Reliability Models for Linear Assets

by

William James McLauchlan Luff

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science

Department of Mechanical and Industrial Engineering
University of Toronto

© Copyright by William JM Luff 2012
Reliability Models for Linear Assets

William James McLauchlan Luff
Master of Applied Science

Department of Mechanical and Industrial Engineering
University of Toronto

2012

Abstract

Linear assets are among the largest and most important engineered systems; their reliability is of the utmost importance. This thesis presents an overview of the reliability estimation methods used for the various types of linear assets, both observation- and statistically-based. While observation-based reliability monitoring and estimation methods are necessarily particular to a certain type of asset, statistically-based methods developed for one type can potentially inform those used for another.

Therefore, this thesis looks to point out commonalities in the methods for the statistical evaluation of the reliability of various types of linear assets, develop and extend reliability models and methods with this knowledge, and suggest how maintenance strategies may be improved. To help illustrate and test the models described in this paper a case study was conducted with a utility operator; this thesis shows the modelling results from the study, and demonstrates the model’s use in a maintenance decision model.
Acknowledgements

I would like to express my thanks to my research supervisor, Prof. Andrew Jardine, as well as the other research staff at C-MORE, for their support and helpful advice as I completed this thesis. In particular, I would like to express appreciation to Dr. Dragan Banjevic for his helpful suggestions and feedback concerning my research, and to Dr. Elizabeth Thompson for her dedicated work editing this document.

I would also like to thank my family and friends for their continual encouragement and support as I conducted the work described here. They have been very kind and patient with me as I inched towards completion. Thanks especially to my parents and to Laura.
# Table of Contents

**Front Matter**

Abstract .......................................................................................................................... ii  
Acknowledgements ......................................................................................................... iii  
Table of Contents ........................................................................................................... iv  
List of Tables .................................................................................................................. vii  
List of Figures ................................................................................................................. ix  
List of Appendices ......................................................................................................... xi  

**Chapter 1: Introduction** ............................................................................................ 1  
1.1 Background & Motivation ....................................................................................... 2  
1.2 Objectives ............................................................................................................... 4  
1.3 Contributions ......................................................................................................... 4  
1.4 Overview ................................................................................................................. 5  
References ..................................................................................................................... 6  

**Chapter 2: Reliability Modelling and Theory** .......................................................... 7  
2.1 Reliability Defined ................................................................................................ 8  
    Stress-Strength Model and Static Reliability ............................................................. 10  
2.2 Modelling the Reliability of Systems .................................................................. 11  
2.3 Proportional Hazards Models .............................................................................. 13  
    Fundamental Assumptions ....................................................................................... 13  
    Weibull PHM ......................................................................................................... 14  
    Alternative Baseline Models ................................................................................... 17  
    Relaxing a Key Assumption: Time-Dependent Hazard Ratios ................................. 19  
2.4 The Role of Reliability Models in Maintenance Strategy and Decision-Making .... 22  
References ..................................................................................................................... 24
Chapter 3: A Review of Reliability Estimation Methods for Linear Assets

3.1 Introduction

3.2 Reliability Modelling of Specific Linear Asset Types
   Electrical and Communications Lines
   Water Supply Lines
   Energy Pipelines
   Reliability Modelling of Transport Lines

3.3 Common Features of Linear Asset Reliability Modelling
   Transmission-Distribution Distinction
   Frequency of Failures
   Frequently-Used Covariates
   Commonly-Used Statistical Models
   Areas for Improvement

References

Chapter 4: Modeling an Ideal Linear Asset

4.1 Introduction

4.2 A Linear Asset as a System of Simple Line Sections
   Modelling the Ageing Factor in the Failure Intensity of a Linear Asset
   On the Relationship between Asset Length and Failure Intensity
   Including Asset Information & Condition Measurements in a Reliability Model

4.3 A Linear Asset as a System of Three-State Line Sections

4.4 Development of Maintenance Strategies
   Selection of Segment/Component Types
   Determining Whether to Inspect a Linear Asset after it Experiences a Damaging Event

4.5 Summary and Potential Areas of Future Exploration

References

4.A Appendix
# Chapter 5: Case Study with Telephone Company

- **5.1 Introduction** ................................................................. 78
- **5.2 Project Motivation** .......................................................... 80
- **5.3 Current state of Reliability Modelling at TC** .......................... 82
- **5.4 Data Sets Used for Case Study** .......................................... 84
- **5.5 Initial Data Analysis** .......................................................... 87
  - Local Area Data Set Analyses .................................................. 87
  - Operating Area Data Set Analyses ........................................... 90
- **5.6 A New Model** ................................................................... 94
  - Basic Considerations .............................................................. 94
  - Identifying a New Model .......................................................... 100
  - New Operating Area Analysis ................................................ 103
    - Square Function Results ...................................................... 103
    - logExp Results ................................................................... 106
  - New Local Area Analysis ......................................................... 109
- **5.7 Discussion** ...................................................................... 111
- **5.8 Decision Model** ............................................................... 112
- **5.9 Conclusions and Future Work** .......................................... 114
- **References** ........................................................................... 116

# Chapter 6: Summary & Conclusions

- **6.1 Thesis Contributions** .......................................................... 120
  - Guidelines for Building Simple Reliability Models of Linear Assets ........................................ 121
- **6.2 Suggestions for Future Work** .............................................. 123
  - Suggestions for Reliability Engineering Theorists .......................................................... 123
  - Suggestions for TC’s Reliability Engineers ................................................................. 124
  - Suggestions for the EXAKT Software’s Developers ....................................................... 125
- **Bibliography** .......................................................................... 127
List of Tables

Table 1-1: Linear assets and their components in various industries........................................2

Table 4-1: Fractional deviation of the virtual age model failure hazard from the explicit failure hazard calculation following the first failure.......................................................... 54

Table 5-1: Failure event cause codes for Telephone Company datasets.................................84

Table 5-2: Summary statistics for data sets under study, for events of interest.......................85

Table 5-3: Variables included in each weather record/inspection; a checkmark indicates that the computed statistic for each element of the record was in the data set.................86

Table 5-4: Estimated parameters for Weibull baseline model on local area data (short histories excluded).................................................................................................................. 87

Table 5-5: Estimated parameters for constant baseline model on local area data (short histories excluded).................................................................................................................. 88

Table 5-6: Estimated parameters for Weibull baseline model on local area data (short histories included).................................................................................................................. 89

Table 5-7: Estimated parameters for constant baseline model on local area data (short histories included).................................................................................................................. 89

Table 5-8: Estimated parameters for Weibull baseline model on operating area data (short histories included)........................................................................................................... 91

Table 5-9: Estimated parameters for constant baseline model on operating area data (short histories included)........................................................................................................... 91

Table 5-10: Functions used to describe the time-dependent hazard ratio...............................97

Table 5-11: Hazard ratio for various functional forms of F(T)..................................................98

Table 5-12: Estimated parameters for model on operating area data with square function (a=350 days) as TDHR function (short histories included)..............................................104

Table 5-13: Estimated parameters for model on operating area data with square function (a=350 days) as TDHR function and Weibull function as baseline (short histories included). .................................................................................................................. 105
Table 5-14: Estimated parameters for model on operating area data with logExp function (a=275 days, b=2.5) as TDHR function as TDHR function (short histories included). .......... 106

Table 5-15: Estimated parameters for model on operating area data with logExp function (a=275 days, b=2.5) as TDHR function and Weibull function as baseline (short histories included). ................................................................................................................. 107

Table 5-16: Estimated parameters for model with new baseline on operating area data (short histories included). .............................................................................................................................................. 109
List of Figures

Figure 2-1: Reliability block diagram for a series system .................................................. 11
Figure 2-2: Reliability block diagram for a parallel system ............................................ 11
Figure 2-3: Example of an EXAKT control chart .............................................................. 16
Figure 2-4: Typical effect of preventive maintenance on intensity function ......................... 20
Figure 2-5: Key areas of maintenance decisions .............................................................. 22
Figure 4-1: Schematic diagram of an electrical distribution line (top), and a water distribution line (bottom) ........................................................................................................ 48
Figure 4-2: Reliability block diagram for an electrical distribution line (top), and a water distribution line (bottom) .......................................................... 49
Figure 4-3: (Left) Standard model of segment failure behaviour; (Right) Proposed three-state model of segment failure behaviour ......................................................... 63
Figure 4-4: Labelled three-state model of segment failure behaviour ............................... 64
Figure 4-5: Failure time probability density function ....................................................... 72
Figure 5-1: Residuals of PHM on local area data with Weibull Baseline (left); transformed residuals against expectation (right) ................................................................. 88
Figure 5-2: Residuals of PHM on operating area data with Weibull Baseline (left); transformed residuals against expectation (right) ......................................................... 92
Figure 5-3: (Left) Standard model of system failure behaviour; (Right) Proposed three-state model of line failure behaviour ................................................................. 95
Figure 5-4: OutputVars code used to calculate TDHR function ...................................... 99
Figure 5-5: K-S Statistic for various TDHR functions, calculated by EXAKT from TC operating centre data .......................................................... 100
Figure 5-6: P-Value for TDHR functions, calculated by EXAKT from TC operating centre data .......................................................... 101
Figure 5-7: Residuals of PHM on operating area data with square function (a=350 days) as TDHR function (left); transformed residuals against expectation (right).

Figure 5-8: Residuals of PHM on operating area data with square function (a=350 days) as TDHR function and Weibull function as baseline (left); transformed residuals against expectation (right).

Figure 5-9: Residuals of PHM on operating area data with logExp function (a=275 days, b=2.5) as TDHR function (left); transformed residuals against expectation (right).

Figure 5-10: Residuals of PHM on operating area data with logExp function (a=275 days, b=2.5) as TDHR function and Weibull function as baseline (left); transformed residuals against expectation (right).

Figure 5-11: Residuals of PHM on local area data with new model (left); transformed residuals against expectation (right).
List of Appendices

4.A Appendix ................................................................................................................................. 76
Chapter 1
Introduction

1.1 Background & Motivation ................................................................. 2
1.2 Objectives ....................................................................................... 4
1.3 Contributions .................................................................................. 4
1.4 Overview ......................................................................................... 5
References .......................................................................................... 6
1.1 Background & Motivation

Linear assets\(^1\) are fundamental to modern life: they satisfy the energy, communications, water, and mobility needs of people in Canada and throughout the world. Thus, the safe and reliable operation of these physical assets is of the utmost importance to managers, maintainers, and engineers. These systems extend over large distances in a linear fashion, connecting nodes of customers, suppliers, or destinations. Linear assets consist of a primary linear component – a conduit through which the items being carried pass – as well as the support structures required for the functioning of the asset.

