. The distribution-free baseline hazard rate is often used to check the appropriateness of continuous distributions used for the baseline hazard rate.

2.3 The Fully Parametric PHM

By assuming a continuous distribution for the form of the baseline hazard rate, the PHM is completely parameterized. The very versatile Weibull distribution (and the only one considered in this research project) is most often used for the parameterization because of its flexibility. It is impossible to estimate the baseline hazard rate independently for the fully parametric PHM[1] and the distribution- and regression parameters have to be estimated simultaneously. Considerably less numerical problems arise for the Weibull PHM and its is much more numerically convenient.

2.3.1 Statistical Model

The time-dependent Weibull distribution is given by:

$$f(t)=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\cdot\exp(-(t/\eta)^{\beta}), \tag{2.15.}$$

with its corresponding hazard rate function:

$$h(t)=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}, \tag{2.16.}$$

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

If the Weibull distribution is used as the baseline hazard rate of the PHM as presented in (2.3.), the model becomes:

$$h(t,\overline{z(t)})=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\cdot\exp(\bar{\gamma}\cdot\overline{z(t)}), \tag{2.17.}$$

which is a fully parametric model.

From reliability theory we know that the reliability, R, of a component under the influence of ageing only, just before renewal at T_i is:

$$R(T_i)=\exp\left(-\int_0^{T_i}h(t)\mathrm{d}t\right)=\exp(-(T_i/\eta)^{\beta}) \tag{2.18.}$$

If $U_i=(T_i/\eta)^{\beta}$, then U_i has a unit negative exponential distribution. Similar to (2.18.), the reliability at a time T_i for a component under the influence of time-independent covariates according to the PHM can be estimated by:

[1] Also referred to as the Weibull Proportional Hazards Model in this dissertation.

$$R(t,\overline{z}) = \exp\left[-\int_0^{T_i} \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} dt \exp(\overline{\gamma}\cdot\overline{z})\right], \\ = \exp\left[-(T_i/\eta)^{\beta}\exp(\overline{\gamma}\cdot\overline{z_i})\right] \tag{2.19.}$$

with $U_i = (T_i/\eta)^{\beta}\exp(\overline{\gamma}\cdot\overline{z_i})$, again with unit exponential distribution. For the case of time-dependent covariates, the covariates have to be included in the integration to estimate the reliability of a component at time T_i:

$$R(t,\overline{z(t)}) = \exp\left[-\int_0^{T_i} \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp(\overline{\gamma}\cdot\overline{z_i(t)})dt\right], \\ = \exp\left[-\int_0^{T_i} \exp(\overline{\gamma}\cdot\overline{z_i(t)})d(t/\eta)^{\beta}\right] \tag{2.20.}$$

with $U_i = \int_0^{T_i} \exp(\overline{\gamma}\cdot\overline{z_i(t)})d(t/\eta)^{\beta}$, also with unit negative exponential distribution. In practice, (2.20.) is often approximated by:

$$R(t,\overline{z_i(t)}) = \exp\left\{\sum_{k=1}^{i} \exp(\overline{\gamma}\cdot\overline{z_i^*(t_k)})\cdot\left[\left(\frac{t_{k+1}}{\eta}\right)^{\beta} - \left(\frac{t_k}{\eta}\right)^{\beta}\right]\right\}, \tag{2.21.}$$

where $0 = t_0 < t_1 < .... < T_i$ are inspection points where covariate measurements were taken and $\overline{z_i^*(t_k)} = 0.5\cdot(\overline{z_i(t_k)} + \overline{z_i(t_{k+1})})$.

2.3.2 Parameter Estimation

The method of maximum likelihood is used to estimate the model's parameters. The full likelihood is obtained by:

$$L(\beta,\eta,\overline{\gamma}) = \prod_i h(T_i,\overline{z_i(T_i)})\cdot\prod_j R(T_j,\overline{z_j(t)}) \tag{2.22.}$$

where i indexes failure times and $j=1,2,...,n$ indexes failure and suspension times. If (2.17.) and (2.20) are substituted in (2.22.), the full likelihood becomes:

$$L(\beta,\eta,\overline{\gamma}) = \prod_i \frac{\beta}{\eta}\left(\frac{T_i}{\eta}\right)^{\beta-1} \exp\left(\overline{\gamma}\cdot\overline{z_i(T_i)}\right)\cdot \prod_j \exp\left[-\int_0^{T_j} \exp(\overline{\gamma}\cdot\overline{z_j(t)})\mathrm{d}(t/\eta)^{\beta}\right] \qquad (2.23.)$$

The same values for β, η and $\overline{\gamma}$ that maximize (2.23.) will also maximize $\log(L(\beta,\eta,\overline{\gamma}))$ or $l(\beta,\eta,\overline{\gamma})$, the log-likelihood. It is numerically much more attractive to maximize $l(\beta,\eta,\overline{\gamma})$ given by:

$$l(\beta,\eta,\overline{\gamma}) = r\ln(\beta/\eta) + \sum_i \ln[(T_i/\eta)^{\beta-1}] + \sum_i \overline{\gamma}\cdot\overline{z_i(T_i)} - \sum_j \int_0^{T_j} \exp(\overline{\gamma}\cdot\overline{z_j(t)})\mathrm{d}(t/\eta)^{\beta}, \qquad (2.24.)$$

with r being the number of failure renewals.

Several maximization techniques were tested on (2.24.) with success. This includes (a) a Nelder-Mead type simplex search method as is commonly found in the literature, (b) a BFGS Quasi-Newton method with a mixed quadratic and cubic line search procedure, (c) Snyman's dynamic trajectory optimization method[20,21] and (d) a modified Newton-Raphson procedure that gives fast convergence.

3 Covariates

The ability of the PHM to include covariates in its estimations and predictions is a very attractive attribute of the model. Covariate behavior and their effects on the PHM are not trivial issues however and a proper background knowledge of covariates is required before the model can be used with confidence. Some comments on these issues are presented in this section.

3.1 Effects of Interaction and Omission of Covariates

Interaction (dependency) of covariates can influence parameter estimation significantly and therefore the presence of this phenomenon should be checked. The easiest, most practical way to test for interaction of covariates is by introducing a new temporary covariate. (This is only to get a feel for the behavior of the data and is not very scientific). The temporary covariate is simply the product of the covariates under discussion and the new covariate's effect on the model is then tested. (Usually the interaction of only two covariates will be checked at a time). If interaction is present, the new covariate will be statistically significant in the model. The results of statistical tests for interaction, i.e. testing for significance of the new covariate in the model, should also be practically justified if possible, to identify possible inconsistencies in the data.

Bendell *et al.*[51] investigated covariate interaction by dividing data into groups (strata) based on major differences in the data and then estimating the regression coefficients for each group. It was shown that a suitable test statistic can be defined to test the effect of a particular covariate on different groups. See also reference [70] for a discussion on 'stratum-covariate' interactions.

Omission of influential covariates from the PHM also effects the regression coefficients of the model. Suppose z_1 and z_2 are two significant covariates with corresponding regression coefficients γ_1 and γ_2. If only z_1 is considered in the model and its coefficient is γ_1, then $|\gamma_1| < |\gamma_2|$. The estimates for γ_1 are asymptotically biased towards γ_2 and have smaller asymptotic variance than γ_2 [71]. The magnitude of the bias depends on the relative importance of the omitted covariate to that of the included covariate. Omission of influential covariates could also lead to overestimation or underestimation of the baseline hazard rate[72].

3.2 Effects of Measurement Error and Misspecification of Covariates

Significant errors in the estimation of the regression vector $\bar{\gamma}$ are possible if an inappropriate parametric form of the baseline hazard rate is specified or if errors in covariate measurement are distributed in an dismal manner. The regression vector is influenced in the same manner as in the of case of linear regression when measurement errors in the covariates are present. It is possible to test the significance of the effects of a covariate, in spite of its measurement error and misclassification, provided that sufficient information or assumptions are available relating to the covariate error or misclassification distribution[73]. Lagakos[74] determined that the efficiency of the partial likelihood estimator may be very low compared to the correct model if covariates are misspecified.

3.3 Effects of Monotonicity and Multicolinearity of Covariates

The 'usual' difficulties of regression analysis like multicolinearity, monotonicity and large covariate values are often encountered in the PHM as well. In such cases, maximization procedures used for parameter estimation often fail to converge.

Monotone increasing or decreasing covariate values in a data set, when ordered according to the magnitude of times to failure, is the biggest cause for divergence of maximization procedures. Bryson and Johnson[75] suggest seven steps to avoid the problems associated with monotonicity of covariates. In a censored data set, it may occur that the covariate at each failure time is either the largest or the smallest of all covariates in the risk set at that time. In such cases the regression vector estimate is often infinite.

It is important to formulate the covariates in such a manner that colinearity is avoided during the estimation of the regression vector since results may be very inaccurate. The data set could be analyzed in different groups based on a trail-and-error method to address this problem. Peduzzi *et al.*[76] published a general procedure for the selection of covariates in a nonlinear regression analysis to avoid colinearity. This is useful when there is a large number of covariates and it is difficult to determine the priority of selection of covariates for the model to their confusing effects on the times to renewal[59].

3.4 Time-dependent Covariates

Numerous papers have been published on the theory of time-dependent covariates in the PHM. This include topics like efficiency of estimation techniques for regression coefficients of time-dependent covariates[3], a two-step PHM model to accommodate time-dependent covariates more accurately[77], techniques to detect time-dependent effects of fixed covariates[78] and graphical techniques based on partial residuals suitable to detect time varying effects of covariates[79]. The detailed theory and analysis of time-dependent covariates are not important for this dissertation but rather its practical use and therefore this discussion is limited to some practical calculation issues.

Covariates in the Weibull PHM in (2.20.) are allowed to be time-dependent and assumed to be known for all values of time, t. This assumption is not entirely valid for the case of vibration covariates because vibration inspections are normally done on a discrete periodic basis. It is thus necessary to estimate covariate values between inspections for the model in (2.20.) to hold. Experience has shown that there is no fixed rule for this estimation (especially not for vibration covariates) and every situation should be considered separately. In (2.21.) an estimation technique was presented where the covariate values between any two inspections were taken to be the average of the covariate values between the two particular inspections. Jardine *et al.*[4] describe a similar method that was used with success on aircraft and marine engine failure data.

In some situations where covariates have a monotonic behavior, conventional interpolation techniques can be used with confidence such as linear, hyperbolic, parabolic, geometric or exponential interpolation, depending on the particular situation. It can be much more difficult to predict the values of other, non-monotonic, situations between inspections. In these cases the most sensible option is usually to consider the covariate behavior as a continuous right jumps process, where covariate values only increase or decrease at inspections and remain constant between inspections.

4 Numerical Model Fitting Procedures

Four optimization techniques were implemented successfully to fit the Weibull PHM with the method of maximum likelihood, 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.
ii. A Standard BFGS Quasi-Newton method with a mixed quadratic and cubic line search procedure.
iii. Snyman's dynamic trajectory optimization method[20,21]
iv. A modified Newton-Raphson procedure.

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 (1) and (2) maximized the likelihood function successfully but performed fairly mediocre. Snyman's method was found to be somewhat expensive but extremely robust which is a very valuable attribute. The modified Newton-Raphson method proved to be by far the most economical and fairly robust as well. Of the four above mentioned techniques, this technique is certainly the most suitable for optimization of the maximum likelihood function.

For the above mentioned reasons, only Snyman's method and the modified Newton-Raphson method are considered in this discussion on numerical model fitting procedures. Snyman's method is presented in fairly general terms to illustrate its robustness, but the Newton-Raphson method is described in detail since this method is used in the case study in Chapter 4.

4.1 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 general presentation of Snyman's technique, the objective function is $l(\overline{\theta})$, the maximum log-likelihood function as presented in (2.24.), where $\overline{\theta} = [\beta, \eta, \overline{\gamma}]$.

4.1.1 Characteristics

The attributes of Snyman's technique can be summarized as follows:

i. It uses only gradient information, i.e. $\overline{\nabla}[l(\overline{\theta})]$.
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 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.

4.1.2 Basic Dynamic Model

Assume a particle of unit mass in a m-dimensional conservative force field with potential energy at $\overline{\theta}$ given by $l(\overline{\theta})$, then the force experienced by the particle at $\overline{\theta}$ is given by:

$$m\overline{a} = \ddot{\overline{\theta}} = -\overline{\nabla}[l(\overline{\theta})], \qquad (4.1.)$$

from which it follows that for the time interval $[0, t]$:

$$\frac{1}{2}\left\|\dot{\bar{\theta}}(t)\right\|^2 - \frac{1}{2}\left\|\dot{\bar{\theta}}(0)\right\|^2 = l(\bar{\theta}(0)) - l(\bar{\theta}(t)) \qquad (4.2.)$$

Equation (4.2.) can be simplified by expressing it in terms of kinetic energy as:

$$T(t) - T(0) = l(0) - l(t) \qquad (4.3.)$$

From (4.3.) it evident that $l(t) + T(t) = \text{constant}$, 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.

4.1.3 Basic Algorithm

Suppose $l(\bar{\theta})$ has to be minimized from a starting point $\bar{\theta}(0) = \bar{\theta}_0$, then the dynamic algorithm is as follows:

i. Compute the dynamic trajectory by solving the initial value problem, $\ddot{\bar{\theta}}(t) = -\bar{\nabla}[l(\bar{\theta}(t))]$, $\dot{\bar{\theta}}(0) = 0$ and $\bar{\theta}(0) = \theta_0$. In practice the numerical integration of the initial value problem is often done by the 'leap-frog' method. Compute for $k = 0,1,2,....$ and time step Δt, the following: $\bar{\theta}^{k+1} = \bar{\theta}^k + \dot{\bar{\theta}}^k \Delta t$ and $\dot{\bar{\theta}}^{k+1} = \dot{\bar{\theta}}^k + \ddot{\bar{\theta}}^k \Delta t$, where $\ddot{\bar{\theta}}^k = -\bar{\nabla}[l(\bar{\theta}^k)]$ and $\overline{\dot{\theta}_0} = (1/2)\overline{\ddot{\theta}_0}\Delta t$.

ii. Monitor $\dot{\bar{\theta}}(t)$, the velocity of the particle. As long as the kinetic energy $T = \frac{1}{2}\left\|\dot{\bar{\theta}}(t)\right\|$ increases, the potential energy decreases, i.e. $l(\bar{\theta})$ decreases.

iii. As soon as T decreases, the particle is moving uphill and the objective function is increasing, i.e. $\|\dot{\bar{\theta}}^{k+1}\| \leq \|\dot{\bar{\theta}}^k\|$. 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 $\dot{\bar{\theta}}^k = (1/4)(\dot{\bar{\theta}}^{k+1} + \dot{\bar{\theta}}^k)$ and $\bar{\theta}^{k+1} = (1/2)(\bar{\theta}^{k+1} + \bar{\theta}^k)$ after which a new $\bar{\theta}^{k+1}$ is calculated and the algorithm is continued.

iv. To accelerate convergence of the method, the algorithm should allow for magnification and reduction of the stepsize, Δt, depending on the particle's position.

4.2 Modified Newton-Raphson Optimization Method

As mentioned earlier the modified Newton-Raphson optimization method was found to be the most suitable to perform the likelihood maximization according to the criteria defined above. This method is used in the case study in Chapter 4 and it will hence be discussed in detail in this chapter.

4.2.1 Data

To simplify the discussion of the optimization technique, the form in which the data should be arranged is defined here with specific reference to vibration monitoring.

Suppose we have n cases of renewal, called histories, in our data and i is used to indicate the history number, i.e. $i=1,2,...,n$. Let T_i denote time to failure or suspension in a particular history and use $c_1, c_2, ..., c_n$ as event indicators such that $c_i = 1$ if T_i is a failure time and $c_i = 0$ in case of suspension. The number of failures present in the data is thus $r = \sum c_i$.

To be able to develop the model for time-dependent covariates we set k_i to be the number of inspections or vibration measurements at moments t_{ij} during a certain history i over the period $(0, T_i]$ for $j = 1,2,....,k_i$ such that:

$$0 = t_{i0} < t_{i1} < t_{i2} < .... < t_{ik_i} = T_i \quad (4.4.)$$

Suppose that covariate vector $\overline{z_j^i} = (z_{j1}^i, z_{j2}^i, ..., z_{jm}^i)$ consisting of m covariates, is measured during history i. For convenience of estimation, it could be assumed that $\overline{z_l^i}(t_{\text{int}}) = \overline{z_l^i}(t_{ij})$, where $t_{ij} \le t_{\text{int}} < t_{i(j+1)}$. The data for history i can be summarized as follows:

Time	***Covariates***			
t_{i0}	z_{01}^i	z_{02}^i	....	z_{0m}^i
t_{i1}	z_{11}^i	z_{12}^i	....	z_{1m}^i
:	:	:	....	:
t_{ik_i}	$z_{k_i 1}^i$	$z_{k_i 2}^i$	....	$z_{k_i m}^i$

Table 4.1.: *Data summary for particular history*

Some covariates can be time-independent, i.e. vary with history i, but not with time t, or in mathematical terms $z_s^i = z_0^i$ for any valid value of s.