Most of these systems are components in larger networks\(^2\) which exhibit a hierarchical structure in terms of length, capacity, and connectedness. Transmission lines are at the apex of this hierarchy: they are the most important in terms of individual effects on system behaviour, and they have the highest capacity and most highly-connected components, though they generally are the least numerous components in a network. At the bottom of this hierarchy are distribution lines: they are the most numerous parts of the network, and the individual effect of each on the network as a whole is limited. Examples of these systems and their components for various industries are seen in Table 1-1 below; the classification in this table is not exhaustive and is merely produced for guidance.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Goods Carried</th>
<th>Transmission</th>
<th>→</th>
<th>Distribution</th>
<th>Support Structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roadways</td>
<td>People and goods/Cars and Trucks</td>
<td>Expressways</td>
<td>Arterials</td>
<td>Local Streets</td>
<td>Bridges, signals, interchanges</td>
</tr>
<tr>
<td>Railways</td>
<td>People and goods/Trains/Locomotives</td>
<td>Mainlines</td>
<td>Spur lines</td>
<td>Last mile/home lines</td>
<td>Junctions, bridges, signals</td>
</tr>
<tr>
<td>Communications</td>
<td>Electrical Signals and Light Signals</td>
<td>Fibre trunk, satellite transceivers</td>
<td>Intermediate Class</td>
<td>MV and home service lines</td>
<td>Switching stations, routers</td>
</tr>
<tr>
<td>Electrical</td>
<td>Electric current</td>
<td>High-Voltage &amp; EHV Lines</td>
<td></td>
<td></td>
<td>Poles, Transformers Distribution Stations</td>
</tr>
<tr>
<td>Water</td>
<td>Water and Sewage</td>
<td>Aqueducts, large-diameter mains</td>
<td></td>
<td></td>
<td>Local mains</td>
</tr>
<tr>
<td>Natural Gas</td>
<td>Gas</td>
<td>Large Pipelines</td>
<td>Home service</td>
<td></td>
<td>Pumping Stations</td>
</tr>
<tr>
<td>Oil</td>
<td>Oil</td>
<td>Large Pipelines</td>
<td>Distribution Nodes</td>
<td></td>
<td>Pumping Stations</td>
</tr>
</tbody>
</table>

\(\text{Table 1-1: Linear assets and their components in various industries.}\)

\(^1\) Linear Assets include such assets as telecommunications and electrical lines, pipelines, roadways, and railways.

\(^2\) Hence the alternative appellation, “network assets”, for those in the class.
As with all physical structures, linear assets must be maintained, either through repair or replacement, or some combination thereof. Physical maintenance policies for these systems are normally driven by the failure behaviour of the linear component; this part performs the system’s primary function and is typically the most technically difficult and/or financially expensive part to access, repair, or replace. Much study has been done on inspection and replacement policies for these primary components as well as their degradation mechanisms (e.g. study of the wear of the pipe wall in pipelines [1], corrosion of the conductors for electrical systems [2], road surface wear [3], etc.). In addition, several optimal maintenance policies for support structures have been defined in previous studies [4][5][6].

Research into maintenance strategies for linear assets as whole systems, however, is more limited. This may be due to the difficulty in modelling complex multicomponent systems or networks of them, or perhaps because the costs and consequences of failure – and the resulting loss of service of an individual asset – are hard to quantify. Moreover, because they typically exist as part of a network, study of the performance and reliability of linear assets has typically focused on the performance of networks [7][8]. This has mainly been done to better plan for system-wide maintenance and capital needs [8], among other reasons, and partly because of insufficient data or analytic capacity. Furthermore, investigation into methods for estimating the reliability of various assets in the class has been conducted separately, and systematic evaluations of the reliability features of linear assets are few.

While observation-based reliability monitoring and estimation\(^3\) methods are particular to a certain type of linear asset, statistically-based methods developed for one type can potentially inform those used for another. Despite the common underlying reliability structure of all linear assets [9], statistical reliability estimation methods for the various types have largely been developed in isolation, with little effort made to apply lessons learned in one class to another.

\(^3\)This paper takes “reliability estimation” to mean a calculation of the risk of an asset experiencing a failure in a given time interval.
1.2 Objectives

A primary objective of this thesis is to develop general probabilistic models that can be used to estimate the reliability of linear assets. Doing so involves establishing which features are common to models previously used for the various types of linear assets. The models developed will be reviewed to determine their accordance with reality, and recommendations will be made on their use in reliability estimation projects and to support maintenance strategies and decision-making.

A secondary objective of this thesis is to conduct a case study to estimate the reliability of a linear asset using the results of the aforementioned model development. This will help determine if the new model has improved predictive power over the one(s) currently used. To this end, a project with a utility operator, Telephone Company (TC), is described, its results are shown and a maintenance decision model demonstrated. As TC’s distribution network was the subject of a prior thesis in this department [10], the case study was conducted with the aim of utilizing the EXAKT software tool [11] used in the prior thesis.

1.3 Contributions

The main contributions of this thesis are:

- the presentation and study of a reliability model for an idealised linear asset;
- the explanation of features commonly found in statistical reliability models of linear assets interpreted as features of this ideal model;
- an explanation of how this model can be used to make decisions about maintenance and reliability strategies.

The benefits of this work are demonstrated through a case study with Telephone Company. The case study makes use of the knowledge gained from the study of the idealised reliability model to better estimate the risk of failure of TC’s lines than can be done using its present model. Furthermore, the use of this model may reduce the amount of effort needed for TC to perform its reliability estimation. It has the added benefit of reducing the
amount of failure history data discarded in the analysis, thus reducing the amount of lost information. This case study also demonstrates a simple cost model that can be used by TC to inform its maintenance policies and to support its decision-makers.

1.4 Overview

To provide the reader with the requisite background to the work presented in this paper, Chapter 2 reviews reliability modelling and theory in general. This review emphasises the various methods to estimate the risk of an asset experiencing a failure in a given time interval. Expanding on this, Chapter 3 provides an overview of the reliability estimation methods used for the various types of linear assets, both those that are statistically-based and those using direct observation. The chapter concludes with a discussion of the common features of linear assets, and their related statistical reliability models.

The core of this study follows on from these review chapters. Chapter 4 introduces an ideal model of a linear asset that attempts to incorporate the features discussed in the previous chapter. This chapter studies this idealized linear asset, and using it brings forth and assesses statistical reliability models that may be appropriate for real linear assets. Chapter 4 concludes by showing some potential applications of these reliability models and explores the use of the ideal model in developing maintenance strategies. Chapter 5 presents a detailed case study conducted with Telephone Company relating to part of its local network. A previous case study with TC [10] is described briefly, as are problems that remained unsolved in that work and those that have been encountered with TC’s reliability modelling projects since then. This chapter includes an application of one of the models developed in Chapter 4, and compares the results of this new model to those from the current one used by TC. A benefit of the new model is demonstrated with a decision tool based on its results. Finally, Chapter 6 summarizes and discusses the results of the thesis and provides suggestions for future work.


References


Chapter 2
Reliability Modelling & Theory

2.1 Reliability Defined................................................................. 8
  Stress-Strength Model and Static Reliability................................ 10
2.2 Modelling the Reliability of Systems...................................... 11
2.3 Proportional Hazards Models.................................................. 13
  Fundamental Assumptions......................................................... 13
  Weibull PHM.............................................................................. 14
  Alternative Baseline Models....................................................... 17
  Relaxing a Key Assumption: Time-Dependent Hazard Ratios.......... 19
2.4 The Role of Reliability Models in Maintenance Strategy and Decision-Making.... 22
References .................................................................................. 24
Chapter 2: Reliability Modelling & Theory

2.1 Reliability Defined

Three interconnected concepts are fundamental to the management of physical assets: an asset’s reliability, its availability, and its maintainability [1].

Briefly, availability is the fraction of potential operating time when an asset is operating or available to be operated. It often has a basically arithmetical definition:

\[ A = \frac{\text{Time Operating} + \text{Time Available to Operate but Not Operating}}{\text{Time Operating} + \text{Time Available to Operate but Not Operating} + \text{Time Under Repair}}. \]

Alternatively, it can have a probabilistic interpretation [1].

Maintainability is a more qualitative term, relating to how easily the functions that an asset provides can be restored after it experiences a failure. Common statistical measures of maintainability are the mean time to repair an asset (MTTR) and the mean downtime (MDT) [2]. Often, these numbers are affected by factors extrinsic to the asset itself.

Finally, reliability is the probability that an asset will perform its expected or desired function over a certain time and with a certain set of operating conditions. The most common statistical measure of reliability is the mean time to failure (MTTF) for assets that are removed from service upon failure. For assets that are repaired upon failure, a common measure is the mean time between failures (MTBF). These measures are often imprecise [2] as they are average values; a better description of an asset’s reliability is given by its reliability function, which can incorporate knowledge about how its risk of failure changes as it ages.

The reliability function \( R(t) \) is expressed mathematically by first defining a continuous random variable \( T \) as the time to failure of an asset, where \( T \geq 0 \). Then

\[ R(t) = \Pr(T \geq t), \]

where we impose the conditions \( R(t = 0) = 1 \) (i.e. the asset is operating with certainty at time zero), and \( R(t) \geq 0 \) [1]. This definition is also closely linked to the survival (or
survivor) function used in the study of disease incidence and mortality in such fields as epidemiology [3] or the force of mortality in actuarial science [4], so there is much overlapping knowledge between these fields and the study of physical asset reliability.

If we take the failure of an asset to be an event of interest, then the reliability function of an asset is the complement to the cumulative distribution function of its probability of failure before time $T$. That is, if we define a failure function

$$F(t) = \Pr(T < t),$$

we can definitively write

$$R(t) = 1 - F(t).$$

The reliability function of an asset is fundamentally linked to another measure of interest to physical asset managers, the instantaneous rate of failure of an asset, $\lambda(t)$, as

$$\lambda(t) = \frac{-dR(t)/dt}{R(t)} = \frac{dF(t)/dt}{1 - F(t)} = \frac{f(t)}{R(t)},$$

where $f(t)$ is the probability density function of the asset’s probability of failure before time $T$, the time differential of $F(t)$. Sometimes $\lambda(t)$ is referred to as an asset’s “failure rate”; it is called the “failure hazard rate”\(^1\), if certain probabilistic conditions are met [5]. In the case of repairable systems, a similar measure is the “rate of occurrence of failures” (ROCOF) or, again depending on certain conditions being met, the failure intensity\(^4\), though for repairable systems the relationship between the reliability function and the failure intensity is much less straightforward [6].