4.2.2 Definition of the Objective Function

The Weibull PHM as introduced in (2.17.) is repeated here for convenience as equation (4.5.):

$$h(t,\overline{z(t)}) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \cdot \exp(\overline{\gamma} \cdot \overline{z(t)}) \qquad (4.5.)$$

If the model in (4.5.) is transformed with an auxiliary equation $a = -\beta \ln \eta$, (4.5.) becomes:

$$h(t,\overline{z(t)}) = \beta \cdot t^{\beta-1} \cdot \exp(a + \overline{\gamma} \cdot \overline{z(t)}), \qquad (4.6.)$$

which is more convenient for calculation procedures. We will now construct the maximum log-likelihood function as explained in section (2.3.2) with (4.6.) as a function of $\overline{\theta}$, where $\overline{\theta} = (a, \beta, \gamma_1, \gamma_2, \ldots, \gamma_m)$. Also assume that $c_1 = c_2 = \ldots = c_r = 1$ and $c_{r+1} = \ldots = c_n = 0$ for simplification of notation. The log-likelihood will be of the form:

$$l(\overline{\theta}) = v(\overline{\theta}) - u(\overline{\theta}), \qquad (4.7.)$$

where

$$\begin{aligned} v(\overline{\theta}) &= \sum_{i=1}^{r} \ln h(t_i, \overline{z(t_i)}) \\ &= ra + r\ln\beta + (\beta - 1) \cdot A + \sum_{b=1}^{m} \gamma_b B_b \end{aligned}, \qquad (4.8.)$$

with

$$A = \sum_{i=1}^{r} \ln t_i \; ; \quad \text{and} \quad B_b = \sum_{i=1}^{r} z_{k_i b}^{i} \qquad (4.9.)$$

Similarly for the second part of equation (4.7.):

$$u(\overline{\theta}) = \sum_{i=1}^{n} \int_0^{t_i} h(s, \overline{z_i(s)})\mathrm{d}s$$
$$= \exp(a) \cdot \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \exp(\sum_{g=1}^{m} \gamma_g z^i_{jg}) \cdot (t^{\beta}_{i(j+1)} - t^{\beta}_{ij}) \quad (4.10.)$$

The most appropriate value of $\overline{\theta}$ is found where the objective function $l(\overline{\theta})$ is a maximum, i.e. where all the partial derivatives with respect to $\overline{\theta}$ are zero or $\partial l(\overline{\theta}) / \partial \overline{\theta}) = 0$.

4.2.3 Partial Derivatives

The first and second partial derivatives for $l(\overline{\theta}) = v(\overline{\theta}) - u(\overline{\theta})$ are required in the Newton-Raphson method.

First and second partial derivatives of $v(\overline{\theta})$ are:

$$\frac{\partial v(\overline{\theta})}{\partial a} = r, \quad \frac{\partial v(\overline{\theta})}{\partial \beta} = \frac{r}{\beta} + A, \quad \frac{\partial v(\overline{\theta})}{\partial \gamma_b} = B_b \quad (4.11.)$$

$$\frac{\partial^2 v(\overline{\theta})}{\partial a^2} = 0, \quad \frac{\partial^2 v(\overline{\theta})}{\partial \beta^2} = -\frac{r}{\beta^2}, \quad \frac{\partial^2 v(\overline{\theta})}{\partial {\gamma_b}^2} = 0, \quad \frac{\partial^2 v(\overline{\theta})}{\partial \theta_i \theta_j} = 0 \ (i \neq j) \quad (4.12.)$$

First and second partial derivatives of $u(\overline{\theta})$ with respect to a are:

$$\frac{\partial u(\overline{\theta})}{\partial a} = u(\overline{\theta}), \quad \frac{\partial^2 u(\overline{\theta})}{\partial a^2} = u(\overline{\theta}), \quad \frac{\partial^2 u(\overline{\theta})}{\partial a \partial \beta} = \frac{\partial u(\overline{\theta})}{\partial \beta}, \quad \frac{\partial^2 u(\overline{\theta})}{\partial a \partial \gamma_b} = \frac{\partial u(\overline{\theta})}{\partial \gamma_b} \quad (4.13.)$$

First and second partial derivatives of $u(\overline{\theta})$ with respect to β are:

$$\frac{\partial u(\overline{\theta})}{\partial \beta} = \exp(a) \cdot \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \exp(\sum_{g=1}^{m} \gamma_g z^i_{jg}) \cdot (t^{\beta}_{i(j+1)} \ln t_{i(j+1)} - t^{\beta}_{ij} \ln t_{ij})$$
$$\frac{\partial^2 u(\overline{\theta})}{\partial \beta^2} = \exp(a) \cdot \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \exp(\sum_{g=1}^{m} \gamma_g z^i_{jg}) \cdot (t^{\beta}_{i(j+1)} \ln^2 t_{i(j+1)} - t^{\beta}_{ij} \ln^2 t_{ij})$$
$$\frac{\partial^2 u(\overline{\theta})}{\partial \beta \partial \gamma_b} = \exp(a) \cdot \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \exp(\sum_{g=1}^{m} \gamma_g z^i_{jg}) \cdot z^i_{jb} (t^{\beta}_{i(j+1)} \ln t_{i(j+1)} - t^{\beta}_{ij} \ln t_{ij}) \quad (4.14.)$$

First and second partial derivatives of $u(\overline{\theta})$ with respect to γ_b are:

$$\frac{\partial u(\overline{\theta})}{\partial \gamma_b} = \exp(a)\cdot\sum_{i=1}^{n}\sum_{j=0}^{k_i-1}\exp(\sum_{g=1}^{m}\gamma_g z_{jg}^i)\cdot z_{jb}^i(t_{i(j+1)}^{\beta} - t_{ij}^{\beta})$$

$$\frac{\partial^2 u(\overline{\theta})}{\partial \gamma_b^{\ 2}} = \exp(a)\cdot\sum_{i=1}^{n}\sum_{j=0}^{k_i-1}\exp(\sum_{g=1}^{m}\gamma_g z_{jg}^i)\cdot(z_{jb}^i)^2\cdot(t_{i(j+1)}^{\beta} - t_{ij}^{\beta})$$

$$\frac{\partial^2 u(\overline{\theta})}{\partial \gamma_b \gamma_s} = \exp(a)\cdot\sum_{i=1}^{n}\sum_{j=0}^{k_i-1}\exp(\sum_{g=1}^{m}\gamma_g z_{jg}^i)\cdot(z_{jb}^i)\cdot(z_{js}^i)\cdot(t_{i(j+1)}^{\beta} - t_{ij}^{\beta}) \qquad (4.15.)$$

4.2.4 Numerical Procedure

The objective of the numerical procedure is to find the value of $\overline{\theta}$ where all the partial derivatives are zero. Let $F(\overline{\theta}) = \partial l(\overline{\theta})/\partial\overline{\theta} = (\partial l/\partial a, \partial l/\partial\beta, \partial l/\partial\gamma_1, ..$ $...., \partial l/\partial\gamma_m)$ and $G(\overline{\theta}) = \partial^2 l/\partial\overline{\theta}^2 = \partial^2 l/\partial\theta_i\partial\theta_j$. (Matrix notation is suppressed for convenience). The following approximation for $F(\overline{\theta})$ can be used: $F(\overline{\theta}) \approx F(\overline{\theta_0}) + G(\overline{\theta_0})\cdot(\overline{\theta} - \overline{\theta_0})$ where $\overline{\theta_0}$ is an initial estimate. It is required to solve $F(\overline{\theta_0}) + G(\overline{\theta_0})\cdot(\overline{\theta} - \overline{\theta_0}) = 0$ to determine the optimal value of $\overline{\theta}$.

The conventional Newton-Raphson procedure would solve for $\overline{\theta}$ as follows:

i. Estimate a meaningful initial value for $\overline{\theta}$, i.e. $\overline{\theta_0}$.
ii. Calculate $F(\overline{\theta_0})$ and $G(\overline{\theta_0})$.
iii. Solve for Δ_0 in the system $G(\overline{\theta_0})\overline{\Delta_0} = -F(\overline{\theta_0})$.
iv. Set $\overline{\theta_1} = \overline{\theta_0} + \overline{\Delta_0}$ 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, $G(\overline{\theta})$ is not calculated directly but an approximation of $G(\overline{\theta})$ is used which is chosen to be always positive definite, thereby eliminating the possibility of singular matrices. The approximation of $G(\overline{\theta})$ is explained in detail in reference [80].

To accelerate convergence and increase accuracy of the procedure, the data is transformed to more numerically convenient forms before the iteration process is started. All recorded times, including inspection times, times to failure and times to suspension as well as the scale parameter η are divided by a value C, where:

$$C = (1/N)\sum_{i,j} t_{ij}\ , \qquad (4.16.)$$

and N is the total number of recorded times. With these scaled observations, initial values of $\eta_0 = 3 \cdot C$ and $\beta_0 = 1.5$ were found to be very reliable in the numerical procedure. The covariates are also standardized to have the same relative magnitude with the following relation:

$$z_{jl}^{*i} = \frac{z_{jl}^{i} - z_{b}^{avg}}{s_b}, \qquad (4.17.)$$

where

$$z_b^{avg} = \frac{1}{N_b} \sum_{i,j} z_{jb}^{i} \quad \text{and} \quad s_b^2 = \frac{1}{N_b} \sum_{i,j} (z_{jb}^{i} - z_b^{avg})^2, \qquad (4.18.)$$

with N_b the number of recordings of a specific covariate. This standardization of the covariates accelerates convergency considerably with initial values $\overline{\gamma_0} = \overline{0}$.

Methods to vary step sizes of the procedure as well as stopping rule procedures are discussed in Press *et al.*[80].

5 Goodness-of-fit Tests

Goodness-of-fit tests for the PHM are all aimed at evaluating the assumptions (see section (2.3.)) on which the model is based. Methods to evaluate the first assumption, the i.i.d. assumption, were discussed in section (2.1.) of Chapter 1 in considerable detail and are not repeated here.

For the second assumption, graphical methods are usually employed to test whether an influential covariate has been omitted from the model. Plots of estimated cumulative hazard rates versus the number of renewals for different strata should be approximately linear with slope equal to one, if no influential covariate has been omitted[81].

Two approaches can be used to test the validity of the third assumption. Graphical techniques have been used most widely and is considered to be the first approach.[81,82,83] The other approach is either based on hierarchical models or makes use of analytical techniques. In the case of hierarchical models, a time-dependent covariate is introduced into the model and tests are then performed to establish whether the estimate of the effect of this covariate is significantly different from zero.[2,84,85,86] A review of these methods is given by Kay[87].

5.1 Graphical Methods

Graphical methods suitable for testing the assumptions of the PHM can generally be categorized into three groups: cumulative hazard plots, average hazards plots and residual plots. In this discussion the emphasis is on residual plots because of its versatility and enormous level of inherent information.

5.1.1 Cumulative Hazard Plots

Measured values of a certain covariate may often be grouped into different levels, also referred to as *strata*, *r*. For example, the covariate $z_r(t)$ which occurs on s different levels and for which the proportionality assumption is to be tested is assigned to one of the s strata. Therefore, the hazard rate in this case can be written as:

$$h_r(t,\overline{z_r(t)}) = h_{0r}(t)\cdot\exp(\sum_{j=1,j\neq r}^{m}\gamma_j\cdot z_j(t)) \qquad (5.1.)$$

To explain cumulative hazard plots further, consider a binary covariate $z_r(t)$ of which the level indicator s has two values, *0* and *1*. This yields the following in terms of total hazard rate:

$$h(t,\overline{z(t)}) = h_0(t)\cdot\exp(\sum_{j=1,j\neq r}^{m}\gamma_j\cdot z_j(t))\cdot\exp(\gamma_r); \quad (s=1) \qquad (5.2.)$$

$$h(t,\overline{z(t)}) = h_0(t)\cdot\exp(\sum_{j=1,j\neq r}^{m}\gamma_j\cdot z_j(t)); \quad (s=0) \qquad (5.3.)$$

Therefore, to satisfy the assumption that $h(t,z_x(t))\,\alpha\,h(t,z_y(t))$ we obtain from (5.2.) and (5.3.) that $h_{0r}(t)=c_r h_0(t)$ for $r=1,2,\ldots,s$ where $c_1,c_2,\ldots,c_s$ are constants equal to $\exp(\gamma_r\cdot z_r(t))$ for all strata. A similar relation holds for the cumulative hazard rate, i.e. $H_{0r}(t)=c_r H_0(t)$. If plots of the logarithm of the estimated cumulative hazard rates against time are constructed, they will be shifted by an additive constant γ_r, the regression parameter of a specific stratum. Thus, if the proportionality assumption is valid, the two plots should be approximately parallel and separated according to the different values of the covariates. Figure 5.1. below illustrates this concept for a case where the PHM is valid.

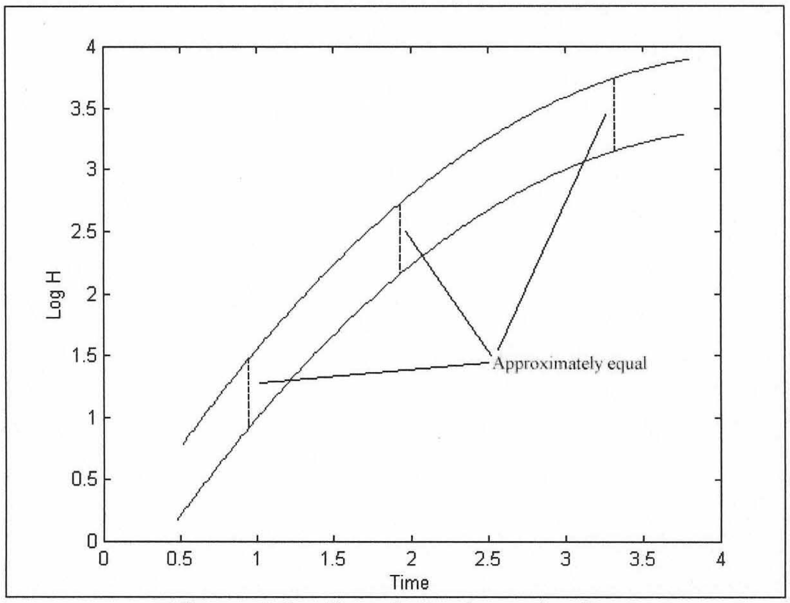

Figure 5.1.: *Cumulative hazards plots*

5.1.2 Average Hazards Plots

Average hazards plots for different strata are based on the assumption that $h_{rj}(t)$ is a slowly varying function that can be approximated by piecewise constant functions, i.e. $h_{rj}(t) = h_{rj}$ in the time interval between inspection number *i(j-1)* and *ij*, where *i* is the history number. Cox[82] defined $S_{rj}^{(i)}$ to be the total estimated operational time at risk in the strata *r* between inspections *i(j-1)*th and *j*th. The operational time is the timescale obtained when the contribution at each inspection point associated with a specific covariate is weighted by $\exp(\gamma \cdot \overline{z_i(t)})$. An auxiliary random variable is now defined as:

$$Z_{rj}^{(p)} = \frac{1}{(2i - 1/3)} - \ln\left(\frac{S_{rj}^{(i)}}{i}\right), \qquad (5.4.)$$

which is independent with mean $\ln(h_{rj})$ and variance $1/(i-0.5)$ [82]. If plots of $Z_{rj}^{(i)}$ against the midpoint of the time interval are constructed, the different plots should be parallel and spaced according to the estimated value of the covariate $\overline{z_r(t)}$ defining the strata, if the model fits the data.

5.1.3 Residual Plots

Residual plots are constructed with Cox-generalized residuals for the PHM. Cox-generalized residuals are given by:

$$r_i = \begin{cases} u_i, \text{if } t_i \text{ is a failure time} \\ u_i + 1, \text{if } t_i \text{ is a suspension time} \end{cases} \quad (5.5.)$$

where $i = 1,2,....,n$.

In (5.5.) u_i is defined as:

$$u_i = \frac{1}{\eta^\beta}\sum_{j=0}^{k_i-1} \exp(\bar{\gamma}\cdot\bar{z}_j^i)\cdot\left[t_{i(j+1)}^\beta - t_{ij}^\beta\right], \quad (5.6.)$$

with the same notation used as in section (4.2.). The calculation of u_i can be checked by noting that:

$$\sum_{i=1}^{n} u_i = r; \text{ and } \sum_{i=1}^{n} r_i = n, \quad (5.7.)$$

with r denoting the number of failures in this case. The unknowns in (5.6.) are determined during the model fitting procedure as described in paragraph (4.2.). With these Cox-generalized residuals known, several plots can be constructed to assess the goodness of fit visually.

5.1.3.1 Residuals Against Order of Appearance

Here the residual for every history $i = 1,2,....,n$ is plotted against the corresponding history number, i.e. $(x_i; y_i) = (i; r_i)$.

The residuals should all be scattered around the straight line $y = 1$. Note that the residual values of suspended cases will always be greater than *1*. If an upper limit of $r_i = 3$ (95%) and a lower limit of $r_i = 0.05$ (5%) are chosen, it is expected that at least 90% of the residuals will fall inside these limits if the model fits the data.