\(^1\)Formally, the [simple] failure hazard rate of an asset is defined as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T < t + \Delta t | T \geq t]}{\Delta t},$$

i.e. it is conditioned on the asset not experiencing a failure up to time $t$ and is the rate at which an asset’s “first” failure occurs. Conversely, the [simple] failure intensity is defined as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[N(t + \Delta t) - N(t) \geq 1]}{\Delta t},$$

where $N(t)$ is the number of failures experienced up to time $t$.\(\)
Often the reliability function of an asset is not only a function of time, but also a function of the operating conditions under which the asset is functioning, and we can rewrite the failure rate as

\[ \lambda(t) = \lambda(t; \bar{z}(t)), \]

where \( \bar{z}(t) \) is a vector of all pertinent information about an asset at time \( t \) including, potentially, information about its prior failure history; this vector is often referred to as a “covariate vector”. If we integrate this function over time, we can recover the asset’s reliability function:

\[ R(t) = \exp \left[ - \int_{\tau=0}^{\tau=t} \lambda(\tau; \bar{z}(\tau)) d\tau \right]. \]

This thesis is concerned with finding efficient and useful methods for modelling the failure rate, \( \lambda(t; \bar{z}(t)) \), of linear assets, and, hence, methods for estimating their reliability.

**Stress-Strength Model and Static Reliability**

A frequently-encountered concept in asset management, related to the above definition of reliability, is the idea of static reliability. Static reliability is often expressed in the stress-strength model \([7]\), where the stochastic distribution of the variable \( X \) that measures the strength of an asset is compared to the distribution of the stress \( Y \), which may cause it to fail over some time period \( T \). If there is some knowledge of these distributions over an interval \([0, T]\), then the reliability of the asset over the interval \([0, T]\) is given as:

\[ R(t) = \Pr[X > Y], \ t \in [0, T]; \]

here, the reliability is constant, as we assume that the distributions are stationary in time.

---

2 To make this calculation several conditions are required, primarily that the future evolution of the components of \( \bar{z}(t) \) are known; if these conditions are not met or \( \bar{z}(t) \) is stochastic in nature a more difficult calculation may be used to similar effect.

3 For example, \( X \) may be the yield strength distribution of a circular pipe given an uncertain set of pipe parameters, whereas \( Y \) may be the expected distribution of pressures of the fluid carried within the pipe over a given time interval \( T \); the chance that the pipe will fail over each time interval \( T \) is \( \Pr[X > Y] \).
2.2 Modelling the Reliability of Systems

While it is generally straightforward to determine the reliability function of a single component or a simple system, most systems are made up of many components or subsystems [8]. A common method of determining the overall reliability of such a system is the reliability block diagram technique [9]. The two simplest types of systems in reliability block diagram form are so-called series systems and parallel systems; reliability block diagrams for these two are shown below in Figure 2-1 and Figure 2-2 respectively.

**Figure 2-1: Reliability block diagram for a series system.**

**Figure 2-2: Reliability block diagram for a parallel system.**
A series system, in reliability terms, is a system in which the failure of any of its subsystems results in the failure of the system as a whole. Mathematically, the relationship of the reliability function of a series system to that of its subcomponents is expressed as the product of all of the subcomponent reliability functions, \( \text{viz.:} \)

\[
R_{\text{Series}}(t) = \prod_{n} R_{n}(t),
\]

where the \( R_{n}(t) \) are independent of each other. A series system could be simply described as a system which requires all of its subcomponents in order to function. Conversely, in reliability terms, a parallel system is one in which all subsystems must fail before the system does. Similarly, a parallel reliability system’s failure distribution is expressed as the product of all of the subsystem failure functions (which, again, must all be independent of each other), \( \text{viz.:} \)

\[
F_{\text{Parallel}}(t) = \prod_{n} F_{n}(t).
\]

If we prefer to use only reliability functions,

\[
R_{\text{Parallel}}(t) = 1 - \prod_{n} \left(1 - R_{n}(t)\right),
\]

using the relationship expressed in (2-4).

Many systems, however, do not correspond to the above forms. These complicated systems require more advanced mathematical methods to evaluate the reliability of the system as a whole. Some methods used to evaluate the reliability of such systems are the following: the decomposition method, in which the system’s reliability block diagram is analysed in sections; methods based on Bayes’ Theorem \[8\]; techniques such as the minimal path method or the cut set method \[10\]. Interested readers are directed to Ebeling’s text \[1\] for an accessible review of such useful methods.
2.3 Proportional Hazards Models

The proportional hazards model (PHM) was initially developed by Cox [11][12] in 1972 to help model the impact of a leukemia treatment on times of remission. The model has great utility in incorporating condition measurements into a reliability function or estimate of risk of failure; thus, the model has found widespread application in the reliability field [13][14][15].

Fundamental Assumptions

Cox’s original model was based upon the work of Kaplan and Meier [16] on nonparametric survival functions, that is, direct estimation of the survival function of a population from historical life tables. Cox assumed that the total hazard rate $\lambda(t; \tilde{z})$ of an individual member of a population could be given as the product of the following: an unknown function, $\lambda_0(t)$, that gives the hazard of a member whose vector of pertinent information, $\tilde{z}$ is the zero vector$^4$; some general function of the information vector $\tilde{z}$ and a vector of unknown parameters, $\tilde{y}$, that has the same dimension as the information vector. In his classic paper Cox [11] assumed that the latter function was simply the exponential of the inner product of the information vector and the vector of unknown parameters. Therefore, the Cox PHM has the typical form:

$$\lambda(t; \tilde{z}) = \lambda_0(t) \cdot \exp[\tilde{z} \cdot \tilde{y}]$$ \hspace{1cm} (2-12)

Cox noted that the vector $\tilde{z}$ need not be constant in time [11] and showed [17] that his model is practicably estimated when the covariate vector contains time-varying components. However, it still fundamentally assumed that the vector $\tilde{y}$ contains only time-constant parameters, so (2-12) is often written as:

$$\lambda(t; \tilde{z}(t)) = \lambda_0(t) \cdot \exp[\tilde{z}(t) \cdot \tilde{y}]$$ \hspace{1cm} (2-13)

$^4$Hence its name, the “baseline” hazard function, as it is from a base $\tilde{z}$ of zero.
A method for estimating this function was developed by Cox [17], and modern methods in a reliability context are found, for example, in Kumar and Klefsjo [14] and Bendell et al. [18]. The components of the vector $\tilde{y}$ are related to a measure of the relative risk between two assets with differing values of $\tilde{z}$. If two members $i$ and $j$ of a population differ only in the value of the $k^{th}$ component of $\tilde{z}$, then the hazard ratio between the two members, i.e. the proportionality between their hazards, is given as:

$$HR_{ij} = \exp[y_k(z_j - z_i)].$$

This ratio is of great interest in many fields, especially when the covariates are categorical, as it can be used to determine how effective efforts to control the determinants of these parameters may be in reducing the risk of failure of an asset or the death of individuals [3].

**Weibull PHM**

The function $\lambda_0(t)$ is often specified in reliability engineering contexts; the Weibull function is frequently used, as suggested by Cox [11], and in this case, the baseline function has the form:

$$\lambda_0(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}.$$  \hspace{1cm} (2-15)

In the Weibull baseline, the parameter $\beta$ is called the shape; it determines whether the baseline hazard is increasing with time ($\beta > 1$), constant in time ($\beta = 1$, corresponding to the exponential reliability function), or decreasing with time ($\beta < 1$). The parameter $\eta$ is the scale, and it scales the function in time. Independent of Cox’s model, the Weibull function had been used in reliability engineering contexts to provide a measure of the effects of age on an asset’s risk of failure and to solve problems such as whether to replace an asset at a certain age [19]. The Weibull model has the convenient form for its reliability function of:

$$R(t) = \exp \left[-\left(\frac{t}{\eta}\right)^\beta\right].$$  \hspace{1cm} (2-16)
A PHM with a Weibull baseline is referred to as a Weibull PHM; it has found widespread usage in the analysis of failure histories of mechanical equipment, to estimate their reliability [20]. The full hazard function is often written as

\[
\lambda(t; \bar{z}(t)) = \frac{\beta}{\eta} \left( t \frac{1}{\eta} \right)^{\beta-1} \cdot \exp \left[ \sum_{i=1}^{n} \gamma_i \cdot z_i(t) \right],
\]

where \( \gamma_i \) and \( z_i(t) \) are the components of the parameter vector \( \bar{y} \) and the covariate vector \( \bar{z}(t) \) respectively; \( n \) is the dimension of the two vectors.

If the covariates related to an asset are known through time, it is possible to estimate its reliability function by using (2-7) and (2-17). Software was proposed [21] to do this, and a method for estimating the time evolution of the covariate vector was combined with maximum likelihood estimation of the PHM parameters to produce a software package capable of estimating a reliability function for a variety of assets [22]. The reliability function calculated by this software package, EXAKT [23], has the form:

\[
R(t, \bar{z}(\tau)) = \exp \left[ - \int_{\tau=0}^{\tau=t} \frac{\beta}{\eta} \left( \frac{\tau}{\eta} \right)^{\beta-1} \cdot \exp \left[ \sum_{i=1}^{n} \gamma_i \cdot z_i(\tau) \right] d\tau \right],
\]

and to meet the conditions as in (2-7) it calculates the \( z_i(\tau) \) probabilistically. For interested readers, the theory behind the EXAKT software was fully presented in Banjevic & Jardine [5] and Banjevic et al. [22].

The primary goodness-of-fit test used by EXAKT is the Kolmogorov-Smirnov (K-S) statistical test [24]; this test is useful for larger sets of failure histories, when the number of censored histories is small relative to the total, and so long as maintenance policies do not induce particular artefacts into the data set [23].

The software also permits graphical analysis of the residuals of the hazard model estimated from the failure data by EXAKT; the plots produced are those of the ordered residuals, the residuals against their expected values, transformed residuals, and variance-stabilized residuals [24].
EXAKT has been used by a number of organizations to optimize reliability engineering decisions, particularly repair/replace decisions, and to model the reliability of a variety of assets with a multitude of condition measurements [25]. The software produces a control chart, an example of which is shown below in Figure 2-3, which combines the reliability function in (2-17) and (2-18) with estimates of the cost of inspecting an asset, repairing it after an unplanned or planned failure, and the typical inspection interval to advise on whether to remove an asset from service, based upon its current age and the current value of the covariate vector. In the below chart, the value of \( \bar{z}(t) \circ \hat{y} \) is called the composite covariate, \( Z \).

\[ Z = 1.1374 \times \text{Nerf} + 1.6902 \times \text{Pitation} \]

*Figure 2-3: Example of an EXAKT control chart, taken from [23].*

For a description of the optimization criteria and algorithms used to produce the above chart, the reader is directed to the paper by Banjevic et al. [22].
Alternative Baseline Models

While the Weibull PHM has proved useful in the reliability field to estimate the risk of failure of assets and, hence, to optimize maintenance decisions, alternatives have been proposed to overcome certain shortcomings. One problem, in particular, is the assumption that a failed asset is renewed upon failure; that is, its age returns to zero after repair. This is not always an appropriate assumption.