5.1.3.2 Ordered Residuals Against Expectation

If the calculated residuals $r_1, r_1, ..., r_n$ are ordered in ascending order we get $r_1^* \leq r_2^* \leq ..... \leq r_n^*$, the ordered residuals. The expected values of the residuals are:

$$E_i = E_{i,n} = \frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{n-i+1} \quad (5.8.)$$

For a Weibull PHM that fits the data well, the points $(x_i ; y_i) = (E_i ; r_i^*)$ will be distributed around the line $y = x$. Note however that the difference in consecutive expectations increases and that there will be a concentration of cases below the value *1* (around 50% - 60%). The points on the right side of the $y = x$ line need not necessarily be close to the line to indicate an appropriate model. This is because the variability of the residuals increases with order number. It is possible to improve the situation by using suitable transformations for the residuals.

To transform all points to lie between *0* and *1* with an approximately equally spaced *x*-axis we could use $x_i = 1 - \exp(-E_i)$ and $y_i = 1 - \exp(-r_i^*)$. It is possible to stabilize the variance by using $x_i = (2/\pi)\sin^{-1}\{\exp[-(E_i/2)]\}$ for the *x*-axis and $y_i = (2/\pi)\sin^{-1}\{\exp[-(r_i^*/2)]\}$ for the *y*-axis. All points will again lie between *0* and *1*. For further discussion on this method see reference [88, 89].

5.2 Analytical Goodness-of-Fit Tests

Graphical goodness-of-fit tests are often interpreted totally different by different analysts. For this reason, analytical goodness-of-fit tests have tremendous value since it is totally objective. Several analytical tests have been used on the PHM in the literature, amongst others, the χ^2 test, the log rank test, Z-test (normal distribution test), Kolmogorov-Smirnov test, Wald test, the doubly cumulative hazard function, the likelihood ratio test, score tests and generalized moments specification tests. See references [3,84,90,91,92,93,94,95,96,98]. Three of these tests are discussed below.

5.2.1 *Z*-Test

Before the *Z*-test can be presented, some comments have to be made about transformation rules in statistics. Transformation rules describe the changes in the mean, variance and standard deviation of a distribution when every item in a distribution is either increased or decreased by a constant amount. These rules also describe the changes in the mean, variance and standard deviation of a distribution when every item in the distribution is either multiplied or divided by a constant amount.

- **Rule 1:** Adding a constant to every item in a distribution adds the constant to the mean of the distribution, but it leaves the variance and standard deviation unchanged.

- **Rule (2):** Multiplying every item in a distribution by a constant multiplies the mean and standard deviation of that distribution by the constant and it multiplies the variance of the distribution by the square of the constant.

The *Z*-test makes use of a special application of the transformation rules, the *Z*-score statistic from which inferences are made. The *Z*-score for an event, indicates how far and in what direction that event deviates from its distribution's mean, expressed in units of its distribution's standard deviation. The mathematics of the *Z*-score transformation are such that if every event in a distribution is converted to its *Z*-score, the transformed scores will necessarily have a mean of zero and a standard deviation of one.

For the PHM, the *Z*-test can be used by letting $r_1^* \le r_2^* \le ..... \le r_n^*$ be the ordered residuals, as before, and define $Z_i = r_i^* / r_n^*$ and $m = n - 1$. The *Z*-score for the PHM is then:

$$Z = \frac{\sum_{i=1}^{m} Z_i - m/2}{\sqrt{m/2}} \qquad (5.9.)$$

Inferences about the value of *Z* can be made by calculating the *p*-value of *Z*, using the normal distribution.

5.2.2 Kolmogorov-Smirnov Test

This testing procedure is classified as a frequency test of the degree of agreement between distributions of a sample of generated random values (or a sample of empirically gathered values) and a target distribution. Since it is known that the residuals of the PHM should have an exponential distribution, the Kolmogorov-Smirnov test is performed on the PHM residuals to check the model fit.

The null hypothesis is that the cumulative density function of the PHM residuals is equal to the cumulative density function of an exponential distribution fitted on the residuals. To be able to test the null hypothesis, a test statistic, called the *D-statistic*, is introduced. This statistic is defined as the largest absolute difference between the Weibull PHM residuals and the cumulative exponential distribution and inferences on the goodness-of-fit of a model is made based on this statistic. The procedure seems simple but becomes fairly complicated for the PHM with censored data[97].

As before, assume $r_1^* \leq r_2^* \leq ..... \leq r_n^*$ to be the residuals ordered by magnitude and c_i to be an event indicator as defined in section (4.2.1). Let s_i and a_i be sequences defined by:

$$s_{i+1} = s_i \cdot \left(1 - \frac{1}{n-i}\right)^{c_i} \tag{5.10.}$$

$$a_{i+1} = a_i + \frac{n}{(n-i)\cdot(n-i-1)} \cdot c_i , \tag{5.11.}$$

with $s_0 = 1$, $s_n = \exp(-r_{n-1}^*)$, $a_0 = 0$, $a_n = 0$ and $i = 0,1,2,....,n-2$. The D-statistic is then:

$$D = \max_{0 \leq i \leq n-1} \left\{ \max \left\{ \frac{\left|s_i - \exp(-r_i^*)\right|}{(1+a_i)\exp(-r_i^*)}, \frac{\left|s_{i+1} - \exp(-r_i^*)\right|}{(1+a_{i+1})\exp(-r_i^*)} \right\} \right\} \tag{5.12.}$$

From the D-statistic, a p-value can be determined to evaluate the quality of the fit.

5.2.3 Wald Test

A test specifically developed for testing the quality of parameter estimations by the method of maximum likelihood, is the Wald test. This test is categorized under likelihood ratio tests and can be used to evaluate the appropriateness of specific coefficients in the estimated regression vector and not only the total goodness-of-fit of a model. This attribute of the Wald test is very useful for the PHM because the contribution of different covariates to the quality of the model can be assessed.

The Wald test statistic for a specific regression coefficient is then given by:

$$W_i = \frac{n \cdot (\theta_i)^2}{\text{Var}(\theta_i)}, \tag{5.13.}$$

where $\text{Var}(\theta_i)$ is the variance of the regression coefficient and n is the sample size. Inference on the Wald test statistic is made by calculating p-values from the χ^2 distribution. See [98] for further details.

6 Optimal Decision Making with the Proportional Hazards Model

The PHM supplies us with an accurate estimate of a component's present risk to fail (hazard rate), based on its primary use parameter and the influence of covariates. This educated knowledge of the hazard rate should be utilized to the full to obtain economical benefits, otherwise the PHM estimation exercise is futile.

Economical benefits from a statistical failure analysis can be guaranteed with a high confidence level if the minimum long term life cycle cost (LCC) of a component is determined and pursued, i.e. if renewal always takes place at either the statistical minimum LCC or in the case of failure prior to the minimum LCC. The instant of minimal LCC can easily be specified in terms of time for statistical models where time is the only age parameter but for models including covariates, this makes no sense because the covariates influence survival time.

For optimal decision making with the PHM in reliability, only one model is used in practice, a model specifically developed for the PHM by Makis and Jardine[13,14]. This specifies the optimal renewal policy in terms of an optimal hazard rate which will lead to the minimum LCC. At every inspection the latest hazard rate is calculated and if it exceeds the optimal hazard rate the component is renewed, otherwise operation is continued. If the recommendations of the decision model is obeyed, the LCC will strive to a minimum over the long run.

To be able to determine the hazard rate that will lead to the minimum LCC it is required to predict the behavior of covariates. Makis and Jardine's model does this by assuming the covariate behavior to be stochastic and approximating it by a non-homogeneous Markov chain in a finite state space. The Markov chain leads to transition probabilities of covariates from where the optimal hazard rate is calculated.

An approach of predicting the useful remaining service life of a component and acting preventively on the prediction rather than pursuing a statistical optimum sounds intuitively meritorious but no research on techniques for such an approach in conjunction with the PHM has been published up to date.

6.1 The Long Term Life Cycle Cost Concept

LCC is a concept used widely in statistical failure analysis. Several models to achieve this minimum for repairable systems and renewal situations which depends only on time can be found in the literature. See [15] and [22] for an overview. The minimum LCC in renewal situations arise from two important quantities in practice namely the cost of unexpected renewal or failure of a component, C_f, and the cost of preventive replacement C_p. It is normally much more expensive to deal with an unexpected failure than it is to renew preventively. A balance has to be obtained between the risk of having to spend C_f and the advantage in the cost difference between C_f and C_p without wasting useful remaining life of a component. The optimum economic preventive renewal time will be at this balance point. LCC's are usually compared when expressed as cost per unit time.

The LCC concept will be illustrated with a simple Weibull model, i.e. a model without covariates. If a component is renewed either preventively after t_p time units or at unexpected failure for every life cycle over the long term, the total expected cost for a life cycle would be:

$$C_t = C_p R(t_p) + C_f [1 - R(t_p)] \qquad (6.1.)$$

Since, the LCC is usually expressed in cost per unit time the average life expectancy has to be calculated as well:

$$L_e = (t_p + T_p) R(t_p) + (t_f + T_f)[1 - R(t_p)], \qquad (6.2.)$$

where t_f is the expected length of a failure cycle under the condition that failure occurred before t_p and T_p and T_f are the times required for preventive renewal and failure renewal, respectively. When (6.1.) and (6.2.) are expressed in terms of the Weibull reliability functions and divided into each other, the LCC per unit time, if renewed at t_p over the long term, is:

$$C(t_p) = \frac{C_p \cdot e^{-(t_p/\eta)^\beta} + C_f \cdot [1 - e^{-(t_p/\eta)^\beta}]}{(t_p + T_p) \cdot e^{-(t_p/\eta)^\beta} + \int_0^{t_p} \left(t \cdot \frac{\beta}{\eta} \cdot \left(\frac{t}{\eta}\right)^{\beta-1} \cdot e^{-(t/\eta)^\beta} \right) dt + T_f \cdot [1 - e^{-(t_p/\eta)^\beta}]} \qquad (6.3.)$$

The preventive renewal time that will lead to the minimum LCC, t_p^*, is found where $D_t[C(t_p)] = 0$. An example of (6.3.) is shown in Figure 6.1. for Weibull parameters of $\beta = 1.80$ and $\eta = 430$ days.

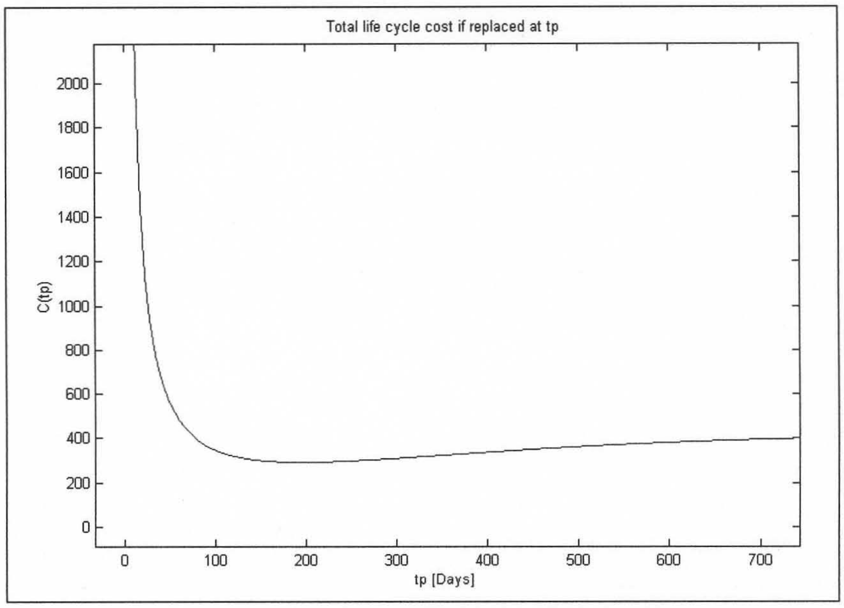

Figure 6.1.: *Example plot of LCC function*

A distinct minimum for $C(t_p)$ in Figure 6.1. exist at 221 days.

6.2 Prediction of Covariate Behavior

Techniques used to predict covariate behavior in Makis and Jardine's model is discussed in this section.

6.2.1 Transition States and Covariate Bands

Transition states have to be defined for the covariates before it can be modeled with a Markov chain. 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 (TPM). For numerical convenience, 4 or 5 bands are usually selected between upper and lower bands except for the last band which need not have an upper bound.

6.2.2 Markovian Chains and the Transition Probability Matrix

Suppose that $\{X_0, X_1, X_2, ...\}$ is a multidimensional Markov process which makes up a component's renewal history such that $X_k = (z_{k1}(t), z_{k2}(t), ..., z_{km}(t)) \in \Re^m$ where m is the number of covariates, and $z_{ki}(t)$ is the k^{th}

observation of variable i before renewal, performed at time $t = k\Delta$, ($k = 0,1,2,....$) while Δ is a fixed inspection interval. A stochastic process $\{X_0, X_1, X_2,....\}$ is assumed to be Markovian if, for every $k \geq 0$,

$$P\{X_{k+1} = j \mid X_k = i, X_{k-1} = i_{k-1}, X_{k-2} = i_{k-2},......, X_0 = i_0\} = P\{X_{k+1} = j \mid X_k = i\} \quad (6.4.)$$

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:

$$P_{ij}(k) = P_{ij}(k,\Delta) = P(X_{k+1} = j \mid T > (k+1)\Delta, X_k = i), \quad (6.5.)$$

where T denotes time to renewal as before and i and j denote any two possible states.

Suppose we have a sample $X_{i0}, X_{i1}, X_{i2},....$ 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 many histories:

$$n_{ij}(k) = \#\{X_k = i, X_{k+1} = j\} \quad (6.6.)$$

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

$$n_i(k) = \#\{X_k = i\} = \sum_j n_{ij}(k) \quad (6.7.)$$

It is now 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:

$$\hat{P}_{ij}(k) = \frac{n_{ij}(k)}{n_i(k)}, \quad k = 0,1,2,... \quad (6.8.)$$

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

$$\hat{P}_{ij}(k) = \frac{\sum_{a \le k \le b} n_{ij}(k)}{\sum_{a \le k \le b} n_i(k)}, \quad a \le k \le b \tag{6.9.}$$

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:

$$\hat{P}_{ij} = \frac{n_{ij}}{n_i}, \text{ where } n_{ij} = \sum_{k \ge 0} n_{ij}(k), \; n_i = \sum_j n_{ij} \tag{6.10.}$$

It is not realistic to assume that the transition probabilities of vibration covariates are independent of time. For this reason continuous time is divided into w intervals, $[0, a_1], (a_1, a_2], ..., (a_w, \infty)$, in which the transition probabilities are considered to be homogeneous. This manipulation simplifies the calculation of the TPM tremendously without loosing much accuracy.

The estimations of the TPM above all 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 than Δ have to be omitted from TPM 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 $i|_{t=0} \rightarrow j|_{t=t}$ is $P_{ij}(t) = P(X(t) = j \mid X(0) = i)$ and the rate at which the transition will take place is $D_t[P_{ij}(t)] = \lambda_{ij}$, $(i \ne j)$. For the case where $i = j$ the transition rate can be derived with the following argument. Suppose the system is in state $i|_{t=0}$ and state $j|_{t=t}$ with r possible states. If the sum over all probabilities over t is taken:

$$\begin{aligned} P_{i0}(t) + P_{i1}(t) + P_{i2}(t) + ... + P_{ir}(t) = 1 \\ \sum_j P(X(t) = j \mid X(0) = i) = 1 \\ \text{or } \sum_j P_{ij}(t) = 1 \end{aligned} \tag{6.11.}$$

If we take the time derivative,

$$\sum_j \frac{\partial}{\partial t}\left[P_{ij}(t)\right]=0$$

$$\therefore \lambda_{i0}+\lambda_{i1}+\lambda_{ii}+.....+\lambda_{ir}=0 \qquad (6.12.)$$

$$\lambda_{ii}=-\sum_{i\neq j}\lambda_{ij}$$

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

$$\hat{\lambda}_{ij}=\frac{n_{ij}}{\Omega_i}, \qquad n_{ij}=\sum_k n_{ij}(k) \qquad (6.13.)$$

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 t with:

$$P'_{ij}(t)=\sum_l P_{il}(t)\lambda_{lj} \qquad (6.14.)$$

Equation (6.14.) 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:

$$P(t)=\exp(A\cdot t), \qquad (6.15.)$$

where $P(t)=(P_{ij}(t))$ and $A=(\lambda_{ij})$. (Brackets denote matrices). This can be calculated by the series:

$$P(t)=\sum_{n=0}^{\infty}A^n\frac{t^n}{n!}, \qquad (6.16.)$$

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.

6.2.3 Calculation of the Optimal Decision Policy

Two different renewal possibilities are considered in Makis and Jardine's model: (i) Variant 1, where preventive renewal can take place at any moment; (ii) Variant 2, where preventive replacement can only take place at moments of inspection. Only Variant 1 will be discussed since Variant 2 is only a simplification of Variant 1.

A basic renewal rule is used: if the hazard rate 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 working age and covariates into account.

The expected average cost per unit time is a function of the threshold risk level, d, and is given by [13,14]:

$$\Phi(d) = \frac{C_p + KQ(d)}{W(d)}, \tag{6.17.}$$

where $K = C_f - C_p$. $Q(d)$ represents the probability that failure replacement will occur, i.e. $Q(d) = P(T_d \geq T)$ with T_d the preventive renewal time at threshold risk level d or $T_d = \inf\{t \geq 0 : h(t, \overline{Z(t)}) \geq d / K\}$. $W(d)$ is the expected time until replacement, regardless of preventive action or failure, i.e. $W(d) = E(\min\{T_d, T\})$. The optimal threshold risk level, d^*, is determined with fixed point iteration to get:

$$\Phi(d^*) = \min_{d>0} \Phi(d) = d^*, \tag{6.18.}$$

if the hazard function is non-decreasing, e.g. if $\beta \geq 1$ and all covariates are non-decreasing and covariate parameters are positive. If covariates are not monotonic, then the fix point iteration does not work, and $\min_{d>0} \Phi(d)$ should be found by a direct search method. During the calculation of d^* it is necessary to calculate $Q(d)$ and $W(d)$ which is no a trivial procedure. To do this we define the covariate vector $\overline{z(t)} = [z_1(t), z_2(t), ..., z_m(t)]$ as before with $\overline{i(t)} = [i_1(t), i_2(t), ..., i_m(t)]$ the state of every covariate at time t. Thus, for every coordinate l let $X^l(i_l(t))$ be the value of the l^{th} covariate in state $i_l(t)$ (representative of the state) at moment t, and $X(\overline{i(t)}) = \{X^1(i_1(t)), ..., X^m(i_m(t))\}$. We could express the hazard rate now as:

$$h(t, \overline{i(t)}) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp(\overline{\gamma} \cdot X(\overline{i(t)})) \tag{6.19.}$$

From (2.20.), the conditional reliability function can be defined as $R(j, i, t) = P(T > j\Delta + t \mid T > j\Delta, \overline{i(t)})$, which becomes after substitution:

$$R(j, \overline{i(t)}, t) = \exp\left\{ -\exp(\overline{\gamma} \cdot \overline{X}) \cdot \left[\left(\frac{j\Delta + t}{\eta}\right)^{\beta} - \left(\frac{j\Delta}{\eta}\right)^{\beta} \right] \right\} \tag{6.20.}$$

with $0 \le t \le \Delta$. If $h(t,\overline{i(t)})$ is a non-decreasing function in *t*, and if we define $t_i = \inf\{t \ge 0 : h(t,\overline{i(t)}) \ge d/K\}$ and the k_i's as integers such that $(k_i - 1)\Delta \le t_i < k_i\Delta$ we can calculate the mean sojourn time of the system in each state with:

$$\tau(j,\overline{i(t)}) = \begin{cases} 0, \; j \ge k_i \\ \tau(j,i,a_i), \; j = k_i - 1 \\ \tau(j,i,\Delta), \; j < k_i - 1 \end{cases}, \tag{6.21.}$$

where $a_i = t_i - (k_i - 1)\Delta$ and $\tau(j,i,s) = \int_0^s R(j,i,t)dt$. Similarly, the conditional cumulative distribution function for this situation is:

$$F(j,\overline{i(t)}) = \begin{cases} 0, \; j \ge k_i \\ 1 - R(j,i,a_i), \; j = k_i - 1 \\ 1 - R(j,i,\Delta), \; j < k_i - 1 \end{cases} \tag{6.22.}$$

Let for each *j*, $\tau_j = (\tau(j,i))_i$ and $F_j = (F(j,i))_i$ are column vectors, and $(P_j) = (R(j,i,\Delta)P_{il}(j))_{il}$ is a matrix. From here, the column vectors $W_j = (W(j,i))$ and $Q_j = (Q(j,i))$ are calculated as follows:

$$\begin{aligned} W_j &= \tau_j + P_j W_{j+1} \\ Q_j &= F_j + P_j Q_{j+1} \end{aligned} \tag{6.23.}$$

Then $W = W(0,i_0)$ 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*, where $W_{j+1} = Q_{j+1} = 0$ and working back to 0, it is possible to solve for *W* and *Q* from (6.23.). The above calculation procedure is described in detail in [14]. A forward version of this backward calculation is numerically more convenient and much faster (see [23]), which can be suitably adjusted for non-monotonic hazard functions also.

Thus, once the optimal threshold level is determined we renew the item at the first moment *t* when:

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\left(\overline{\gamma} \cdot \overline{z(t)}\right) \ge \frac{d^*}{K}, \tag{6.24.}$$

or, which is practically more convenient, when

$$\overline{\gamma} \cdot \overline{z(t)} \geq \delta^* - (\beta - 1)\ln t, \tag{6.25.}$$

where $\delta^* = \ln(\frac{d^* \eta^\beta}{K\beta})$.

A warning level function is defined only in terms of time by:

$$g(t) = \delta^* - (\beta - 1) \cdot \ln(t), \tag{6.26.}$$

with $g(t)$ strictly decreasing if $\beta > 1$.

Chapter 4

Vibration Covariate PHM Application

1 Introduction

The only way to truly contribute to the reliability modelling field with this dissertation is to apply the theory discussed in Chapter 3 successfully to an applicable situation in the industry. In Chapter 4, data collected from the industry is analyzed and modeled with the Proportional Hazards Model to make such a contribution.

Up to date, no successful case study on vibration covariates in the PHM has been published, mainly due to a lack of suitable data. While searching for suitable data in South Africa, a number of serious shortcomings in vibration data recording practices were discovered in general vibration monitoring programs. From the shortcomings it was possible to compile a structured list of data requirements that have to be fulfilled before vibration covariates can be used in the PHM.

Data satisfying the determined requirements was found at SASOL Coal's Twistdraai plant at Secunda[1]. The Twistdraai plant is a coal beneficiation plant that seperates raw coal into different coal products according to client specifications. In September 1996 the plant was formally started up and ever since a vibration monitoring maintenance strategy has been used on 8 Warman® axial in, radial out pumps used to circulate a water and magnetite solution which is used in the washing process. Data recorded from these

[1] SASOL Coal has granted full permission to publish their name, data obtained from them as well as modelling results.

pumps during their operation was retrieved from the plant's meticulous Computerized Maintenance Management System (CMMS) and is used in this research project.

The data was modeled and analyzed in detail according to the theory described in Chapter 3 with close involvement of the vibration technicians who are monitoring the vibration of the pumps at the plant. Experience of these technicians was utilized in the mathematical modelling process by including their knowledge in the selection of covariates. Results obtained from mathematical models were also continually presented to the technicians for their interpretation and comments for improvement to make the final model truly useful in practice.

2 Preliminaries for PHM Analysis

Several futile searches for suitable data were undertaken before the data at SASOL Coal was discovered. During these searches a structured list of requirements for a PHM analysis was set up and used to assess the potential of a possible data source effectively and quickly. In this section the list of requirements is presented together with the shortcomings in general vibration data recording practices that were identified in the industry.

2.1 Requirements for a Vibration Covariate PHM Analysis

Requirements are defined under two main headings: (1) The suitability of an item for a vibration covariate PHM analysis; and (2) The availability of certain observations (data) throughout the item's working life.

2.1.1 Identification of a Suitable Item

First of all, a suitable item must be important enough for periodic diagnostic data collection, i.e. vibration measurements must be taken periodically (preferably at fixed intervals). If the item is important enough to be included in a vibration monitoring program, the cost of unexpected failure is usually considerably higher than the cost of preventive renewal. This is only a rule of thumb and it is not true in all cases. For situations where this is true, the optimal renewal time will be very distinct compared to a much more insensitive optimum for other scenarios.

The specific item must have been renewed on a number of occasions in the past, preferably because of failure. (The renewal assumption is thus made implicitly). Failure does not necessarily refer to a physical shutdown or destruction of the item, but to any condition where the item was unable to perform according to requirements, whereafter it had to be renewed. The preference of failure does not mean that preventive renewals are not important and all available data should be included in an analysis and handled suitably. If certain parts of available data is ignored, important information could be lost and biased estimates of the life time distribution will be the result, such as underestimation of the mean time to failure.

2.1.2 Required Information

Two main types of information have to be available:

(i) The operational age of an item at significant events (explained below) as well the event type. An operational age instant can be expressed in any suitable use parameter – in this case time will be used.
(ii) Diagnostic information (vibration levels) at every significant event.

Significant events mentioned above are any of the following:

(i) The moment when the item is brought into service.
(ii) Every point where diagnostic information is available.
(iii) Points in time where minor maintenance is done that could affect (usually reduce) covariate values, for example realignment, increased lubrication or balancing. The information on expected covariate values at these points should also be included in the data.
(iv) Time of renewal and the state of the item at renewal, i.e. failed or suspended.
(v) Data cutoff date where all operating units will be treated as calendar suspensions.

2.2 Shortcomings

Numerous shortcomings in data collection practices and data retrieving mechanisms of companies were discovered while searching for suitable data. Some of the major shortcomings most often encountered, are:

(i) Unfriendly or improperly organized Computerized Maintenance Management Systems or Enterprise Asset Management Systems.

(ii) Only the calendar age of a component is recorded and not the operational age, i.e. the real usage of the component.

(iii) Irregular inspections.

(iv) Records of maintenance done on a component that could have influenced its vibration levels are not recorded.

(v) The state of the item at the time of renewal is seldom recorded, i.e. whether a preventive renewal or failure renewal was performed.

(vi) A general lack of commitment exists regarding proper vibration monitoring documentation amongst managers of vibration monitoring programs.

The shortcomings mentioned above are all direct, major impairments of a successful vibration covariate PHM analysis, although improvements to these shortcomings could hold benefits for conventional vibration analysis techniques as well.

2.3 Concluding Remark

The information requirements stated in section (2.1.2.) were derived from the PHM theory described in Chapter 3. These requirements are defined for a best case scenario. It does not mean if these requirements are not met flawlessly that a PHM analysis is totally impossible. Mathematical manipulations and approximations allow for some deviation of the requirements as was also described in Chapter 3. Section (2.1.2.) should thus rather be used by analysts not familiar with PHM theory as detailed guidelines for a PHM feasibility analysis in his/her situation, than as strict prerequisites for a successful PHM analysis.

3 SASOL Data

Useful data was found at one of SASOL Coal's coal beneficiation plants, a part of the Twistdraai mine, situated at Secunda. The data does not strictly satisfy all the requirements outlined above – the main deviation being that no regular inspection frequency is used, although this is not a rigid prerequisite. This data set was the best one found following a fairly 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 1st, 1996 to November 1st, 1998 which gives an analysis time horizon of 791 days. The information recorded over the 791 days was used to estimate the PHM and for finding the optimal decision policy. Thereafter a second data set was collected from November 1st, 1998 to February 28th, 1999 that was used to evaluate the model's performance as if it was used to make renewal decisions in a real life situation. (Other techniques, apart from the second data set, were also utilized to test the optimal policy).

3.1 Background

A total of 8 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 8 pumps work under exactly the same conditions and it was assumed that renewals on the various pumps were generated by the same renewal process. Figure 3.1. below shows the pump installation layout, with the 8 pumps. Figure 3.2 shows a close-up of one of the pumps.

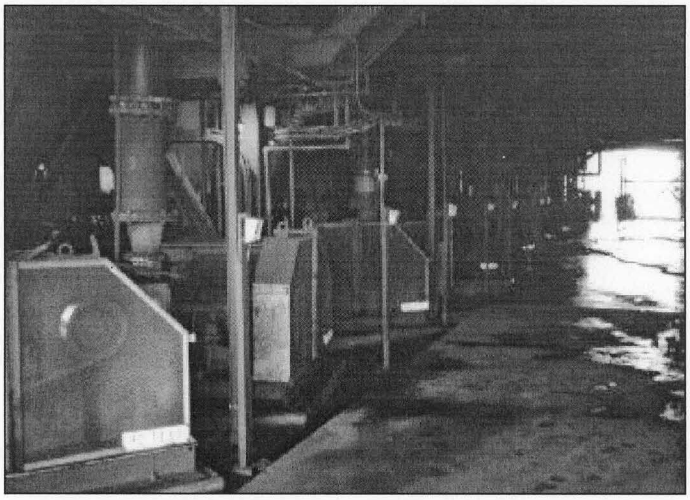

Figure 3.1.: *Pumps in operation*

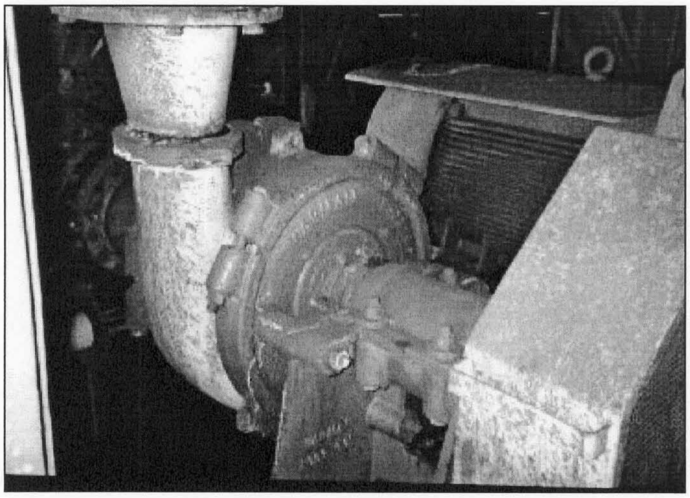

Figure 3.2.: *Warman pump*

When there is referred to a *pump* in this chapter, all the elements visible in Figure

3.2. are implied, except for the 220 kW electrical motors used to drive the pumps. A pump consists of an impeller housing, impeller, bearing housing, 2 SKF 938 932 bearings, a drive shaft, V-belt pulley and seals.

Because of the aggressive nature of the fluid being circulated and the robust environment of the pumps, total destructive failures are encountered frequently. These destructive failures often occur very abruptly, i.e. a pump's state literally change overnight from being in an acceptable condition to being completely failed. 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.
iv. Broken drive shaft.

When a pump has failed due to one of the reasons above, it is overhauled completely to an as-good-as-new condition 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.

During the analysis time horizon, the plant's management prescribed a condition based preventive renewal 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 3.3. on the next page shows the horizontal measuring positions.

The 'wet-end' bearing (the bearing closest to the impeller) is labeled as bearing number *3* while the 'dry-end' bearing is labeled as bearing number *4*. Measuring positions 3H and 4H are thus the horizontal measurements on bearing number 3 and 4, respectively. Only the horizontal measurements were used in this PHM analysis – reasons are presented later.

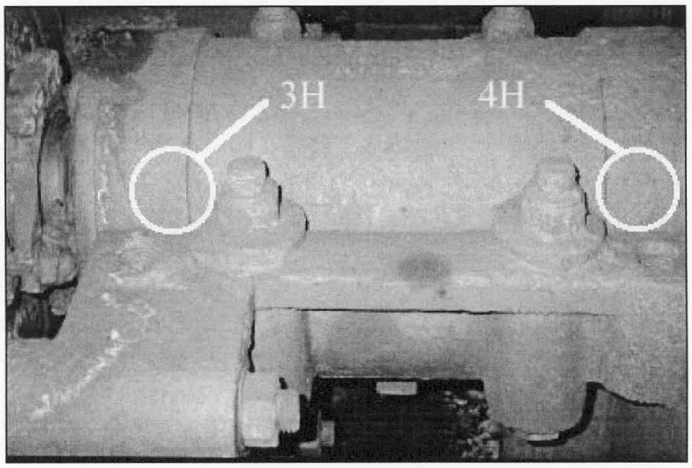

Figure 3.3.: *Monitoring positions*

As in most typical vibration monitoring programs, the renewal decisions of pumps were based on spectral vibration analysis. Several important frequencies are enveloped or benchmarked and renewal is performed as soon as two or three of the benchmarks are exceeded. Benchmarks 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 measurements are stored in the database and the chosen covariate levels (discussed later) could be retrieved easily and accurately.

The vibration measurement database does not contain information regarding events during a pump's life, nor does the plant's computerized maintenance management system (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, which probably does not effect the covariate levels too severely.

Failure analysis records obtained from the CMMS provided insight on the state of a pump when it was renewed, i.e. whether it was in the failed state or was suspended (preventively renewed).

3.2 Covariates

Covariate selection was based primarily 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 good enough indication of a pump's condition and that not much additional information is obtained from the vertical measurements. This corresponds to vibration theory and hence only the horizontal vibration measurements are considered.

As mentioned earlier, the vibration monitoring program that was used on the pumps was based on spectral analysis. A number of important frequencies (as defined by theory and experience) are monitored and a pump is renewed as soon as two or three of the frequencies' amplitudes exceed certain benchmarks. It was decided to use all of these frequencies as covariates in the PHM, thereby incorporating vibration theory and prior experience with the pumps in the model. Table 3.1. summarizes the 12 selected covariates.

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 may cause the proportional hazards model to become mathematically unstable or difficult to estimate, especially when the sample size is fairly small.

	Covariate Abbreviation	**Description**
1.	RF043H	0.4 x Rotational frequency amplitude, measured on bearing 3, indicative of a bearing defect.
2.	RF13H	1 x Rotational frequency amplitude, measured horizontally on bearing 3, indicative of unbalance in the pump.
3.	RF23H	2 x Rotational frequency amplitude, measured horizontally on bearing 3, indicative of misalignment in the pump.
4.	RF53H	5 x Rotational frequency amplitude, measured horizontally on bearing 3, indicative of cavitation in the pump.
5.	HFD3H	High frequency domain components between 1200-2400 Hz, measured on bearing 3, indicative of bearing defect. This is a subjective covariate where 1 indicates the presence and 0 the 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 the presence and 0 the absence of the mentioned components.