A variety of models have been proposed to better reflect the reality of asset management. Examples include Kijima’s virtual age models [26][27], in which the age of an asset is decreased by the fractional “effectiveness” of a repair, or other reduction in intensity models [28] which arithmetically reduce failure intensity upon repair following failure or other maintenance actions. Interested readers are referred to Doyen and Gaudoin’s paper [28], or Percy’s chapter in [29], for a fuller discussion of alternative reliability models in a maintenance context. We will briefly discuss Kijima’s models here.

The basic assumption of Kijima’s models is that the failure intensity of an asset is dependent not on its calendar age (or some other appropriate age metric) but on its virtual age, which is calculated based on both the asset’s age and its maintenance and failure history. It is intended to be a model for systems that undergo general repair, that is, a repair that may be neither a renewal (i.e. the asset is returned to an age of zero) nor a minimal repair (i.e. the asset is returned to its age immediately before failure). Kijima’s paper presented two models: the type I model assumes that the nth repair of an asset can only affect the age accumulated since the (n-1)th repair, i.e. a repair can only reduce a proportion of the virtual age of the asset that has been accumulated since the last repair; the type II model assumes that the nth repair of an asset can affect damage incurred over the whole operating history of the asset, i.e. a repair can reduce a proportion of the total virtual age of the asset [30].

Mathematically we may express the Kijima models as follows, using the formalism described by Guo et al. [30]. We first assume that we know the ordered times of failure of an asset, \( t_1, t_2, \ldots, t_n \), and the times between each of these failures, the inter-arrival times \( x_i \),
where $x_i = t_i - t_{i-1}$, and $x_1 = t_1$. For illustrative purposes, only one repair type is assumed here, but the model can be generalised to systems with multiple types of repair; it is assumed that the repair effectiveness factor, $q$, is known, and that $0 < q < 1$ (the lower bound corresponds to a minimal repair and 1 to renewal), though the lower restriction is not necessary, as some repairs actively harm an asset’s risk of failure and can be interpreted as increasing virtual age \[31\]. The type I model defines the virtual age, $v_i$, at the start of the period during which $t_i < t < t_{i+1}$, in a recursive fashion as:

$$v_i = v_{i-1} + q \cdot x_i.$$  \hspace{1cm} (2-19)

If it is desired that the virtual age be expressed without recurrence relations, the definition is:

$$v_i = q \cdot t_i.$$  \hspace{1cm} (2-20)

The type II model defines the virtual age in recursive terms as:

$$v_i = q \cdot (v_{i-1} + x_i).$$  \hspace{1cm} (2-21)

In expanded (i.e. non-recursive) form, the definition is:

$$v_i = q \cdot (q^{n-1} \cdot x_1 + q^{n-2} \cdot x_2 + \cdots + x_i).$$  \hspace{1cm} (2-22)

Interestingly, in the type II model, the virtual age may oscillate about a constant value in the long run (i.e. it is possible that $v_i < v_{i-1}$) \[32\], so long as the modelled failure intensity is increasing in virtual age, while the type I model will monotonically increase in virtual age (i.e. $v_i \geq v_{i-1}$ for all $i$) as time goes on \[30\].
Relaxing a Key Assumption: Time-Dependent Hazard Ratios

As noted above, a fundamental assumption of the Cox PHM is that all of the hazard ratios in the model, or the set of covariate parameters \( \hat{\beta} \), are constant in time. This proportional hazards assumption may not hold in all cases [33]. For example, if a condition measurement \( z_i \) states whether an asset is in a dangerous location or not, the relative risk of damage to the asset caused by being in a poor location may decrease with time. This sort of problem regularly arises in epidemiological studies and has been noted in several papers, e.g. Abrahamowicz et al. [33] and Gray [34].

One potential solution to this problem is to relax the proportional hazards assumption and, for certain covariates in a population, to permit the hazard ratio to have some time-varying component [33] beyond that implied by the time-dependence of the covariates: the PHM then contains a time-dependent hazard ratio (TDHR) for some covariates\(^5\). This can be expressed in various ways, but perhaps the most easily understood is the following (this is similar to the presentation in [34]):

\[
\lambda(t; \tilde{z}(t), \tilde{z}_{TD}(t)) = \lambda_0(t) \cdot \exp \left[ \tilde{z}(t) \circ \tilde{\beta} + \sum_i f_i \left( z_{TD,i}(t) \right) \right],
\]

where \( \tilde{z}(t) \) and \( \tilde{\beta} \) are defined as earlier, but where the components of \( \tilde{z}_{TD}(t) \) are those covariates that have a time-dependent hazard ratio, and the \( f_i \left( z_{TD,i}(t) \right) \) describe the functional form of the hazard ratio's time-dependence for the \( i \)th covariate in the vector. An alternative presentation that perhaps makes the concept clearer is:

\[
\lambda(t; \tilde{z}(t), \tilde{T}) = \lambda_0(t) \cdot \exp[\tilde{z}(t) \circ \tilde{\beta}(\tilde{T})],
\]

where \( \tilde{T} \) is the vector of times that determine the shape of the hazard ratios through an asset's history (e.g. the set of failure times or the times of other important events in the asset's history that might effect the hazard ratio, up to time \( t \)).

\(^5\) This perhaps-confusing nomenclature comes from the fact that in epidemiology most covariates are binary, and thus the hazard ratio is typically expressed simply as \( \exp[\gamma_i] \) for the \( i \)th covariate. We more properly mean to discuss time-dependent covariate parameters, but to be consistent with the prior literature we will refer to the concept as that of time dependent hazard ratios.
Methods for estimating and interpreting proportional hazards models containing time-dependent hazard ratios have been developed by various groups in the epidemiological field and are described in works by Abrahamowicz and MacKenzie [35], Lehr and Schemper [36], Klein and Moeschberger [3], and others. Interested readers are directed to those works.

The incorporation of time-dependent hazard ratios into proportional hazards models or proportional intensity models of the reliability of mechanical equipment has not been described, to the author’s knowledge, in the reliability engineering literature as fully as in the epidemiology field. The modelling problems noted in the epidemiological studies, however, have been described [27] and at least one paper by Park et al. [37] did use TDHRs in modeling the failure of water mains. In addition solutions to these modelling problems, similar in effect to the use of TDHRs, have been proposed [38]. For example, Percy et al. graphically described [38] the effect of preventive maintenance as causing a reduction in failure intensity that returns to some underlying level some time after the maintenance action; this is shown in Figure 2-4.

Figure 2-4: Typical effect of preventive maintenance on intensity function, from Percy et al. [38].
Percy et al. modelled this effect by using a multiplicative function of the baseline hazard, a function that varied from some number less than one (that modelled the action’s effectiveness) and exponentially decayed to unity in the limit [38]. This method can easily be presented within a proportional intensities model that utilizes TDHRs to model maintenance actions, thus it would seem that the use of time dependent hazard ratios offer great promise in modeling the effect on risk of failure of many maintenance actions while permitting the continued use of the proportional hazards/intensities framework.
2.4 The Role of Reliability Models in Maintenance Strategy and Decision-Making

While the above models are of great mathematical and theoretical interest, models of assets’ risks of failure are invaluable in the field of evidence-based physical asset management [39]. Figure 2-5 shows the types of asset management decisions that can be optimized; all can be informed by knowledge of assets’ reliability.

![Optimizing Equipment Maintenance & Replacement Decisions Diagram](image)

*Figure 2-5: Key areas of maintenance decisions, from Jardine [39].*

Reliability models can be used to determine when to expect to replace or repair an asset, given the cost of experiencing a failure relative to the cost of replacing or repairing the asset before it does so. They can also be used to determine the optimal inspection time for an asset, or whether changes in operating conditions or methods of construction may produce a reduction in the risk of failure [39]. In addition, the PHM has been used to
estimate the differing spare parts requirements of mines with different operating conditions [40], among many other innovative applications.

Accurate reliability models are fundamental to many physical asset management methodologies, and the need for them is expressed in several official standards, including the British Publicly Available Specification 55 - Optimal management of physical assets (PAS55), in the performance, assessment and improvement stage [41] and the United States Air Force military standard on the Reliability Program Requirements for Space and Launch Vehicles [42].

Improvement in methods for evaluating the reliability of physical assets has regularly led to improved asset management practices [39]. Thus, the continued improvement of these models and methods is potentially of great benefit to asset managers, maintainers, and society as a whole. It is for this reason that many industries direct substantial efforts to modeling the reliability of their own assets (see e.g. Billinton and Allan [43] for the case of the electrical power industry or Antaki [44] or Muhlbauer [45] for the pipeline industry); a review of such efforts in industries that operate and maintain linear assets follows in Chapter 3.
References


# Chapter 3

## A Review of Reliability Estimation Methods for Linear Assets

3.1 Introduction......................................................................................................................... 28
3.2 Reliability Modelling of Specific Linear Asset Types ......................................................... 29
   Electrical and Communications Lines ................................................................................ 29
   Water Supply Lines .............................................................................................................. 32
   Energy (Oil & Gas) Lines ...................................................................................................... 35
   Reliability Modelling of Transport Lines ........................................................................... 38
3.3 Common Features of Linear Asset Reliability Modelling .................................................... 39
   Transmission-Distribution Distinction .................................................................................. 39
   Frequency of Failures ........................................................................................................ 39
   Frequently-Used Covariates ............................................................................................... 40
   Commonly-Used Statistical Models .................................................................................... 40
   Common Areas for Improvement ....................................................................................... 41
References ............................................................................................................................. 42
3.1 Introduction

In this chapter, we discuss reliability estimation techniques for the various types of linear assets and provide a brief overview of the methods based on direct observation for each type of linear asset. However, as the focus of this thesis is on the statistically-based evaluation of the reliability of linear assets, we will devote the bulk of the chapter to that topic. There is a wide body of literature on each type of linear asset, and we hope to provide a sufficient overview, rather than an exhaustive catalogue.

This chapter is organised as follows. First, we will outline the reliability estimation techniques used for each type of linear asset. Because of their common features and the frequently overlapping literature, we have grouped together similar types of linear assets: electrical distribution/transmission lines and telecommunications lines; oil and natural gas pipelines; water transmission and distribution lines, and wastewater lines; transportation lines. We will then outline common reliability features of linear assets, and compare the statistical models used to evaluate their reliability. Finally, we will discuss potential areas for improvement.
3.2 Reliability Modelling of Specific Linear Asset Types

Electrical and Communications Lines

The reliability of electrical power systems is a well-studied topic, due to the great importance of such systems to the public, and the immediately visible effects of their failure [1]. Many of the methods developed for these systems have been applied in the telecommunications industry [2] because of their assets’ similar nature¹. Because of the widespread nature of their networks, significantly more effort has been channelled into measuring the relative reliability performance of networks of electrical lines than individual lines [3][4], and on ensuring that networks are designed to provide service reliably under expected conditions [4]. This has resulted in significant research dedicated to estimating and predicting system-wide metrics, such as the system average interruption frequency index (SAIFI) [1], defined as

\[
SAIFI = \frac{\text{total number of customer interruptions}}{\text{total number of customers served}},
\]

and many other similar measures.