7.	RF044H	0.4 x Rotational frequency amplitude, measured on bearing 4, indicative of a bearing defect.
8.	RF14H	1 x Rotational frequency amplitude, measured horizontally on bearing 4, indicative of unbalance in the pump.
9.	RF24H	2 x Rotational frequency amplitude, measured horizontally on bearing 4, indicative of misalignment in the pump.
10.	RF54H	5 x Rotational frequency amplitude, measured horizontally on bearing 4, indicative of cavitation in the pump.
11.	HFD4H	High frequency domain components between 1200-2400 Hz, measured on bearing 4, indicative of bearing defect. This is a subjective covariate where 1 indicates the presence and 0 the 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 the presence and 0 the absence of the mentioned components.

Table 3.1.: *Summary of covariates*

3.3 Data

The data collected include the pump unit identification, dates of inspection, vibration frequency spectrum at each inspection, date of failure or suspension and the state at renewal, 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.

A total of 27 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 – Begin or pump startup; (2) ES – Event suspension; and (3) EF – Event failure. Events that occurred to the pumps are listed in Table 3.2. below:

Pump Identification	Age (days)	Date	Event
PC1131	0	9/1/96	B
PC1131	397	10/3/97	ES
PC1131	397	10/3/97	B
PC1131	554	3/9/98	EF
PC1131	554	3/9/98	B
PC1131	690	7/23/98	ES
PC1131	690	7/23/98	B
PC1131	765	10/6/98	EF
PC1131	765	10/6/98	B
PC1131	791	11/1/98	ES
PC1132	0	9/1/96	B
PC1132	491	1/5/98	EF
PC1132	491	1/5/98	B
PC1132	544	2/27/98	ES
PC1132	544	2/27/98	B
PC1132	557	3/12/98	ES
PC1132	557	3/12/98	B
PC1132	751	9/22/98	EF
PC1132	751	9/22/98	B
PC1132	791	11/1/98	ES
PC1231	0	9/1/96	B
PC1231	563	3/18/98	EF
PC1231	563	3/18/98	B
PC1231	578	4/2/98	ES
PC1231	578	4/2/98	B
PC1231	791	11/1/98	ES
PC1232	0	9/1/96	B
PC1232	599	4/23/98	ES
PC1232	599	4/23/98	B
PC1232	791	11/1/98	ES
PC2131	0	9/1/96	B
PC2131	184	3/4/97	EF
PC2131	184	3/4/97	B
PC2131	470	12/15/97	ES
PC2131	470	12/15/97	B
PC2131	631	5/25/98	EF
PC2131	631	5/25/98	B
PC2131	774	10/15/98	EF
PC2131	774	10/15/98	B
PC2131	791	11/1/98	ES
PC3131	0	9/1/96	B
PC3131	450	11/25/97	EF
PC3131	450	11/25/97	B
PC3131	791	11/1/98	ES
PC3132	0	9/1/96	B
PC3132	506	1/20/98	EF
PC3132	506	1/20/98	B
PC3132	791	11/1/98	ES
PC3232	0	9/1/96	B
PC3232	563	3/18/98	EF
PC3232	563	3/18/98	B
PC3232	723	8/25/98	ES
PC3232	723	8/25/98	B
PC3232	791	11/1/98	ES

Table 3.2.: *Events table*

Detailed inspection data of all the covariate measurements between events is provided as an appendix to this chapter. Covariate values immediately after the occurrence of an event were all taken to be zero. Further detailed comments on the data are presented below:

1. 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 noticeable fact is that these values occurred at suspensions.

   The most logical physical explanation for these values lies in the wear mechanism present in the bearing. (RF043H is indicative of a particular bearing defect). It could be that the bearings that produced these extreme values were able to withstand the wear associated with RF043H, i.e. did not abrade with the introduction of the RF043H vibration, which would have kept the vibration levels within normal limits. The vibration levels continued to rise up to their outrageously high values, which persuaded management to renew the pumps preventively.

2. Subjective covariates HFD3H, HFD4H, LNF3H and LNF4H indicated the presence of their associated phenomena with a simple 0 or 1. These phenomena certainly appear in different degrees of severity and it could be argued that covariates that quantify the severity could lead to a more accurate model. 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 try 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.

   Intuitively it is expected that whenever one of the considered 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 will be present in a certain inspection but will be absent in the following, 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 600-1200Hz 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.

   Interaction between the subjective covariates and the quantitative covariates is an area which should be investigated in a study such as this.

3. Failure times are distributed such that 6 failures occurred below 200 days and the remaining 5 failures above 450 days (not randomly distributed). Suspension times are randomly distributed with some being very short like 53, 15 and 13. The question is whether these renewal patterns can be explained by the covariates.

4. 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 (3).

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 $C_f = R\,162\,200$ will be used and a preventive renewal cost of $Cp = R\,25\,000$. These costs were average costs sustained by the Twistdraai plant over the two years over which the data was collected. No details are available.

4 Weibull PHM Fit

There is no straightforward procedure to select the most appropriate covariates for a good 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 used to get to the best possible model.

Some important facts and guidelines concerning vibration covariates and vibration covariate selection for the PHM were discovered and established from experience gained in this research project:

(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, it can produce very uncertain estimates (large standard errors) which will make them appear as insignificant, even if one of them could be a very 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 the model, because their parameters could be very inaccurate and produce a misleading model. They could be checked again when more data is collected.

(iv) Positive covariates with negative regression coefficients should be considered with special care, because it indicates that the hazard 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 interpretation. This almost always indicates some data problem, particularly if wrong covariate values are reported at failures, because failure information has a large influence on the maximum likelihood. .

An extensive discussion about practical analysis of covariate data and modeling procedures can be found in [5].

The following theory, described in Chapter 3, will be used to fit the PHM:

- The numerically convenient method of maximum log-likelihood as objective function in the optimization routines. (Chapter 3, section 2.3.2.).
- For estimation of the parameters in the objective function, the Newton-Raphson optimization method because of its rapid convergence. Snyman's method will only have a supportive role in the modelling process because of its robustness. (Chapter 3, section 4.).
- Guidelines on covariate behavior and selection will constantly be referred to in the modelling process. (Chapter 3, section 3.).
- Residual plots as graphical indication of the goodness-of-fit of models because all the PHM assumptions can be evaluated by analyzing the residuals. The Kolmogorov-Smirnov test, Wald test, p-values and standard errors will be used as analytical goodness-of-fit tests. (Chapter 3, sections 5.1. and 5.2.).

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; (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.

The results of the modelling procedures are presented below.

4.1 Phase 1: Simple Weibull Model

The simple Weibull model was calculated to be:

$$h(t) = \frac{1.984}{468.7} \cdot \left(\frac{t}{468.7} \right)^{0.984} \qquad (4.1.)$$

Residual plots of the model in (4.1.) yield the following:

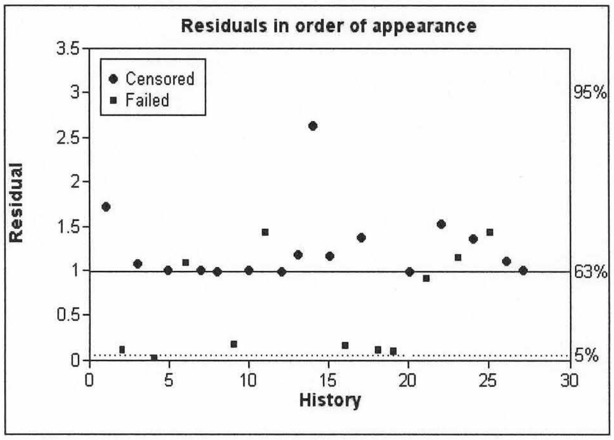

Figure 4.1.: *Residuals in order of appearance*

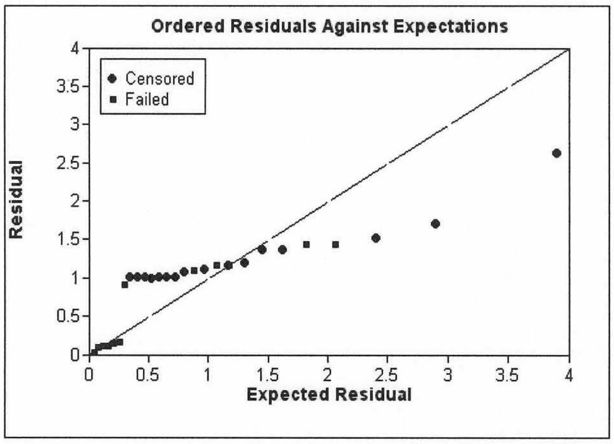

Figure 4.2.: *Ordered residuals agains expectations*

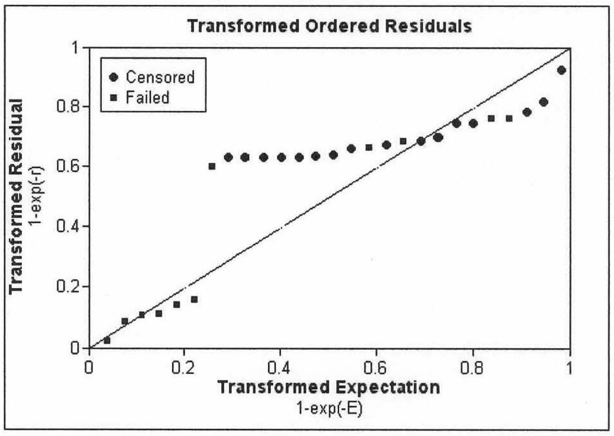

Figure 4.3.: *Transformed ordered residuals*

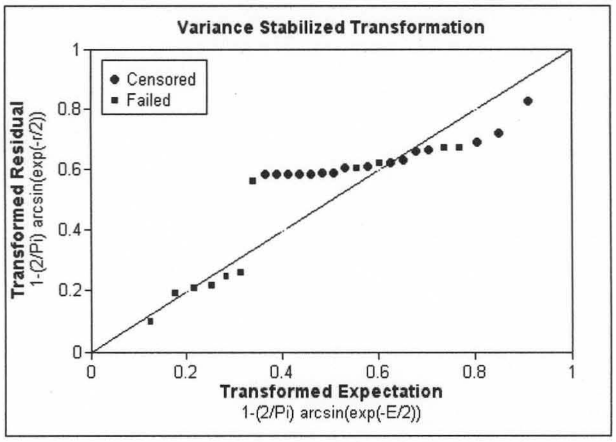

Figure 4.4.: *Variance stabilized transformation*

From the residual plots it is clear that the model does not represent the data very well, especially the suspended observations. Analytical tests revealed a standard error for the shape parameter of 0.46, which is significantly different from 1, while the standard error for the scale parameter is 71.9 days, showing that the model is not very accurate but still a useful estimate. The $MTTF = 415.5$ which is realistic. The Kolmogorov-Smirnov test (KS-test) yielded a value of $KS = 0.3949$ with a $p-\text{value} = 0.000276$, which rejects the fit at a 5 % level of significance. (A time-independent model, i.e. $\beta = 1$ fixed, was also tested but rejected based on an observed value for the Wald test of 4.57 with a $p-\text{value} = 0.0325$).

4.2 Phase 2: Weibull PHM with Subjective Covariates Excluded

For this phase, HFD3H, HFD4H, LNF3H and LNF4H were excluded from the modelling process. A reasonable model fit was obtained when using all of the remaining quantitative covariates, i.e. RF13H, RF14H, RF23H, RF24H, RF53H, RF54H, RF043H and RF044H, with negative regression coefficients for RF13H and RF14H. This is consistent with the observed behavior of the data (see comment 5 of section 3.3.).

Using mainly backward selection with an upper Wald p-value limit of 5 %, the best possible model was obtained by using only the two covariates associated with cavitation, RF53H and RF54H. The model is presented below as (4.2.).

$$h(t,\overline{z(t)}) = \frac{1.464}{1431.8}\cdot\left(\frac{t}{1431.8}\right)^{0.464}\exp\left(0.127\cdot \text{RF53H} + 0.143\cdot \text{RF54H}\right) \qquad (4.2.)$$

The following results were obtained with residual analyses of (4.2.):

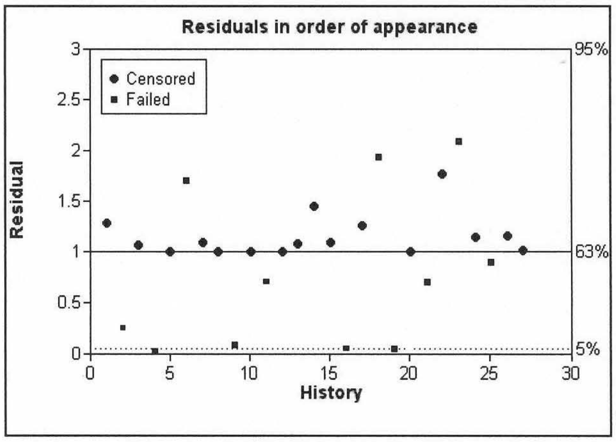

Figure 4.5.: *Residuals in order of appearance*

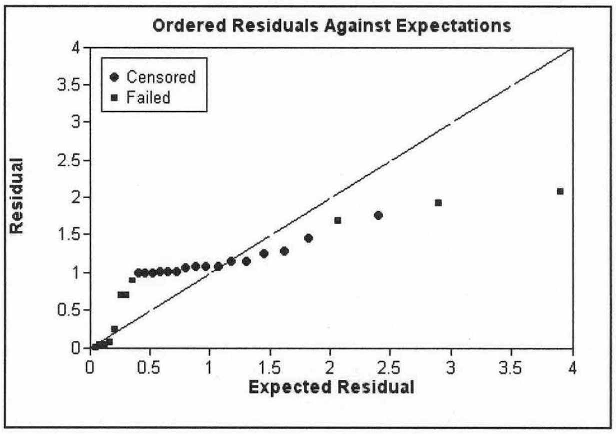

Figure 4.6.: *Ordered residuals against expectation*

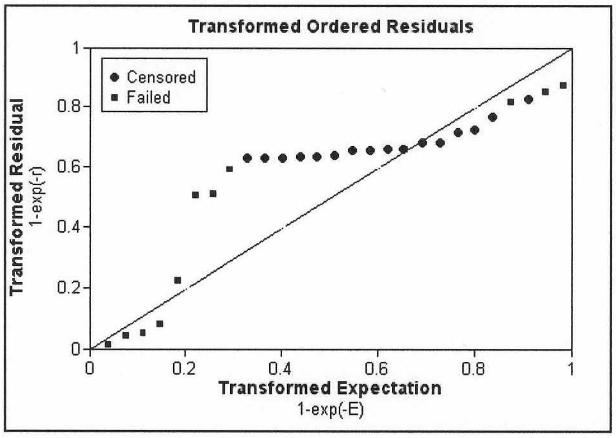

Figure 4.7.: *Transformed ordered residuals*

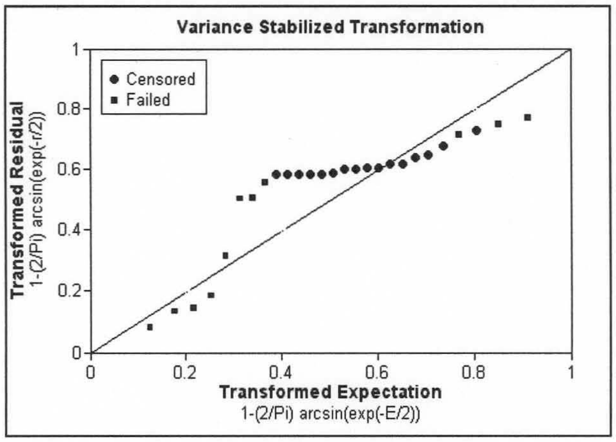

Figure 4.8.: *Variance stabilized transformation*

The results of analytical significance tests on the parameters are summarized in Table 4.1. It is clear that both RF53H and RF54H are very significant in the failure process although the shape parameter did not prove to be significant. The KS-test was determined to be $KS = 0.3180$ with a *p*-value of 0.00628, which is not an extremely good model fit.

	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

Table 4.1.: *Results of analytical goodness-of fit tests performed on (4.2.)*

The graphs obtained from the residual analysis show that 4 of the 6 short failures (see comment 4 in section 3.3.) cannot be explained well by the model (e.g. with high covariate values). The data was analyzed and it was found that no other quantitative covariate contributed significantly to these early failures. Further analyses of the data revealed that the contribution of RF53H and RF54H to the other, longer, failures is evident. The model with RF53H only was also considered, and a better model fit was obtained $KS\ p-\text{value} = 0.145$. Still, this model did not explain the 4 short failures any better.

It was noticed that for all considered models, the shape was not significant although the hypothesis of $\beta = 1$ was never rejected with Wald *p*-values between 0.18 and 0.36, except for the model with RF53H only, where the Wald *p*-value was calculated to be 0.062. (Values of β ranged from 1.4 to 2). After this observation, models with $\beta = 1$ were hence estimated, and in all cases better model fits were obtained than with $\beta \neq 1$ (Wald *p*-values >10 %). It is important to note however that RF53H and RF54H were, as before, the only two significant covariates (after a process of backward covariate selection based on Wald p-values). This implies that time (working age) is not a significant variable in the model and that some of the failures could be better explained by an additional non-observed covariate. Vibration technicians do not agree with this statement and the problem possibly lies in a too short data horizon.

4.3 Phase 3: Weibull PHM With All Covariates

For this phase, the inspection data was analyzed with all the covariates. Inspection showed that subjective covariates are somehow, 'complementary' to the numerical covariates, i.e. at failures the majority of them have the value 1 if numerical covariates are low, and mostly the value 0, if numerical covariates are high.

To get a feel for the behavior of the subjective covariates, they were first analyzed separately. Only LNF4H appeared to be significant, with test model fit KS p-value = 0.17, which is acceptable. It was further noticed when LNF3H and LNF4H are in the model, their regression coefficients have the opposite signs and the same for HFD3H and HFD4H, as they tend to compensate each other. The data was analyzed again and the correlation coefficient for LNF3H and LNF4H was calculated to be 0.57, and for HFD3H and HFD4H to be 0.80. The high correlation between HFD3H and HFD4H could be because of the similar configuration (and operating conditions) of the bearings and a lack of lubrication will affect both bearings.

The next step was to build a model with all covariates included. Estimation procedures (both Snyman and Newton-Raphson) failed to converge initially, with the scale parameter approaching infinity. In such a case, it is not simple to decide which covariate to exclude from the model. By looking at the highest partial derivatives in the model fitting optimization routine, it was decided to exclude RF043H from the model (this could be because of the few unproportionally high observations). Still the estimation procedure would not converge and it was necessary to exclude more covariates from the model, in different combinations, to get convergence. These covariates were considered in the model at later stages, to check whether their removal was not only due to some relationship with other covariates.

A good example of this was RF53H, which was removed in one of the procedures at an early stage, and when later considered showed high significance. When both RF53H and RF54H were included in a model, they appeared as significant, but with a poor model fit, due to the large residuals when both covariates have high values. Their correlation was found to be 0.60 (one measurement with a very high value for RF53H and a very low value for RF54H was excluded). It shows that both these covariates are good predictors of failures but it still has to be decided whether it makes practical sense to include both in the final model.

When RF53H was removed from the model (either at an early stage, as mentioned above, or to improve the model fit), RF54H and LNF4H remained in the model, with a *p*-value for the scale parameter of 0.416, and the model fit KS *p*-value = 0.015, which is not very good. When β was fixed to 1, a much improved model fit was obtained (KS *p*-value = 0.647). From the residuals it appears that some of short failures could be better explained by this model, than by the model without LNF4H. The sum of LNF3H and LNF4H was also included as a covariate in the model with RF54H, and the model showed a very good model fit for both estimated values of β, and β fixed to 1. This definitely shows that subjective covariates could be useful in the pumps' condition diagnosis.

4.4 Final Model

The analyses showed clearly that RF53H and RF54H are the two most significant covariates in the data and will hence be used in the final model. It was also decided that the shape parameter should not be restricted in the final model although better model fits were obtained with $\beta = 1$. The only reason for the good performance of the model with $\beta = 1$ could be because of a shortage of data and it was concluded that a model with $\beta \neq 1$ would be of more practical use.

The final PHM with which the decision models will be constructed is presented below as (4.3.), previously (4.2.):

$$h(t, \overline{z(t)}) = \frac{1.464}{1431.8} \cdot \left(\frac{t}{1431.8} \right)^{0.464} \exp\left(0.127 \cdot \text{RF53H} + 0.143 \cdot \text{RF54H}\right) \qquad (4.3.)$$

5 Decision Model

This section describes the construction of the transition probability matrices (TPM) as well as the calculation of the optimal cost function. The policy is evaluated theoretically by applying it on the observed data but also evaluated practically on a second data set collected from November 1st, 1998 to February 28th, 1999, as if it was used in a real life situation.

5.1 Transition Probability Matrices

The optimal policy is very sensitive to the choice of covariate bands and it is thus very important to choose these bands with great care. As explained in Chapter 3 is it generally recommended to select the lower bands shorter than the upper bands especially for vibration covariates, because vibration covariate values tend to be closely grouped under normal wear-out of a component and outlier values only occur sporadically. The data used for this research also shows this behavior. See the PDF's of the two covariates below, represented by continuous Weibull distributions:

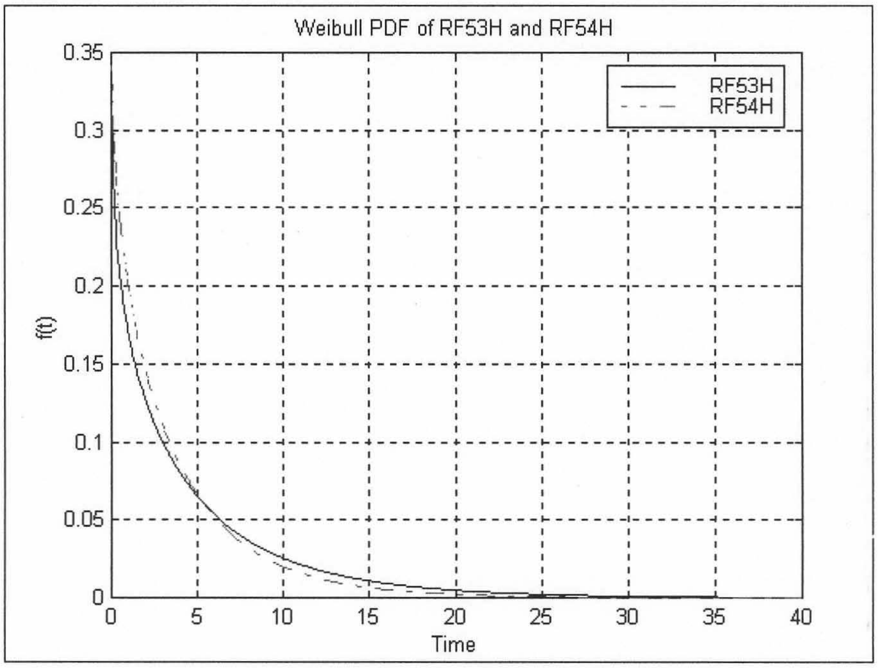

Figure 5.1.: *PDF of observed RF53H values*

After evaluation of Figure 5.1., the following bands were selected that resulted in realistic cost models (explained later):

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

Table 5.1.: *Selected bands and observed frequencies*

With the above covariate bands the transition rates were determined and transition matrices were calculated. For example, the transition probabilities for covariate RF53H for an observation interval of 50 days are given in Table 5.2.

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

Table 5.2.: *TPM for RF53H (for observation interval of 50 days)*

From the table it can be seen that if RF53H is currently between 0 and 5, then after 50 days it will be still within the same limits, with a probability of 91.3%. If it is currently between 5 and 10, it will stay there with a probability 48.1%, but it can also decrease, with the probability 20.8%, i.e. it can improve. This is very realistic in practice since vibration levels most often increase with the deterioration process but it can sometimes decrease because of specific wear mechanisms present in the component as was observed in the data. Similarly was the TPM for RF54H determined. See Table 5.3.

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

Table 5.3.: *TPM for RF54H (for observation interval of 50 days)*

5.2 Cost Function and Optimal Replacement Policy

As mentioned earlier, were the costs provided by the Twistdraai plant, $C_f = R\,162\,200$ and $C_p = R\,25\,000$ 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 Chapter 3, section 6.2.2.) and a future inspection interval of 50 days was used for the cost model. With all preliminary calculations completed, the cost function (equation (6.17.) of Chapter 3) was hence calculated using the backward recursive procedure. The result is shown graphically in Figure 5.2. in terms of the threshold risk level, d (or $h(t, \overline{z(t)}) \cdot K$) for convenience.

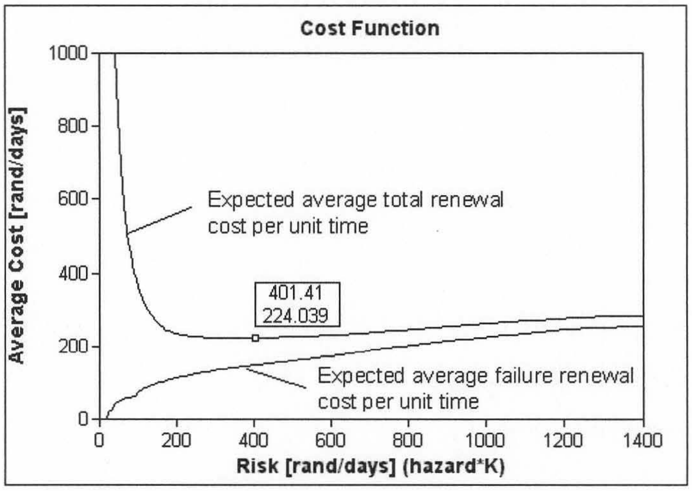

Figure 5.2.: *Estimated cost function in terms of risk*

A distinct optimum exist at a risk of R 401.41 / day or a hazard rate of $h = 0.0029$. This optimum is not very sensitive to slight deviations from the decision rule. With the optimal risk known it is also possible to represent the replacement rule and warning level function graphically (equations (6.24.) and (6.26.) of Chapter 3):

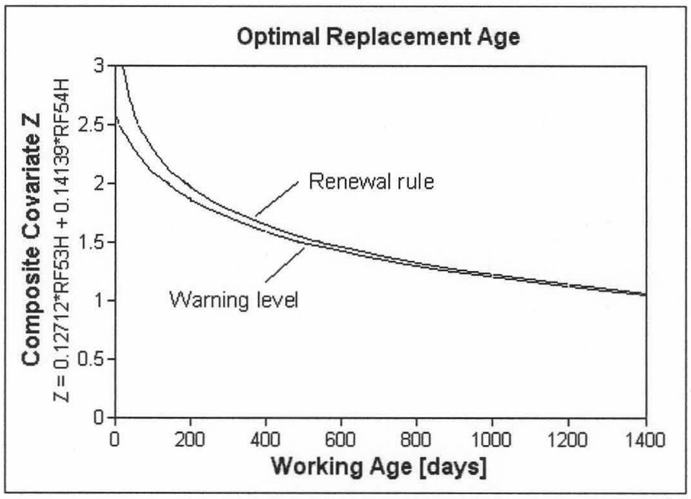

Figure 5.3.: *Decision policy*

5.3 Evaluation of Optimal Renewal Policy

First a summary of the performance of the optimal renewal policy is presented in Table 5.4. below, whereafter detailed comments follow.

	Theoretical Optimal Policy	Renew Only at Failure Strategy	Theoretical Policy Applied	Real 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	76.7%	0%	80.0%	42.1%
Failure Renewals	23.3%	100%	20.0%	57.9%
MTTR	254.49 days	404.08 days	263.6 days	214.6 days

Table 5.4.: *Summary of renewal policy performance (All costs are in R/day)*

The table shows that the theoretical model predicts an average cost of R 224.4 / day when using the calculated optimal renewal policy, with 66.4% of the cost due to failures, although failures only occur 23.3% of the time. This is due to a relatively high renewal cost ratio of R 162 200 / R 25000 = 6.5. The average time between renewals is calculated to be 254.5 days. If no renewal policy is used, except at failures, it would result in a mean time between failures of 404.1 days, close to an estimate of 415.5 days obtained from the simple Weibull model (see section 4.1.), but with an average cost of renewal of R 401.4 / day. This would be 44.2% more expensive than using the optimal policy.

To evaluate the above mentioned theoretical costs, it should be compared with: (a) the real replacement costs realized for the analyzed histories, (b) the cost that would be obtained if the theoretical optimal policy was used for the analyzed histories

(a) It is very important to realize that there are two options when using real histories for the cost calculation. Every failure or suspension (preventive renewal) has a clearly defined cost, either C_f or C_p, but this is not the case for temporary suspensions or calendar suspensions. A conservative approach is to exclude all temporary suspensions from the calculation (TSE method) or a less conservative method is to include them all in the calculation as true suspensions (TSI method). The TSI method could be justified by counting the replacement cost at the beginning of the history as an installation cost, so that the calculated average replacement cost would be a "current" average cost. With a large number of histories, and not many temporary suspensions, both methods will give similar results. With a small number of histories and many temporary suspensions, the TSE method usually gives an overestimation of the real average cost value. Using both methods, the real average replacement cost for the pumps over the analysis horizon was R 345.16 / day (using the TSI method, counting 11 failures and 16 suspensions), or R 385.58 / day (using the TSE method, counting 11 failures and 8 suspensions). So, if the TSI method cost is compared to the theoretical optimal cost, the saving would be (345-224)/345 = 35%. The real policy is slightly better than the policy to replace only at failure, with a saving of (401-345)/401 = 14%. The real average time to renewal is 214.6 days, calculating only completed histories (failures and true suspensions). The theoretical mean time to renewal is 254.5 days which can also be considered as an advantage of the theoretical optimal policy.

(b) The optimal theoretical decision policy is applied on all 27 histories. Three situations are considered: (i) immediate renewal based on the most recent inspection record; (ii) renewal based on an earlier inspection record (with that renewal time counted); and (iii) no renewal based on all inspection records. After appying the theoretical policy, the number of failures was reduced from 11 to 4, which is then 4/19 = 21% of all renewals (temporary suspensions excluded), close to the theoretical value of 23%. Renewal times were not significantly reduced, which resulted in a significant reduction of the average cost. Using the TSI method, the average renewal cost is R 214 / day, close to the theoretical cost of R 224 / day, so the real saving would be (345-214)/345 = 38%. Using the TSE method, the average cost is R 215 / day (one real temporary suspension is included in the calculation as a definite suspension, due to (ii)), surprisingly close to the previous value. The average renewal time is 263.6 days (7 undecided temporary suspensions excluded), close to the theoretical value of 254.5 days.

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 statistical and decision models are reasonable. The method of comparison could be argued because the same data is used to build the model and to test it. The method can be justified however by noticing that the data is first used to build the statistical model and then to calculate the optimal decision policy, without refering to the actual renewal policy. Theoretically, the same statistical model would be obtained (within the range of a statistical error), even if the actual policy was to renew only at failure. With a larger data set (more histories) other methods can be used, such as to use a random sample of histories to build the model, and then the rest as a control group to test the model.

As a final test of the renewal decision policy's performance, more data was collected from the plant from November 1st, 1998 to February 28th, 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 1st, 1998. The pump eventually failed unexpectedly 67 days later on January 6th, 1999 at an age of 259 days. A total of five inspections were done during this time. The latest inspection data is plotted on Figure 5.4. below, together with the 4 inspections from the first data set.

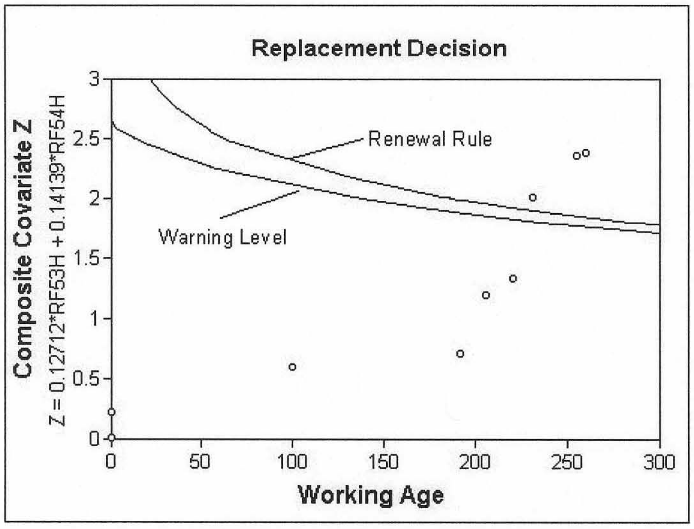

Figure 5.4.: *Example policy on PC1232*

Figure 5.4. shows clearly that the expensive unexpected failure could have been prevented if the calculated decision policy was followed. In terms of cost, the unexpected failure cost meant R 162 200 / 265 days = R 612.07 / day. If the PHM renewal policy was available and there was acted upon, R 25 000 / 235 days = R106.38 / day, would have been the result. This is another confirmation that the model is relevant and practical.

6 Conclusion

Although the final PHM was not statistically speaking a very accurate model, it proved to be of high practical value. The two covariates used in the model, RF53H and RF54H, were correctly identified as good predictors of events, somewhat contrary to statistical recommendations. This shows that a PHM analysis can never be done away from practice otherwise costly misinterpretations may be the result. The calculated optimal policy also withstood thorough evaluation and clearly showed its enormous benefits even with conservative assumptions.