Inspection of communications and electrical power lines is often made simpler by the fact that many such lines are strung above ground, and hence easily inspected visually. Transmission systems are subject to more regular rigorous inspection, not only of the conductors but of the high-voltage insulators [5] and the support towers for the lines as well as other components [6]. In addition to visual monitoring and manual inspection, transmission systems are often inspected by specialised robots that may also perform maintenance tasks [7]. Interested readers are directed to the various proceedings of the IEEE International Conference on Transmission and Distribution Construction, Operation and Live-Line Maintenance (ESMO), which contain regular updates on state-of-the-art inspection tools and methods used for electrical transmission and distribution lines.

¹ For example, both carry electric current along wires, both have distribution systems that run into most homes, etc.
There has been relatively more done on statistically modelling the reliability of distribution lines as individual assets than on transmission lines, in part due to the large number of distribution lines that must be handled. In most texts, such as Billinton and Allan's classic work [1] or more recent ones such as Chowdhury and Koval [4], distribution lines are assumed to be simple series systems with constant failure rates. In these two particular texts, each distribution line's failure rate is expressed as the sum of the failure rates of any components in the line, such as switches, plus a rate proportional to the length of the line. Similarly, in the IEEE Reliability Test System [8], all components are assumed to have constant annual failure rates. Billinton and Allan offer some extended models for common-cause failures, or for determining how many customers on a line are affected by a failure [1], but for the most part, models of distribution systems assume a constant average failure rate over the lives of the assets.

To model the effects of weather on distribution system reliability, Billinton and Allan [1] present a two-state weather model with "normal" weather and "adverse" weather states. They use this to model the effect of weather on system-wide measures, e.g. to account for a system's higher-than-average failure rate during years of adverse weather, and also to account for the problem of weather-induced failure bunches, i.e. failures occurring in rapid succession due to damage caused by adverse weather.

Some studies have used more advanced models of distribution line failure. Tang [9] presents a reliability model of distribution line failures for a network operator, Telephone Company, that uses processed weather covariates in a Weibull PHM of the failures of lines on the network. However, this study is limited by the omission of any physical information about the lines in its reliability model, which if added might significantly improve the accuracy of its forecasts. In addition, it handles the bunching problem by removing those failures from the set of data analysed instead of modelling them. Kumar et al. [10] use a PHM to model the effects of operating conditions on the failure of electrical distribution cables in a fleet of trucks. Additionally, there have been studies into the reliability of individual components of distribution systems, such as a Weibull model of the failure of wood support poles [11] and rights-of-way [12].
As the cost of a failure of an individual transmission line is extremely high, inspection is done regularly on these systems to ensure that they have high reliability. Thus, there is a small pool of failure histories and correspondingly few statistical analyses of electrical and telecommunications line failure and reliability. In one of the few notable examples, Baxter et al. [13], use a PHM model to evaluate the reliability of members of a population of transmission lines in Wales and southwest England, and include such covariates as time (age), line length, weather, etc.. This group has found that a binary weather covariate ("normal" and "adverse") can explain failure bunching. However, their model artificially defines a period of adverse weather to be a period in which bunching is present, and it is therefore difficult for the model to provide predictive power.

Other efforts have been made to model the effects of weather on electrical transmission line reliability. For instance, Billinton and Acharya have proposed and evaluated a multi-state weather model for system reliability [14][15], and use it to predict long-run reliability indices for the system. However, their model does not account for system ageing.

There have been several examinations of the ageing effect in transmission line components, but most are physically-based and seem directed towards efforts at producing more reliable components.
Water Supply Lines

In the analysis of the reliability of a water supply system, it is common to divide the system into distribution and transmission sections [16]. Non-destructive internal inspection of transmission lines is often done; such inspections are not performed as frequently for distribution lines, as it is relatively expensive to do so given the length of the distribution and its inaccessibility [16]. For transmission lines, visual inspection either by technicians or sensor equipment is often practical. Methods of inspection common to both transmission and distribution include acoustic leak detection and ultrasonic measurement. For an exhaustive list of such methods, interested readers are directed to Misiunas [16].

The statistical analysis of the reliability of water supply assets is somewhat complicated by the often-benign nature of their most common failures: a small leak of a water distribution pipe will likely have few damaging effects beyond the loss of the resource and the energy expended in delivering the water [17]. In some cases, the number of breaks that a section of water distribution pipe experiences is used as a measure of its condition; this is not practical for transmission lines, as noted in [18]: “a critical water asset like a trunk transmission line cannot rely on failure history as a proxy for condition. This vital asset should be managed in a proactive manner that prevents failures from occurring because of the dire consequences associated with failure.”.

However, in water distribution systems, statistical analysis is the most feasible option for evaluating reliability. As noted in Rajani and Kleiner: “It appears that currently, only large water mains with costly consequence of failure may justify the accumulation of data that are required for physically based model application. The statistical models seem to be an economically viable approach for the smaller distribution water mains.”[19]. Thus, statistically-based analyses are widely used in predicting the number of pipe breaks in networks (i.e. the failure intensity) [16], and a number of reports outline the use of such methods in the water distribution field. Reviews of the statistical models used in modeling the reliability of water distribution networks may be found in Osman and Bainbridge [18], Rajani and Kleiner [19], or, for more detail, Røstum [20].
Osman and Bainbridge [18] present two alternative models in their case study. The first is based on the transition state-life regression model, as they call it, and involves fitting Weibull hazard models to each of the first, second, third, etc., times to failure of similar populations of pipelines. (The lines are similar in terms of pipe diameter, section length, pipe material, etc.) The second model presented in [18] is the multivariate exponential model which predicts the number of failures in a given time period with the following formula:

\[ N(\tilde{z}_t) = N(\tilde{z}_{t_0}) \cdot \exp(\tilde{a} \cdot \tilde{z}_t), \]

where \( N(\tilde{z}_t) \) is the number of breaks in a pipe in time period \( t \) with covariates \( \tilde{z}_t \), \( \tilde{a} \) is the vector of covariate parameters, and \( N(\tilde{z}_{t_0}) \) is the number of failures that would be expected with all covariates at their reference values. It should be observed that this model has the same predictive result as a proportional hazards model with a constant baseline in each time period. In their model, Osman and Bainbridge segment their pipelines into sections of approximately equal length to avoid introducing an extra proportionality constant; they also estimate separate models for different pipe materials.

The multivariate exponential model used in [18], equation (3-2) was developed by researchers at NRC-IRC and resulted in a software application called D-WARP (Distribution Water Mains Renewal Planner) released for use by water infrastructure managers [21]. After using the tool to model the effects of a variety of parameters on failure rate, NRC-IRC researchers found it to be well correlated with length (most importantly, for long pipes; for shorter ones it was not as significantly correlated), pipe failure history (under certain conditions), pipe material, dynamic covariates such as those related to climate [22]. They found that in some cases a pipe’s location in certain sectors of the distribution network can affect its failure rate, e.g. if it is in a sector that experiences more failures than most other sectors. An extension of the D-WARP project resulted in another software package, I-WARP (Individual Water Main Renewal Planner), based on the non-homogeneous Poisson process (NHPP) instead of the multivariate exponential model [23]. This software is useful in case studies that provide similar results to those established by D-WARP [24].
Røstum [20] uses a NHPP with covariates to predict failures for networks of pipes as well as individual pipes, comparing the results to a modified Weibull proportional hazards model. The selected NHPP has an implicit assumption of minimal repair in the pipes. The covariates significantly related to failure include the length of the pipe (its logarithm), the diameter of the pipe, and in certain cases, the soil type in which the pipe is laid; for the Weibull model, covariates include the number of previous failures of the pipe.

In addition to the case studies mentioned above, several water distribution systems have been analyzed using a standard proportional hazards model. Park et al. [25] use a non-parametric PHM to model pipe failures in a distribution system to estimate the economic residual life of the pipes. Their significant covariates include pipe material, diameter, length, the type of water carried (i.e. treated or untreated) and the surface conditions above the pipe. The same research group has used a PHM to estimate the economic replacement time for pipelines [26].

Other uses of the PHM have been reported in Le Gat and Eisenbeis [27], Le Gauffre et al. [28], and others. Standard Weibull models are used instead of a PHM by Moteleb [29] for different material types. Further alternative statistical models such as the accelerated lifetime model and the time linear model have been used successfully elsewhere; some examples can be found in Lei & Saegrov [30] or Yamijala et al. [31].
Energy Pipelines

The failure of an oil or natural gas transmission pipeline is nearly always a newsworthy event. From time to time we hear of the catastrophic results of leaks in a natural gas distribution pipeline, a testament to the adverse consequences of such failures and also to their relative rarity. As with other types of linear assets, in the analysis of the reliability of oil or natural gas pipelines distribution and transmission systems are treated in different ways.

Because of the immense economic and social disruption caused by their failure, oil and natural gas transmission lines are subjected to rigorous inspection programs and other methods of integrity assurance [32]. Inspection of energy transmission pipelines is a technically advanced and highly developed industry, and is done by devices called “pigs” with sensors using a wide variety of physical detection methods. Interested readers are directed to such works as Kishawy and Gabbar [33], Mohitpour et al. [32] or Tiratsoo [34] for descriptions of pipeline pigging methods and other methods of inspecting them. Statistically-based methods of evaluating the reliability of energy transmission pipelines are less well-described, likely as a result of the effectiveness of inspection.

Energy distribution pipelines, however, are less easily inspected, and so—as is the case with water and electricity distribution lines—are more likely to be the subject of statistical failure analysis, although system-wide averages are often used to determine the relative risk-of-failure of particular segments of pipe [35]. (The works described herein discuss natural gas distribution pipelines, as oil distribution pipelines are relatively rare.) The safe operation of the distribution network is partly achieved by rapid detection of leaks [36], which is facilitated by the inclusion of a strong-smelling chemical in the gas to permit rapid reporting by persons in the area of a leak (i.e. a failure) [33].

---

2 The devices are called “pigs” due to the sound they make as they pass through the pipelines; the act of inspecting a pipeline with such a device is hence called “pigging”.

A common method of reliability estimation for pipelines involves calculating risk scores for each segment\(^3\) of the network [35] and converting the score into an estimated failure rate. Such risk scores can be based on the pipe length, the materials that make up the pipe, pressure within the pipe, the pipe's installation techniques, etc. [35]. In a variation on this method, Cooke et al. [37] use expert judgement to estimate the risk-of-failure for several failure modes, in a value of per km of line per year. Han and Weng [38] compare this method to more quantitative methods for estimating the risk-of-failure, including methods that combine expert judgement with historical data, to establish probability of failure estimates for segments of pipelines. They use the risk-of-failure estimate as an input into a relative ranking of the risk to society and the pipeline operator of each segment of the distribution system.

Methods based on more rigorous statistical analysis are rare. In a paper devoted to repairable systems, Ruggeri [39] describes several attempts to model the number of failures in an urban gas distribution network. In one case study on the city's network, where repairs were made to the lines upon failure, he uses an NHPP in cases where ageing is a factor in determining the failure rate; where ageing was not a factor, the HPP was used. In another case, a superposition of NHPPs is used to estimate the number of failures in a network. In all models, Ruggeri uses length as a linear proportionality constant for different sections of pipe; in some, he uses covariates to determine the effect of certain pipe parameters, the most influential being pipe diameter, depth of laying, and laying location.

In a conference presentation, Zuashkiani and Chiotti [40][41] discuss the failure histories of part of the distribution network of a utility operator and introduce a maintenance optimization program derived from the analysis. Their reliability analysis incorporates expert knowledge and judgement to build their model. The reliability analysis uses a Weibull PHM to model the failure histories and finds a variety of important covariates. Significant static covariates include the coating type of the pipelines, the soil in which they

\(^3\) Many papers note that segmenting a line should be sensitive to the conditions in which the line exists, and that it is often convenient to section a line into approximately-equal-length segments.
are buried, whether corrosion protection is present, whether there is any stray current protection, a function of the number of customers served on the line (i.e. number of service branches), and the length of the line. In addition, they include a function that provides information about the line's leak history.

Soszynska reports [42] a novel model to estimate the failure of a port's oil distribution pipelines by aggregating the failure rates of its component sections, themselves estimated by aggregating the failure rates of their segments. It is modelled as having multiple states reflecting the different operating conditions under which the segments might be functioning.
Reliability Modelling of Transport Lines

Transportation lines are somewhat more difficult to include in this study. They rarely exhibit failures and often operate at various levels of quality-of-service. The transportation asset closest in reliability terms to the assets discussed earlier are railway lines. Tracks are frequently inspected to ensure reliability and are highly reliable; they may have reliability features common to transmission systems. Rail lines must maintain particular track geometries to function correctly, and a misalignment of track is often considered a failure; more serious failures include derailments [43]. In addition, electrified track has catenaries to deliver motive power, and these experience failures in a manner similar to other electric power lines.

Vatn describes [43] maintenance optimization in the rail industry and provides a static reliability model for predicting derailments on sections of railway lines. The model is multiplicative, with a baseline failure rate in failures per kilometre per year, and multiplicative factors that are scaled exponentially with risk influencing factors (RIFs) as judged by the modeller; these factors include number of failures/cracks found previously in the rails, rail quality, track gradient, trackbed quality, and track geometry.

Zhao et al. [44] present a model of the reliability of a railway line’s sleepers, where failure is only considered to have happened when a certain number of individual sleepers fail in a set of sleepers; the sleeper reliability is modelled with a Weibull model. Centrone [44] uses a Weibull proportional hazards model to provide input into decisions about track geometry maintenance.
3.3 Common Features of Linear Asset Reliability Modelling

Transmission-Distribution Distinction

The most obvious common feature in the analysis of the reliability of the various types of linear assets is that all industries using such assets treat distribution and transmission assets in different ways. Transmission assets have extremely high failure costs, and to ensure that they occur as infrequently as possible, regular inspection is done on the lines with follow-up maintenance if required. The main reliability model used for transmission lines, it could be argued, is the stress-strength model; the inspections are used to ensure that the transmission line is operating with a sufficient safety factor, with statistically-based methods used far less frequently.

For distribution assets, however, these inspection techniques are either too expensive or not practical. For the most part, statistically-based reliability models are used to estimate the reliability of these assets, with inspections being used much less frequently. Managers often make use of measures of groups of distribution assets rather than individual lines.

Frequency of Failures

The frequency of failures on a type of line often determines the type of model used. When modelling lines or networks with very frequent failures, a common goal is to estimate the expected number of failures in a given time period, or a related metric for a section of the network, instead of a reliability function or estimate of the next time to failure (e.g. an MTBF estimate) as is more often the case with types of lines that have higher reliability.
Chapter 3: A Review of Reliability Estimation Methods for Linear Assets

**Frequently-Used Covariates**

Other commonalities are the covariates to include when building a reliability model. For example, most models have failure rates proportional (or nearly proportional) to the length of the asset, or some combination of length with the number of certain other components on the line (e.g. number of services attached).

Additionally, factors such as a line’s material of construction or its right-of-way type (e.g. soil type, whether it is underground or above ground, etc.) often enter into a reliability model. The use of internal and external dynamic measurements, such as weather or system pressure, is common to many of the models described.

**Commonly-Used Statistical Models**

The proportional hazards model and a related simplification, the multivariate exponential model, are popular choices for modelling the reliability of linear assets, as for modelling the reliability of other assets. Models similar to the PHM, such as the NHPP, are used in some cases to overcome the limitations of using a hazard model—with its assumption of renewal upon failure—for a complex system like a linear asset.

For assets that experience many failures, i.e. certain types of distribution lines, ageing is frequently not modelled as a factor in reliability models. In such cases, the distribution lines may not be analysed individually, but as part of a section of the network; in addition, the desired output of the reliability model will be the average frequency of failures in a given time period⁴, or some similar measure (e.g. SAIFI) that may affect the organisation’s customers or its maintenance processes.

⁴ Or equivalently, the expected number of failures in a given time period.
Areas for Improvement

There are no perfect reliability models, to the author’s knowledge, to be found in the literature. The above papers offer no highly effective method of incorporating inspection results or other maintenance actions into the reliability model. However, this is a shortcoming of reliability models in general [43].

The lack of statistical models for transmission assets is troubling, as we may derive valuable information from the analysis of their failures. It may be worthwhile for certain industries to pool their transmission asset failure history data to determine if some knowledge can be gleaned from their analysis.

The handling of failure bunching, or failures that occur in rapid succession, is similarly problematic. Analyses thus far handle them in an arbitrary manner or in a way that offers little predictive value. It may be worthwhile exploring alternative methods of incorporating them into reliability models.
Chapter 3: A Review of Reliability Estimation Methods for Linear Assets

References


Chapter 4
Modelling an Ideal Linear Asset

4.1 Introduction

4.2 A Linear Asset as a System of Simple Line Sections
Modelling the Ageing Factor in the Failure Intensity of a Linear Asset
On the Relationship between Asset Length and Failure Intensity
Including Asset Information & Condition Measurements in a Reliability Model

4.3 A Linear Asset as a System of Three-State Line Sections

4.4 Development of Maintenance Strategies
Selection of Segment/Component Types
Determining Whether to Inspect a Linear Asset after it Experiences a Damaging Event

4.5 Summary and Potential Areas of Future Exploration

References

4.A Appendix


4.1 Introduction

We begin our discussion of the reliability modeling of linear assets by considering certain basic structures common to the asset class. First, such assets contain many individual segments of the linear component—e.g. pipe sections, strings of cable, rails etc. The failure of any of these segments effectively results in the failure of the line on which it resides. Second, in local terms at least, these component segments will all experience much the same operating and weather or other external conditions as adjacent or nearby ones. That is, the flow rate of fluid or the level of electrical current running through nearby segments of the linear component on a subsection of a line will not differ substantially under normal operating conditions. Third, for much the same reason, nearby segments will often experience maintenance and inspection activities at approximately the same time [1]. For example, right of way clearing and other maintenance will typically occur sequentially down the length of an asset, not one section at a time.

In the following sections, we discuss several statistically-based reliability models that may be useful in estimating the failure intensity of an idealized linear asset\(^1\). These models are derived with the assumption that the asset may be divided into segments whose hazard rates may be individually described as some function of age\(^2\) and internal and/or external system conditions. The ideal division of a linear asset into different sections or groupings of segments will be discussed in the course of this chapter, but for purposes of exposition we will model as a single asset a line that may have multiple sections in an ideal analysis. For examples of the models, we use variations on the proportional intensity model, for reasons discussed previously in 2.3, and as it is common to use a Weibull baseline hazard/intensity function in reliability modeling [2], it will be used here.

\(^1\) As it is common to estimate failure intensity models for each failure mode or class of failure modes that an asset may experience, it is assumed that this would be done by any prudent reliability modeller and that we are more properly attempting to estimate the failure intensity for a particular mode.

\(^2\) Age, as mentioned previously, may be a function of calendar time or some other measure of the wear that an asset has experienced, e.g. cumulative flow through a pipe.
While it would be more natural and likely more accurate to track and estimate models for each of these segments individually—or even to estimate models for all components in a linear asset—it is often not practical to do so. In the following theoretical exercise, we presume for the sake of argument that each of these segments may be individually described in reliability terms; based on this, we then suppose that in the aggregate, a system of these segments may be accurately described in a certain way. Thus, we hope to make general statements about the reliability behaviour of linear assets based upon their common underlying structural reliability features.

In what follows, we explore how this initial ideal model may be extended to provide greater explanatory power, and how it might be used to better model the reliability of linear assets. We discuss how maintenance strategies may be influenced or supported by the use of improved reliability models and develop some optimization models that may be of use in the management of linear assets. The chapter concludes with suggestions for areas of future exploration in designing improved models of linear asset reliability.
4.2 A Linear Asset as a System of Simple Line Sections

A simple model of a linear asset is as a series system of the asset’s separate line segments, each with failure hazard function $\lambda_i(t_i; \mathbf{z}_i(t))$, where $t_i$ is the time since segment $i$ last experienced a failure (one might write it as $t_i(t)$, to be explicit), and $\mathbf{z}_i(t)$ is the covariate vector of segment $i$, both at time $t_i$. (As noted in 2.2, the covariate vector may include information about the asset’s history, its maintenance policies, and external factors such as weather, operational factors, and so on.) Functionally, a linear asset, such as an electrical line or pipeline, might appear as in Figure 4-1 below, and could be divided into segments as indicated. The reliability schematic for the line would appear as in Figure 4-2 below and is simply a series system.

![Diagram of a linear asset as a system of simple line sections](image)

Figure 4-1: Schematic diagram of an electrical distribution line (top), and a water distribution line (bottom). Each line is sectioned into segments based on the linear component.
Figure 4-2: Reliability block diagram for an electrical distribution line (top), and a water distribution line (bottom). The segments correspond to the ones in Figure 4-1. The reliability for the subcomponents for section \( i \) are indexed as \( R_{\text{Component Initials}(i)} \).