Appendix

Pump Identification	Age (days)	Date	RF043H [mm/s]	RF13H [mm/s]	RF23H [mm/s]	RF53H [mm/s]	HFD3H [mm/s]	LNF3H [mm/s]
PC1131	159	2/7/97	0.00	0.70	0.30	0.80	1	0
PC1131	295	6/23/97	0.15	0.30	0.25	0.55	0	1
PC1131	387	9/23/97	0.30	3.00	0.90	8.00	1	0
PC1131	394	9/30/97	0.80	2.40	1.00	12.30	1	0
PC1131	397	10/3/97	250.00	175.00	20.00	17.00	1	0
PC1131	530	2/13/98	0.10	11.50	3.20	11.00	0	0
PC1131	533	2/16/98	0.30	8.80	3.50	13.00	1	0
PC1131	554	3/9/98	0.50	7.00	3.80	16.00	0	0
PC1131	578	4/2/98	1.00	19.50	1.50	2.00	1	0
PC1131	597	4/21/98	0.30	27.50	1.50	1.60	1	0
PC1131	639	6/2/98	0.50	31.00	6.00	4.00	1	0
PC1131	689	7/22/98	0.00	9.00	2.00	0.80	0	0
PC1131	690	7/23/98	0.00	8.27	1.82	0.67	0	0
PC1131	703	8/5/98	0.05	1.20	0.95	0.20	1	0
PC1131	712	8/14/98	0.05	0.50	0.80	1.40	1	0
PC1131	765	10/6/98	0.05	0.40	0.70	2.70	1	0
PC1131	791	11/1/98	0.50	9.00	2.00	12.00	0	0
PC1132	239	4/28/97	0.00	0.90	0.30	1.50	0	0
PC1132	386	9/22/97	0.10	7.00	0.60	2.10	1	0
PC1132	394	9/30/97	0.20	8.00	0.50	11.00	1	0
PC1132	397	10/3/97	0.10	6.20	0.20	3.00	0	0
PC1132	491	1/5/98	0.10	5.00	0.50	1.00	0	0
PC1132	499	1/13/98	0.10	27.50	2.00	2.50	0	0
PC1132	533	2/16/98	0.10	35.00	2.50	12.00	0	0
PC1132	543	2/26/98	5.00	19.00	26.00	9.00	0	0
PC1132	544	2/27/98	5.61	16.94	28.93	8.56	0	0
PC1132	557	3/12/98	3.00	43.00	9.00	2.00	0	0
PC1132	558	3/13/98	1.00	41.00	14.00	3.00	0	0
PC1132	597	4/21/98	4.00	29.00	3.70	2.60	0	1
PC1132	689	7/22/98	0.10	5.60	1.70	0.30	0	1
PC1132	712	8/14/98	0.10	3.40	0.60	0.90	0	1
PC1132	751	9/22/98	0.99	3.01	0.30	2.99	0	1
PC1132	791	11/1/98	0.08	4.65	0.17	2.01	0	0
PC1231	239	4/28/97	0.30	5.50	1.90	1.00	0	0
PC1231	295	6/23/97	1.30	10.40	2.20	1.00	0	0
PC1231	390	9/26/97	1.00	56.00	12.00	3.00	0	0
PC1231	530	2/13/98	0.30	18.10	6.10	8.50	1	0
PC1231	563	3/18/98	0.09	12.00	1.18	10.24	1	0
PC1231	578	4/2/98	1.00	33.00	18.00	6.00	1	1
PC1231	653	6/16/98	0.22	3.57	0.98	0.57	0	0
PC1231	698	7/31/98	0.68	8.11	1.47	0.61	0	0
PC1231	791	11/1/98	0.73	38.64	7.68	1.86	0	0
PC1232	583	4/7/98	0.50	56.00	9.00	4.00	0	0
PC1232	592	4/16/98	0.40	54.00	4.00	6.50	0	0

PC1232	597	4/21/98	0.60	48.00	9.00	3.50	0	0
PC1232	599	4/23/98	0.05	7.00	2.10	0.60	1	1
PC1232	699	8/1/98	0.33	34.16	5.76	2.48	0	0
PC1232	791	11/1/98	0.24	32.40	2.44	4.09	0	0
PC2131	156	2/4/97	0.00	9.00	1.20	0.40	0	0
PC2131	159	2/7/97	0.10	5.80	2.20	0.60	0	1
PC2131	178	2/26/97	0.20	4.00	3.30	1.35	0	1
PC2131	179	2/27/97	0.00	8.30	2.00	0.90	0	0
PC2131	184	3/4/97	0.00	36.39	2.00	1.00	0	1
PC2131	239	4/28/97	0.09	3.65	1.60	1.55	1	0
PC2131	241	4/30/97	0.05	3.10	0.75	1.70	1	0
PC2131	295	6/23/97	0.10	2.55	2.20	1.40	1	0
PC2131	386	9/22/97	0.40	5.60	7.50	0.70	1	0
PC2131	470	12/15/97	1200.00	120.00	30.00	10.00	0	0
PC2131	535	2/18/98	0.20	20.90	1.60	4.80	0	0
PC2131	583	4/7/98	2.00	77.00	46.00	11.00	0	0
PC2131	597	4/21/98	2.00	66.00	43.00	6.00	0	0
PC2131	604	4/28/98	1.00	74.00	37.50	5.00	1	0
PC2131	611	5/5/98	0.01	20.00	4.10	11.60	1	0
PC2131	631	5/25/98	0.10	18.00	10.00	72.33	1	0
PC2131	640	6/3/98	0.60	10.50	2.80	5.90	1	0
PC2131	689	7/22/98	0.09	1.70	0.40	0.50	1	0
PC2131	768	10/9/98	0.10	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	11/1/98	0.16	13.37	1.08	3.69	0	0
PC3131	241	4/30/97	0.10	6.80	3.90	1.30	1	0
PC3131	295	6/23/97	0.80	29.00	17.00	14.00	1	0
PC3131	386	9/22/97	0.50	37.00	6.50	4.00	1	0
PC3131	450	11/25/97	0.20	20.52	6.00	3.00	1	0
PC3131	550	3/5/98	0.09	7.20	3.74	1.27	1	0
PC3131	651	6/14/98	0.96	33.06	17.34	16.80	1	0
PC3131	750	9/21/98	0.59	40.33	6.43	4.16	1	0
PC3131	791	11/1/98	0.20	19.48	5.82	3.39	1	0
PC3132	239	4/28/97	0.10	2.40	0.15	0.39	1	0
PC3132	295	6/23/97	0.20	9.60	1.80	1.60	1	1
PC3132	386	9/22/97	0.20	24.00	3.00	3.50	1	1
PC3132	450	11/25/97	0.50	32.00	21.00	13.00	0	0
PC3132	506	1/20/98	0.97	37.56	48.37	26.84	0	0
PC3132	566	3/21/98	0.12	2.44	0.16	0.45	1	1
PC3132	711	8/13/98	0.19	11.04	1.92	1.82	1	1
PC3132	791	11/1/98	0.20	27.60	3.27	3.39	1	1
PC3232	239	4/28/97	0.30	11.50	3.80	0.60	1	0
PC3232	295	6/23/97	1.00	43.00	8.00	6.00	1	0
PC3232	386	9/22/97	2.00	39.00	6.00	6.00	1	0
PC3232	535	2/18/98	0.00	66.00	44.00	7.00	0	0
PC3232	563	3/18/98	0.00	75.72	56.86	7.33	1	0
PC3232	591	4/15/98	0.00	235.00	22.00	10.00	0	0
PC3232	604	4/28/98	2.00	175.00	18.00	7.00	0	0

PC3232	639	6/2/98	3.00	74.00	9.00	3.00	0	0
PC3232	722	8/24/98	0.00	20.50	14.80	1.90	1	1
PC3232	723	8/25/98	0.00	21.45	15.10	1.96	1	1
PC3232	748	9/19/98	0.18	7.59	2.96	0.39	1	0
PC3232	783	10/24/98	0.62	26.66	5.44	4.50	1	0
PC3232	791	11/1/98	1.28	28.08	3.72	4.08	1	0

Table A.1.: *Inspection data for bearing 3*

Pump Identification	Age (days)	Date	RF044H [mm/s]	RF14H [mm/s]	RF24H [mm/s]	RF54H [mm/s]	HFD4H [mm/s]	LNF4H [mm/s]
PC1131	159	2/7/97	0.05	0.85	0.30	0.10	1	0
PC1131	295	6/23/97	0.20	0.45	0.25	0.12	0	1
PC1131	387	9/23/97	0.10	4.00	1.70	6.20	1	0
PC1131	394	9/30/97	2.30	4.00	2.10	5.00	0	0
PC1131	397	10/3/97	4.00	4.60	2.80	6.00	1	0
PC1131	530	2/13/98	0.10	13.20	3.50	5.50	0	0
PC1131	533	2/16/98	0.20	10.00	3.80	7.00	1	0
PC1131	554	3/9/98	0.30	5.00	4.20	10.00	0	0
PC1131	578	4/2/98	0.70	42.00	3.00	3.00	1	0
PC1131	597	4/21/98	0.50	52.00	2.00	5.00	1	0
PC1131	639	6/2/98	0.50	47.00	8.00	5.00	1	0
PC1131	689	7/22/98	0.00	14.00	2.00	1.20	0	0
PC1131	690	7/23/98	0.00	13.04	1.73	1.08	0	0
PC1131	703	8/5/98	0.20	2.25	0.90	0.40	1	0
PC1131	712	8/14/98	0.05	0.58	1.30	0.41	1	1
PC1131	765	10/6/98	0.05	0.40	2.10	0.60	1	1
PC1131	791	11/1/98	0.20	12.00	2.00	7.00	0	0
PC1132	239	4/28/97	0.00	1.65	0.30	0.72	0	1
PC1132	386	9/22/97	0.10	12.20	0.70	7.80	1	0
PC1132	394	9/30/97	0.10	14.00	0.90	8.20	1	0
PC1132	397	10/3/97	0.20	12.00	0.90	12.00	1	0
PC1132	491	1/5/98	1.00	10.00	0.80	30.00	1	0
PC1132	499	1/13/98	0.10	66.00	4.00	12.00	0	0
PC1132	533	2/16/98	0.00	65.00	3.00	10.00	0	0
PC1132	543	2/26/98	1.00	120.00	38.00	7.00	0	0
PC1132	544	2/27/98	1.13	126.88	42.38	6.64	0	0
PC1132	557	3/12/98	1.00	34.00	5.00	2.50	1	0
PC1132	558	3/13/98	2.00	27.50	6.50	1.00	0	0
PC1132	597	4/21/98	1.00	24.00	4.20	5.40	0	1
PC1132	689	7/22/98	0.10	4.80	0.70	0.40	0	0
PC1132	712	8/14/98	0.05	2.70	0.30	0.40	0	0
PC1132	751	9/22/98	0.13	1.61	0.06	1.54	0	1
PC1132	791	11/1/98	0.15	7.80	0.56	7.68	1	0
PC1231	239	4/28/97	0.00	9.00	0.60	0.40	0	0
PC1231	295	6/23/97	0.30	16.50	2.30	0.30	0	0
PC1231	390	9/26/97	0.00	67.00	6.00	4.00	0	0
PC1231	530	2/13/98	0.00	21.00	6.00	6.00	1	1

PC1231	563	3/18/98	0.08	10.00	5.05	5.87	1	1
PC1231	578	4/2/98	2.00	51.00	16.00	9.00	1	1
PC1231	653	6/16/98	0.00	6.75	0.41	0.27	0	0
PC1231	698	7/31/98	0.22	10.72	1.35	0.15	0	0
PC1231	791	11/1/98	0.00	46.90	4.14	2.64	0	0
PC1232	583	4/7/98	0.00	71.00	8.00	3.00	0	0
PC1232	592	4/16/98	0.05	53.00	3.00	2.00	0	0
PC1232	597	4/21/98	1.00	57.00	6.00	3.00	0	0
PC1232	599	4/23/98	0.15	7.90	3.50	0.90	0	1
PC1232	699	8/1/98	0.00	49.70	5.28	1.92	0	0
PC1232	791	11/1/98	0.03	36.57	2.04	1.24	0	0
PC2131	156	2/4/97	0.00	15.50	2.10	0.50	0	1
PC2131	159	2/7/97	0.00	7.00	1.80	0.40	0	1
PC2131	178	2/26/97	0.05	6.70	2.30	0.40	0	0
PC2131	179	2/27/97	0.00	12.20	2.20	0.40	0	0
PC2131	184	3/4/97	0.00	47.97	1.51	0.40	0	1
PC2131	239	4/28/97	0.05	9.60	1.10	0.70	0	0
PC2131	241	4/30/97	0.10	8.10	1.00	0.70	1	0
PC2131	295	6/23/97	0.20	6.10	1.50	0.40	1	0
PC2131	386	9/22/97	1.70	21.00	1.40	3.70	1	0
PC2131	470	12/15/97	78.00	48.00	12.00	9.00	0	0
PC2131	535	2/18/98	0.50	27.00	7.40	7.00	0	0
PC2131	583	4/7/98	2.00	62.00	39.00	6.00	0	0
PC2131	597	4/21/98	2.00	64.00	38.00	4.00	0	0
PC2131	604	4/28/98	2.00	61.00	37.00	5.00	1	0
PC2131	611	5/5/98	0.01	24.00	6.00	1.40	1	0
PC2131	631	5/25/98	0.01	10.00	10.00	1.00	1	0
PC2131	640	6/3/98	0.20	26.00	1.00	4.00	1	0
PC2131	689	7/22/98	0.05	4.60	0.25	0.33	1	0
PC2131	768	10/9/98	0.05	4.20	0.30	0.20	1	0
PC2131	774	10/15/98	0.06	5.89	0.37	0.48	1	0
PC2131	791	11/1/98	0.34	17.55	4.66	5.60	0	0
PC3131	241	4/30/97	0.10	8.00	1.70	1.00	1	0
PC3131	295	6/23/97	0.70	35.00	10.00	7.00	1	0
PC3131	386	9/22/97	2.00	33.00	5.00	7.00	1	0
PC3131	450	11/25/97	3.13	20.00	4.00	2.00	1	0
PC3131	550	3/5/98	0.10	8.08	1.81	1.20	1	0
PC3131	651	6/14/98	0.71	39.20	9.80	7.70	1	0
PC3131	750	9/21/98	2.40	36.30	4.90	6.58	1	0
PC3131	791	11/1/98	3.47	21.40	4.08	1.80	1	0
PC3132	239	4/28/97	0.20	3.60	0.25	0.55	1	0
PC3132	295	6/23/97	0.30	12.20	0.90	2.20	1	1
PC3132	386	9/22/97	0.05	35.00	2.50	2.40	1	1
PC3132	450	11/25/97	0.00	81.00	8.00	6.50	0	0
PC3132	506	1/20/98	0.04	141.55	15.78	12.77	0	0
PC3132	566	3/21/98	0.23	4.32	0.25	0.59	1	0
PC3132	711	8/13/98	0.37	15.61	1.06	2.35	1	1
PC3132	791	11/1/98	0.06	39.90	3.25	2.61	1	1

PC3232	239	4/28/97	0.01	16.00	2.30	0.30	1	0
PC3232	295	6/23/97	1.00	48.00	9.00	4.00	1	0
PC3232	386	9/22/97	1.00	52.00	4.00	3.00	1	0
PC3232	535	2/18/98	0.00	91.00	26.00	8.00	0	0
PC3232	563	3/18/98	0.00	102.83	34.32	9.86	0	0
PC3232	591	4/15/98	0.00	280.00	10.00	15.00	0	0
PC3232	604	4/28/98	0.00	150.00	9.00	8.00	0	0
PC3232	639	6/2/98	5.00	73.00	6.00	6.00	0	0
PC3232	722	8/24/98	0.00	27.00	10.00	0.80	0	0
PC3232	723	8/25/98	0.00	27.62	10.14	0.73	0	0
PC3232	748	9/19/98	0.00	12.00	1.84	0.23	1	0
PC3232	783	10/24/98	0.73	30.72	5.85	3.20	1	0
PC3232	791	11/1/98	0.72	31.20	2.96	1.95	1	0

Table A.2.: *Inspection data for bearing 4*

Chapter 5

Closure

1 Overview

This research originated from a lack of means to integrate two established preventive maintenance strategies, statistical failure data analysis and vibration monitoring. Intuitively such an integration could have enormous benefits because of the well known successes that these two techniques have had on their own in the past. After studying typical vibration monitoring practices in the industry, it was decided that the most logical route to overcome the lack of integration between the two techniques, was by using regression models in renewal theory. The strategy followed was as follows:

(i) A thorough literature survey on existing regression models in renewal theory.
(ii) Identification of the most suitable model for the specific application.
(iii) A comprehensive, in-depth study on the chosen model with the emphasis on its practical use.
(iv) Practical evaluation of the theory for this application with a case study from the industry.

The literature study on regression models revealed that the Proportional Hazards Model (as in (1.1.)) was most suitable regression model to integrate failure time data and vibration information (in the form of covariates).

$$h(t,\overline{z(t)}) = h_0(t) \cdot \exp(\overline{\gamma} \cdot \overline{z(t)}) \qquad (1.1.)$$

The three most important reasons for this selection are: (1) The PHM has the widest theoretical foundation; (2) Parameter estimation for the PHM (and specifically the fully parametric PHM) is relatively tractable; and (3) The PHM has been used in similar reliability applications before.

An in-depth study into the PHM showed that the fully parametric Weibull PHM (see (1.2.)) was most suitable for this application and numerically the most attractive.

$$h(t,\overline{z(t)}) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \cdot \exp(\bar{\gamma} \cdot \overline{z(t)}) \qquad (1.2.)$$

Parameters were successfully estimated with four different optimization routines of which the modified Newton-Raphson method proved to be the most economical. Various goodness-of-fit tests for the PHM were found in the literature – all striving to test the PHM assumptions, in one way or another.

Only one decision making method for the PHM could be found, that of Makis and Jardine[13,14]. This method predicts the optimal hazard rate that will result in the minimum long term life cycle cost.

Suitable data to test the theory was found as SASOL, Secunda. The collected data was not perfect according to requirements but was still useful enough to produce, as far as is known, the first publishable case study on vibration covariates in the PHM. The case study revealed that close practical involvement is crucial in the modelling process and that statistical tests alone could easily lead to a misleading model. Only two of a total of twelve covariates were identified to be good predictors of failure by the PHM.