For such assets, the reliability of the system as a whole is simply the product of the reliability of the segments of the asset, which are each the product of the reliability of their subcomponents (assuming that the failures are independent),

\[
R_{\text{Asset}}(t) = \prod_{i \in \text{Asset}} \left[ \prod_{j \in C_i} R_{i,j}(t) \right],
\]

where \( \text{Asset} \) is the set of segments that compose the linear asset under consideration, and \( C_i \) is the set of components that make up segment \( i \); \( R_{i,j}(t) \) is the reliability of component \( j \) of segment \( i \). The relationship may equivalently be expressed thusly: the failure rate of the system as a whole is the sum of the failure rates of the component sections:

\[
\lambda_{\text{Asset}}(t) = \sum_{i \in \text{Asset}} \left[ \sum_{j \in C_i} \lambda_{i,j}(t) \right].
\]

For the diagrammatic examples above, the failure rate would be

\[
\lambda_{\text{Electric,Line}}(t) = [\lambda_{L,1}(t) + \lambda_{SP,1}(t) + \lambda_{I,1}(t)] + [\lambda_{L,2}(t) + \lambda_{SP,2}(t) + \lambda_{I,2}(t)]
+ [\lambda_{L,3}(t) + \lambda_{SP,3}(t) + \lambda_{I,3}(t)] + [...],
\]

in the case of the electrical distribution line, and in the case of the water distribution line, it would be:

\[
\lambda_{\text{Pipeline}}(t) = [\lambda_{PS,1}(t) + \lambda_{BS,1}(t) + \lambda_{W,1}(t)] + [\lambda_{PS,2}(t) + \lambda_{BS,2}(t) + \lambda_{W(1),2}(t) + \lambda_{W(2),2}(t)]
+ [\lambda_{PS,3}(t) + \lambda_{BS,3}(t) + \lambda_{W,3}(t)] + [...].
\]
Modelling the Ageing Factor in the Failure Intensity of a Linear Asset

If each of an asset’s segments or subcomponents is renewed upon its failure, then the observed failure intensity of the system as a whole is that of a set of pooled renewal processes [3]. Perhaps the most useful property of this type of process results from the elementary renewal theorem, which states that in the long run, the average event intensity—in this case the average failure intensity or the rate of occurrence of failures—will be stationary and equal to the sum of the reciprocals of the average interarrival time of each process in the pool [4]. That is,

\[
\lambda_{\text{Asset}}(t \to \infty) = \sum_{i \in \text{Asset}} \left[ \sum_{j \in C_i} \frac{1}{\text{MTBF}_{i,j}} \right],
\]

using the fact that the average interarrival time of failures for an item that is renewed upon failure is the mean time between its failures. (In (4-3) the indices are as in (4-1).) Based on this, we can make our first statement about the failure behaviour of linear assets.

**Statement 1—In the long run, a linear asset which has its components renewed upon failure will experience a failure intensity that tends to a constant average. Thus, the asset’s failure rate may be modelled as the sum of the average failure rate of all of its components.**

This statement must be qualified by noting that each segment must be renewed with an identical or equivalent component; thus, over its life, deterministic conditions, such as maintenance and inspection policies, must also be identical. In addition, we caution against making this approximation for new assets, as the time to [first] failure (TTF) will often be significantly different from the long-run time between failures (TBF) [5]. Further, this raises the question: “What does ‘long-run’ mean?”. As this question has arisen before, for example in the related area of maintenance planning [6][7] for fleets of equipment, we may draw from similar experiences. It has been noted [8] that for large fleets of equipment, the long run or steady-state behaviour will be reached sufficiently quickly for the assumption to be used; therefore, the steady state approximation is useful relatively soon after operation begins. However, as noted by Montgomery [9] and others [5], the steady-state is reached soon after “a number of failures have occurred, with the number of failures
dependent on the underlying failure distribution and the number of components in service" [9], so care must be taken to determine whether this approximation is applicable.

While this is an interesting observation in its own right, some types of linear assets are maintained in this way and can approximately be considered to have reached a steady state. The most common example is an electrical distribution line [10], though other above-ground assets such as telecommunications wires may often be modelled in a similar fashion. Accordingly, for these types of assets, a reliability analysis can be performed without needing to take into account the asset’s ageing factors. This is the method of analysis historically used for such assets as described by Billinton and Allan [10] and by Chowdhury and Koval [11] in the electrical power industry; it is also found in the analysis of other distribution assets, e.g. water [12][13][14], that regularly experience failure and repair. Thus, there is support for the application of these frequently-used methods from renewal theory.

However, many linear assets are not in this steady state, nor do they reach it over their expected service lives; such assets include lines that are newly installed, or those expected to have long service lives and a correspondingly high level of reliability—many transmission lines fall into this category. For such assets, equation (4-2) must be used in its original time-dependent form. Thus, it is necessary to account for subcomponent ageing in the analysis of the reliability of such assets.

Analysis of the failures of assets in this class must include a time or ageing component in the failure rate estimate until at least the first time-to-failure. Often it may be practical to model the times to the first, second, etc., failures with the same distribution as the first. This can be seen if we consider a line composed of \( n \) segments, each of whose aggregate failure rate (i.e. the hazard rate of the segment as a whole) is given as \( \lambda_i(t; \bar{z}_i(t)) \), defined as earlier. The overall hazard rate of the line until its first failure will then be

\[
\lambda_{\text{Line}}(t; \bar{z}(t) | t < T_1) = \sum_{i=1}^{n} \lambda_i(t; \bar{z}_i(t)),
\]

(4-4)
which, if we assume that all of the segments are identical, installed at time zero, and subject to the same conditions, simplifies to

\[ \lambda_{\text{Line}}(t; \tilde{z}(t)|t < T_1) = n \cdot \lambda(t; \tilde{z}(t)). \]

(4-5)

Immediately following the line’s first failure, which occurs at \( T_1 \), and prior to the second (occurring at \( T_2 \)) the failure intensity will be

\[ \lambda_{\text{Line}}(t; |t \geq T_1, t < T_2) = (n - 1) \cdot \lambda(t; \tilde{z}(t)) + \lambda(t - T_1; \tilde{z}(t)). \]

(4-6)

Assuming that the segments exhibit a wear-out effect, the overall system hazard will be lower after the failure than before. We can see that the new system failure intensity will differ from that of a system that has not experienced a failure and repair by:

\[
\Delta \lambda_{\text{Line}}(t; \tilde{z}(t)|\text{First Failure}) = \lambda_{\text{Line}}(t, \tilde{z}(t)|t < T_1) - \lambda_{\text{Line}}(t, \tilde{z}(t)|t \geq T_1, t < T_2)
= \lambda(t; \tilde{z}(t)) - \lambda(t - T_1; \tilde{z}(t)).
\]

(4-7)

If \( n \) is large (i.e. \( n \gg 1 \)), i.e. the line is of substantial length, then the fractional difference between a system that has experienced a single failure and one that has not will be minor, viz.

\[
\frac{\Delta \lambda_{\text{Line}}(t; \tilde{z}(t)|\text{First Failure})}{\lambda_{\text{Line}}(t; \tilde{z}(t))} = \frac{\lambda(t; \tilde{z}(t)) - \lambda(t - T_1; \tilde{z}(t))}{n \cdot \lambda(t; \tilde{z}(t))} = \frac{1}{n} \left( 1 - \frac{\lambda(t - T_1; \tilde{z}(t))}{\lambda(t; \tilde{z}(t))} \right),
\]

which will be strictly less than \( 1/n \) so long as \( \lambda(t; \tilde{z}(t)) \) is monotonically increasing.

As an example, consider the case where the segment failure hazard is given by the Weibull distribution. The fractional difference in hazard between a system that has experienced a single failure and one that has not will be

\[
\frac{\Delta \lambda_{\text{Line}}(t; \tilde{z}(t)|\text{First Failure})}{\lambda_{\text{Line}}(t; \tilde{z}(t))} = \frac{1}{n} \left( 1 - \frac{\beta}{\eta} \left( \frac{t - T_1}{\eta} \right)^{\beta - 1} \cdot \exp[\hat{y} \circ \tilde{z}(t)] \right) = \frac{1}{n} \left( 1 - \left( 1 - \frac{T_1}{t} \right)^{\beta - 1} \right),
\]

(4-9)

which tends to zero as \( t \) becomes large.
Based on this we may conclude that it may be useful to estimate an intensity or hazard function for the second failure using the original hazard function—i.e. it will often be practical and reasonable to assume a minimal repair, for the first few failures after installation at least, when estimating the reliability of a linear asset. However, this assumption will only be useful for a certain number of failures [5]. In the long run, it will likely be more practical to make use of Statement 1; in the intermediate phase of linear assets' lives it may be useful to estimate different forms of the failure intensity so as to ensure the greatest accuracy, even though there may not be a significant improvement by doing so.

As an alternative to these options, we may consider the use of different models to account for the ageing of a line's components over its whole life. Kijima's type II model [15] seems well-suited to this. A reasonable repair effectiveness, $\rho$, might be one equal to the fraction of the line that is renewed upon failure, i.e. $1/n$ if one segment fails and is renewed out of the $n$ in the asset. As before, let us examine the case where each of the line's segments has a failure rate which is described by a Weibull distribution, and all segments are subject to the same conditions. Prior to the first failure the system's estimated failure rate is

$$\lambda_{\text{Line}}(t; \tilde{z}(t) | t < T_1) = n \cdot \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} \cdot \exp[\tilde{y} \circ \tilde{z}(t)], \quad (4-10)$$

while after (and before the second failure) it is:

$$\lambda_{\text{Line}}(t; \tilde{z}(t) | t \geq T_1, t < T_2) = \left[ (n - 1) \cdot \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1} + \frac{\beta}{\eta} \left( \frac{t - T_1}{\eta} \right)^{\beta-1} \right] \cdot \exp[\tilde{y} \circ \tilde{z}(t)]. \quad (4-11)$$

If we use a virtual age model with a Weibull baseline function in a proportional intensities model, the model will have the form

$$\lambda_{\text{Line}_{VA}}(v(t); \tilde{z}(t)) = \frac{\beta_{VA}}{\eta_{VA}} \left( \frac{v(t)}{\eta_{VA}} \right)^{\beta_{VA}-1} \cdot \exp[\tilde{y} \circ \tilde{z}(t)], \quad (4-12)$$

where $v(t)$ is the virtual age, and is calculated as in (2-22). If we select the shape parameter $\beta_{VA}$ for the virtual age baseline to be the same as the shape parameter for the
line’s segments, the virtual age baseline’s scale parameter $\eta_{VA}$ will be $\eta/n^{(1/\beta)}$ if we wish to exactly match the line’s failure intensity before its first failure.