The decision model was proved to be valid and useful by several theoretical evaluations but also to be practical, by additional data. The case study showed that huge cost savings could have been brought about if the recommendations of the PHM were known and were acted upon.

2 Recommendations

From facts discovered and experience gained through this research project, three recommendations for future research/practices can be made:

(i) It is very surprising that so little research has been done on the PWP-model up to date. (Chapter 2, section 4.). This model has enormous potential because of its extreme versatility. It adds a third dimension to the PHM by taking previous replacements of the same item into account and it can handle both renewal and repairable systems situations. This model may have the ability to be a much better representation of practical situations than, for example, the PHM. It is recommended that the PWP-model's abilities are investigated further and it is predicted that a major contribution to reliability modelling would be the result.

(ii) The concept of minimum long term life cycle cost (Chapter 3, section 3.1.) is not accepted very well in the industry. People involved with reliability of items reject this concept as soon as an item lasts longer that the estimated time of minimum long term life cycle cost. Many reliability engineers have expressed their need for a decision making technique that predicts the 'exact' time to failure of an item rather than the minimum long term life cycle cost. Such a technique would certainly be possible if used with a model allowing for covariates, although no such technique was found in the literature during this research. A remaining service life estimation technique based on a regression model will be of wide practical interest if developed successfully.

(iii) Vibration data recording practices in the South African industry were found to be very incomplete and ineffective. (Chapter 4, section 2.2.). This is mainly due to a lack of commitment to vibration monitoring by managers but also due to unfriendly and disorganized CMMS's. It is believed that the successful case study described in this dissertation will improve these practices if published widely enough.

References

1. Bendell, A., *Proportional hazards modeling in reliability assessment*, Reliability Engineering, II, pp. 175-183, 1985.
2. Cox, D.R., *Regression models and life tables (with discussion)*, Journal of the Royal Statistical Society, B34, pp. 187-220, 1972.
3. Kalbfleisch, J.D. and Prentice, R.L., *The Statistical Analysis Of Failure Time Data*, Wiley, New York, 1980.
4. Jardine, A.K.S., Anderson, P.M. and Mann, D.S., *Application of the Weibull proportional hazards model to aircraft and marine engine failure data*, Quality and Reliability Engineering International, 5, pp. 77-82, 1987.
5. Jardine, A.K.S., Ralston, P., Reid, N. and Stafford, J., *Proportional hazards analysis of diesel engine failure data*. Quality and Reliability Engineering International, 5, pp. 207-216, 1989.
6. Kumar, D., Klefsjo, B., *Proportional hazards model: a review*. Reliability Engineering and System Safety. 44, pp. 177-188, 1993.
7. Dale, C., *Application of the proportional hazards model in the reliability field*, Proc. Fourth Nat. Reliab. Conf., pp. 5B/1/1-9,1983.
8. Ascher, H., *Regression analysis of repairable systems reliability in Electronic Systems Effectiveness and Life Cycle Costing* (J.K. Skwirzynski, Ed.), pp. 647-654, Springer-Verlag, 1983.
9. Davis, H.T., Campbell, K. and Schrader, R.M., *Improving the analysis of LWR component failure data*, Los Alamos Scientific Laboratory Report, LA-UR, pp. 80-92, 1980.
10. Jardine, A.K.S. and Anderson, M., *Use of concomitant variables for reliability estimations and setting component replacement policies*, Proc. 8th Advances Reliab. Techn. Symp., pp. B3/2/1 6,1984.
11. Prentice, R.L., Williams, B.J. and Peterson, A.V., *On the regression analysis of multivariate failure time data*, Biometrika, 68, 273-279, 1981.
12. Pijnenburg, M., *Additive hazards models in repairable systems reliability*, Reliability Engineering and System Safety, 31, 369-390.
13. Makis, V. and Jardine, A.K.S., *Computation of optimal policies in replacement models*, IMA Journal of of Mathematics Applied in Business Industry, Vol. 3, pp.169-175, 1992.
14. Makis, V. and Jardine, A.K.S., *Optimal replacement in the proportional hazards model*, INFOR, Vol. 30, pp. 172-183, 1991.
15. Ascher, H.E. and Feingold, H., *Repairable systems reliability*. Lecture notes in statistics - Volume 7. Marcel Dekker, New York, 1984.
16. Nelson, W., *Applied life data analysis*, Wiley series in probability and mathematical statistics, John Wiley & Sons, 1982.

17. Lewis, E.E., *Introduction to reliability engineering*, John Wiley & Sons, 1987.
18. Brinbaum, Z.W. and Saunders, S.C., *A new family of life distributions*, J. Of Appl. Prob., 6, 1969.
19. Bendell, A., Wightman, D.W. and Walker, E.V., *Applying proportional hazards modelling in reliability*, Reliability Engineering and System Safety, 34, pp. 35-53, 1991.
20. Snyman, J.A., *A new and dynamic method for unconstrained minimization*, Appl Math Modeling, Vol 6, 1982, pp 449-462.
21. Snyman, J.A., *An improved version of the original leap-frog dynamic method for unconstrained minimization LFOP1(b)*, Appl Math Modeling, Vol 7, 1983, pp. 216-218.
22. Coetzee, J.L., *Maintenance*, Maintenance Publishers, 1997.
23. Banjevic, D., Ennis, M., Braticevic, D., Makis, V. and Jardine, A.K.S., *Analysis of the optimal policy and the structure of the software for condition-based maintenance*, 1997, COMADEM '97.
24. Jardine, A.K.S., Banjevic, D. and Makis, V., *Optimal replacement policy and the structure of software for condition-based maintenance*, Journal of Quality in Maintenance Engineering, Vol. 3, No.2, 1997, pp. 109-119.
25. Kaplan, E.L. and Meier, P., *Nonparametric estimation from incomplete observations*, J. Amer. Stat. Soc, 53, pp. 457-481.
26. Hines, W.W. and Montgomery, D.C., *Probability and statistics in engineering and management science*, John Wiley & Sons, 1980.
27. Crowder, M.J., Kimber, A.C., Smith, R.L. and Sweeting, T.J., *Statistical analysis of reliability data*, Chapman and Hall, 1991.
28. Shyur, H. and Luxhoj, J.T., *Reliability curve fitting for aging helicopter components*, Reliability Engineering and System Safety, Vol. 48, pp. 229-234, 1995.
29. Rawicz, A.H. and Girling, D.R., *Neural network enhancement for a reliability expert-system*, Proceedings Annual Reliability and Maintainability Symposium, 1994, pp. 468-474.
30. Lakey, P.B., *Reliability modelling and design criteria for the backpropagation artificial neural network*, Proceedings Annual Reliability and Maintainability Symposium, 1993, pp. 114-119.
31. Pike, M.C.A., *A method of analysis of a certain class of experiments in carcinogenesis*, Biometrics, 22, pp. 142-161, 1966.
32. Leemis, L.W., *Reliability, probabilistic models and statistical methods*, Englewood Cliffs, NJ: Prentice Hall, 1995.
33. Newby, M., *Accelerated failure time models for reliability data analysis*, Reliability Engineering and System Safety, 20, pp. 187-197, 1988.
34. Solomon, P., *Effect of misspecification of regression models in the analysis of reliability data*, Biometrika, 71, pp. 291-298, 1984.
35. Lin, D.Y., Wei, L.J. and Ying, Z., *Accelerated failure time models for counting processes*, Biometrika, 85, V3, pp. 605-618, 1998.
36. Martorell, S., Sanchez, A. and Serradell, V., *Age-dependent reliability model considering effects of maintenance and working conditions*, Reliability Engineering and System Safety, 64, pp. 19-31, 1999.

37. Addison, J.T. and Portugal, P., *Some specification issues in unemployment duration analysis*, Labour economics, 5, pp. 53-66, 1998.
38. Shyur, H. and Luxhoj, J.T., *Non-parametric models for analysing failure time data of aging aircraft with multiple stresses*, 5th Industrial Engineering Research Conference Proceedings, pp. 453-458.
39. Younes, N. and Lachin, J., *Link-based models for survival data with interval and continuous time censoring*, Biometrics, 53, 1199-1211, 1997.
40. Williams, B.J., *Proportional intensity models for multiple event times*, Unpublished Ph.D. dissertation, University of Washington, 1980.
41. Ascher, H.E., *Regression analysis of repairable systems reliability*, Springer-Verlag, New York, 1982.
42. Dale, C.E., *Application of the proportional hazards model in the reliability field*, Proceedings of the Fourth National Reliability Conference, Reliability '83, United Kingdom, 1983.
43. Bennett, S., *Analysis of survival data by the proportional odds model*, Statistics in Medicine, Vol. 2, pp. 273-277, 1983.
44. Shen, X., *Proportional odds regression and sieve maximum likelihood estimation*, Biometrica, 85, 1, pp. 165-177, 1998.
45. David, H.A. and Moeschberger, M.L., *The theory of competing risks*, Griffin, London, 1978.
46. Aranda-Ordaz, F.J., *Transformations to additivity for binary data*, Unpublished PhD thesis, Imperial College, London, 1980.
47. Elandt-Johnson, R.C., *Some prior and posterior distributions in survival analysis: a critical insight on relationships derived from cross-sectional data*, J.R. Statis. Soc. (Ser. B), 42, pp. 96-106, 1980.
48. Proschan, F., *Theoretical explanation of observed decreasing failure rate*, Technometrics, 5, pp. 375-383, 1963.
49. Davis, D.J., *An analysis of some failure data*, J. Amer. Stat. Assoc., 47, pp. 113-150, 1952.
50. Jardine, A.K.S., and Anderson, M., *Use of concomitant variables for reliability estimation*, Maintenance Manage. Int., 5, pp 135-40.
51. Bendell, A., Walley, M., Wrightman, D.W. and Wood, L.M., *Proportional hazards modelling in reliability analysis – an application to brake discs on high speed trains*, Quality and Reliability Engineering International, 2, pp. 45-52, 1986.
52. Ansell, R.O. and Ansell, J., *Modelling the reliability of sodium sulphur cells, Reliability Engineering*, 17, pp. 127-137, 1987.
53. Lindqvist, B., Molnes, E. and Rausand, M., *Analysis of SCCV performance data*, Reliability Engineering and System Safety, 20, pp. 3-17, 1988.
54. Mazzuchi, T.A. and Soyer, R., *Assessment of machine tool reliability using a proportional hazards model*, Naval Res. Logistics, 36, pp. 765-777, 1989.
55. Leitao, A.L.F. and Newton, D.W., *Proportional hazards modelling of aircraft cargo door complaints*, Quality and Reliability Engineering International, 5, pp. 229-238, 1989.

56. Love, C.E. and Guo, R., *Using proportional hazards modelling in plant maintenance*, Qual. and Reliab. Eng. Int., 7, pp. 7-17, 1991.
57. Love, C.E. and Guo, R., *Application of Weibull proportional hazards modelling to bad-as-old failure data*, Qual. and Reliab. Eng. Int., 7, pp. 149-157, 1991.
58. Kumar, D., Klefsjö, B and Kumar, U., *Reliability analysis of power transmission cables of electric mine loaders using the proportional hazards model*, Reliab. Eng. & System Safety, 37, pp. 217-222, 1992.
59. Kumar, D. and Klefsjö, B., *Proportional hazards model – an application to power supply cables of electric mine loaders*, J. Reliab., Quality & Safety Eng. (in press).
60. Kumar, D., *Reliability analysis considering operating conditions in a mine*, Licentiate Thesis 1993: 17L, Luleå University of Technology, Sweden, 1993.
61. Booker, J., Campbell, K., Goldman, A.G., Johnson, A.G. and Bryson, M.C., *Application of Cox's proportional hazards model to light water reactor component failure data*, Report no. LA-8834-SR, Los Alamos Scientific Laboratory, Los Alamos, NM, 1981.
62. Gill, R.D., *Understanding Cox's regression model: a martingale approach*, Journal for the American Statistical Association, 79, pp. 441-447, 1984.
63. Cox, D.R., *Partial Likelihood*, Biometrika, 62(2), 269-276, 1975.
64. Fokianos, K., Kedem, B. and Short, D.A., *Partial Likelihood analysis of categorical time series models*, Technical research report: T.R. 95-98, Mathematics Department and Institute for Systems Research, University of Maryland, 1995.
65. Breslow, N.E., *Covariance analysis of censored survival data*, Biometrics, 30, pp. 89-99, 1974.
66. Breslow, N.E., *In discussion of the paper by Cox (1972)*, Journal of the Royal Statistical Society, B34, pp. 216-217, 1972.
67. Link, C.L., *Confidence intervals for the survival function using Cox's proportional hazards model with covariates*, Journal for the American Statistical Association, Biometrics, 40, pp. 601-610, 1984.
68. Effron, B., *The efficiency of Cox's likelihood function for censored data*, Journal of the American Statistical Association, 72, pp. 557-565, 1977.
69. Cooper, N.B. and Darch, G.D., *The use of concomitant variables and a preventive replacement study for the armored vehicle general purpose*, Unpublished report, Department of Engineering Management, RMC Kingston, 1984.
70. Thall, P.F. and Lachin, J.M., *Assessment of stratum covariate interactions in Cox's proportional hazards regression model*, Statistics in Medicine, 5, pp. 73-83, 1986.
71. Struther, C.A. and Kalbsfleich, J.D., *Misspecified proportional hazards models*, Biometrika, 73, pp. 363-369, 1986.
72. Brännäs, K. and Pohjanpalo, J., *Small sample properties of the likelihood estimator in proportional hazard regression models with omitted variables*, Biometrical Journal, 28, pp. 219-228, 1986.
73. Prentice, R.L., *Covariate measurement errors and parameters estimation in a failure time regression model*, Biometrika, 69, pp. 331-342, 1982.
74. Lagakos, S.W., *The loss of efficiency from misspecifying covariates in proportional hazards regression models*, Biometrika, 75, pp. 156-160, 1988.

75. Bryson, M.C. and Johnson, M.E., *The incidence of monotone likelihood in the Cox model*, Technometrics, 23, pp. 381-383, 1981.
76. Peduzzi, P.N., Hardy, R.J. and Halford, T.R., *A stepwise variable selection procedure for nonlinear regression*, Biometrics, 36, pp. 511-516, 1980.
77. Anderson, J.A. and Senthilselvan, A., *A two-step regression model for hazard functions*, Applied Statistics, 31, pp. 44-51, 1982.
78. Gore, S.M., Podcock, S.J. and Kerr, G.R., Regression models and non-proportional hazards in the analysis of breast cancer survival, *Applied Statistics*, 33, pp.176-195, 1984.
79. Pettitt, A.N. and Daud, I.B., *Investigating time-dependence in Cox's proportional hazards model*, Applied Statistics, 39, pp. 313-329, 1990.
80. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., *Numerical Recipes in C : The Art of Scientific Computing*, Cambridge Univ. Press, Second Edition, 1993.
81. Arjas, E., *A graphical method for assessing goodness of fit in Cox's proportional hazards model*, Journal for the American Statistical Association, 83, pp. 204-212, 1988.
82. Cox, D.R., *A note on the graphical analysis of survival data*, Biometrika, 66, pp. 188-190, 1979.
83. Kay, R., *Proportional hazards regression models and the analysis of censored survival data*, Applied Statistics, 26, pp. 227-237, 1977.
84. Schoenfield, D., *Partial residuals for the proportional hazards regression model*, Biometrika, 67, pp. 134-53, 1980.
85. Anderson, P.K., *Testing goodness of fit of Cox's regression and life model*, Biometrics, 38, pp. 67-77, 1982.
86. Moreau, T., O'Quigly, J. and Mesbah, M., *A global goodness-of-fit statistic for the proportional hazards model*, Applied Statistics, 34, pp. 212-218, 1985.
87. Kay, R., *Goodness of fit methods for the proportional hazards regression model: a review*, Rev. Epidémiol. Et Santé Publique, 32, pp. 185-198, 1984.
88. Michael, J.R., *The stabilized probability plot*, Biometrika , Vol. 70, pp. 11-17, 1983.
89. Aitkin, M. and Clayton, D., *The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data using GLIM*, Applied. Statistics, Vol. 29, pp. 156-163, 1980.
90. Lawless, J.F., *Statistical models and methods for lifetime data*, John Wiley, New York, 1982.
91. Holford, T.R., *Life tables with concomitant information*, Biometrics, 32, pp. 587-598, 1976.
92. O'Quigley, J. and Pessione, F., *Score tests for homogeneity of regression effects in the proportional hazards model*, Biometrics, 45, pp.135-144, 1989.
93. Schoenfield, D., *Chi-squared goodness-of-fit tests for the proportional hazards regression model*, Biometrika, 67, pp. 134-153, 1980.
94. McKeague, I.W. and Utikal, K.J., *Goodness-of-fit tests for additive hazards and proportional hazards models*, Scandinavian Journal of Statistics, 18, pp. 177-195.
95. Horowitz, J.L. and Neumann, G.R., *A generalized moments specification test for the proportional hazards model*, Journal for the American Statistical Association, 87, pp. 234-240, 1992.

96. Gill, R. and Schumacher, M., *A simple test of the proportional hazards assumption*, Biometrika, 74, pp. 289-300, 1987.

97. Koziol, J.A., *Goodness-of-fit tests for randomly censored data*, Biometrika, Vol. 67, pp. 693-696, 1980.

98. Azzalini, A., *Statistical Inference based on the likelihood*, Chapman and Hall, 1996.