With these parameter selections, the bounds on the fractional difference between the explicitly calculated failure intensity and the virtual age model are minor after the first failure; the fractional differences (as in 4-8)) for some values of the segment shape parameter are shown below in Table 4-1.

| Segment Shape Parameter, $\beta$ | $\frac{\Delta \lambda_{Line,VA}(t, \tilde{z}(t)|First\ Failure)}{\lambda_{Line,Explicit}(t, \tilde{z}(t))}$ |
|----------------------------------|-------------------------------------------------|
| 1                                | 0                                               |
| 2                                | 0                                               |
| 3                                | $\leq \frac{1}{n}$                             |

*Table 4-1: Fractional deviation of the virtual age model failure hazard from the explicit failure hazard calculation following the first failure. See Appendix 4.A for a derivation of the above.*

In the long run, the virtual age model also approximates the steady-state condition, as the virtual age in the Kijima type II model may oscillate about a certain virtual age, where the interarrival time of failures equals or exceeds the product of the virtual age and the repair effectiveness, i.e. $v(t) \cdot \rho$. Thus, the virtual age model offers the potential of providing a useful failure baseline over a line’s whole life and may present a viable alternative to estimating multiple models for different parts of a line’s life, with the added benefit of reduced effort.

Based on the preceding, we can make our next statement about the failure behaviour of linear assets.

**Statement 2**—*Component ageing significantly affects the failure intensity of a linear asset, and its effect must be included in any reliability model of a system that is not very aged (i.e. an asset where Statement 1 does not apply). The effect of component replacement—as a result of a failure or otherwise—on the failure intensity of a linear asset depends on several factors, in particular the time-in-service of the asset, the number of failures it has previously experienced, and the number of components in the asset. Often the assumption of minimal repair may be made without great loss of accuracy. It may be useful, however, to use different models to estimate the failure intensity of a linear asset over certain parts of its life; certain reliability models (e.g. the virtual age model) may be considered for use over when considering the whole life of the asset.*
On the Relationship between Asset Length and Failure Intensity

One may note that in equation (4-2) the failure rate of a line is strongly dependent on the number of segments in the asset—the number of components that may fail. This relationship between number of segments and failure rate is noted in Xie et al. [16], where a pipeline is divided into \( n \) segments for the purposes of reliability analysis. Upon further reflection, one might consider that the commonly found \([17][10][18]\) proportionality of a linear asset’s length to its failure rate may simply be due to asset length being an effective proxy measurement for number of components in an asset, rather than some other factor.

If certain components are uniformly distributed among segments, i.e. there is a certain amount of pipe or wire and one pole or weld respectively per segment, the length of the asset will be strongly correlated to the number of these components in the asset. Thus, in a practical sense, it may be useful to simplify the reliability analysis of a linear asset by using length as a proportionality term instead of directly using the number of these components.

For illustrative purposes, consider a population of linear assets composed of three types of components—types A, B, and C. Let us compare two such assets: Line 1 has \( n_A, n_B, \) and \( n_C \) components of types A, B, and C respectively; Line 2 has \( m_A, m_B, m_C \) components of types A, B, and C respectively. Line 1 is \( L_1 \) km long, while Line 2 is \( L_2 \) km long. The failure rates for the two lines are given below; for simplicity, let us assume that the failure rate of each component may be modelled as a function of time only.

\[
\lambda_{\text{Line } 1}(t) = \sum_{i=1}^{n_A} \lambda_{A,i}(t) + \sum_{j=1}^{n_B} \lambda_{B,j}(t) + \sum_{l=1}^{n_C} \lambda_{C,l}(t),
\]

\[
\lambda_{\text{Line } 2}(t) = \sum_{i=1}^{m_A} \lambda_{A,i}(t) + \sum_{j=1}^{m_B} \lambda_{B,j}(t) + \sum_{l=1}^{m_C} \lambda_{C,l}(t).
\]

If the lines are newly installed, the hazard rates of the lines prior to their first failures will be as follows:

\[
\lambda_{\text{Line } 1}(t | t < T_1) = n_A \lambda_A(t) + n_B \lambda_B(t) + n_C \lambda_C(t),
\]

\[
\lambda_{\text{Line } 2}(t | t < T_1) = m_A \lambda_A(t) + m_B \lambda_B(t) + m_C \lambda_C(t).
\]
If we suppose that the lines’ components have similar wearout characteristics (e.g. if they could all be modelled with a Weibull distribution with identical shape parameters), all proportional to some function $\lambda(t)$ with some constant $\alpha$, we may express the initial hazard rates as follows:

\[
\lambda_{\text{Line } 2}(t|t < T_1) = m_A \lambda_A(t) + m_B \lambda_B(t) + m_C \lambda_C(t).
\]

(4-15)

\[
\lambda_{\text{Line } 1}(t|t < T_1) = (n_A \alpha_A + n_B \alpha_B + n_C \alpha_C) \lambda(t),
\]

\[
\lambda_{\text{Line } 2}(t|t < T_1) = (m_A \alpha_A + m_B \alpha_B + m_C \alpha_C) \lambda(t).
\]

Let us then suppose that the number of type C components is independent of the number of the other two components and is easily determined, and that type A and type B components are roughly uniformly distributed throughout both lines, such that

\[
\frac{n_A}{m_A} \approx \frac{n_B}{m_B} \approx \frac{L_1}{L_2},
\]

(4-16)

If we similarly assume some proportional ratio between length and number of components of each type, viz.,

\[
\frac{L_1}{n_A} \approx \frac{L_2}{m_A} \approx a, \text{ and } \frac{L_1}{n_B} \approx \frac{L_2}{m_B} \approx b,
\]

(4-17)

we may rewrite the initial hazards as:

\[
\lambda_{\text{Line } 1}(t|t < T_1) \approx \left(\frac{\alpha_A}{a} + \frac{\alpha_B}{b}\right) L_1 + n_C \alpha_C \lambda(t),
\]

\[
\lambda_{\text{Line } 2}(t|t < T_1) \approx \left(\frac{\alpha_A}{a} + \frac{\alpha_B}{b}\right) L_2 + m_C \alpha_C \lambda(t).
\]

(4-18)

Note that the hazards of the lines are proportional to a function of length and the number of some other component. This may be expressed in the common notation used in the proportional hazards/intensities models as follows, with some rearrangement:

\[
\lambda_{\text{Line }}(t|t < T_1, L_{\text{Line}}, n_{c,\text{Line}}) \approx \exp\left\{\ln\left[\left(\frac{\alpha_A}{a} + \frac{\alpha_B}{b}\right) L_{\text{Line}} + \alpha_C \cdot n_{c,\text{Line}}\right]\right\} \lambda(t).
\]

(4-19)
If all components in a line are approximately uniformly distributed, there will be a direct proportionality between length and failure rate, with the constant of proportionality being equal to:

\[
\left( \frac{\alpha_A}{a} + \frac{\alpha_B}{b} + \frac{\alpha_C}{c} + \cdots \right).
\]  

(4-20)

In the proportional hazards model, we might bring this constant of proportionality into the baseline hazard function \( \lambda_0(t) \). Accordingly, the model for the line would be:

\[
\lambda_{\text{Line}}(t|t < T_1, L_{\text{Line}}) \approx \exp\{\gamma_0 \cdot \ln[L_{\text{Line}}]\} \lambda_0(t);
\]

(4-21)

\( \gamma_0 \) should be close to unity in a properly-estimated PHM/PIM.

We may repeat the above exercise by assuming that components in an asset are all of some “average” age \( \bar{\ell} \) at time \( t \); in this case, the equations in (4-14) would be:

\[
\lambda_{\text{Line}}^1(t|t < T_1) \approx n_A \lambda_A(\bar{\ell}(t)) + n_B \lambda_B(\bar{\ell}(t)) + n_C \lambda_C(\bar{\ell}(t)),
\]

\[
\lambda_{\text{Line}}^2(t|t < T_1) \approx m_A \lambda_A(\bar{\ell}(t)) + m_B \lambda_B(\bar{\ell}(t)) + m_C \lambda_C(\bar{\ell}(t)),
\]

and equation (4-19) could be re-expressed as:

\[
\lambda_{\text{Line}}(t|t < T_1, L_{\text{Line}}, n_{c,\text{Line}}) \approx \exp\left\{\ln \left[ L_{\text{Line}} + \frac{a \cdot b \cdot \alpha_C}{b \cdot \alpha_A + a \cdot \alpha_B} \cdot n_{c,\text{Line}} \right] \right\} \cdot \left( \frac{\alpha_A}{a} + \frac{\alpha_B}{b} \right) \cdot \lambda(\bar{\ell}(t)).
\]

(4-23)

The proportional hazards/intensities framework might result in a model of the form

\[
\lambda_{\text{Line}}(t|t < T_1, L_{\text{Line}}) \approx \exp\{\gamma_0 \cdot \ln[L_{\text{Line}} + x \cdot n_{c,\text{Line}}]\} \lambda_0(t),
\]

(4-24)

where \( x \) is the parameter stating the relative proportion of line failures caused by components whose number is not proportional to length (in this example, Type C components). The constant \( x \) would need to be determined when estimating the model. This type of relationship was found in a study of the failure of gas distribution pipes [17], and the proportionality of failure rate with length alone is well-documented. The exclusion of the additional terms, i.e. the number of components not proportional to the length of the asset, might explain results in Kleiner et al.’s analyses [12], where the failure rate of long
pipelines was proportional with length, but the correlation of length with failure decreased for short pipelines.

Based on the preceding, we can make our next statement about the failure behaviour of linear assets and modeling the failure behaviour of linear assets.

**Statement 3**—*The failure rate of a linear asset is strongly dependent on the number of components that make up the asset; the number of components in a linear asset is strongly affected by its length. When modeling the reliability of a population of linear assets, it is necessary to model the proportionality of the failure rate to the number of components. While the lengths of the lines are often useful proxy measurements for the number of certain components within the linear assets, it is most accurate to explicitly use the number of components rather than the length.*
Including Asset Information & Condition Measurements in a Reliability Model

Including Asset Information in a Reliability Model

It is often useful to estimate a reliability model for lines with differing characteristics (e.g. different construction methods, right-of-way types, or materials), just as it is useful in epidemiology to estimate a survival model for the whole varied population of a country. One of the more common uses of the proportional hazards/intensities models in the reliability field is, similarly, in using the same model to estimate the failure rate of components or assets of similar but not identical type and operation. It may be used for the same purposes in the case of linear assets.

Let us begin this discussion by supposing that a line’s linear component may be composed of two types of material—materials A and B. This may be reflected in each segment type’s reliability model with a PHM including a binary covariate as follows, with material A being considered the “base” material:

\[
\lambda_A(t; \bar{z}(t), z_1 = 0) = \lambda_0(t) \cdot \exp[y_1 \cdot 0] \cdot \exp[\gamma \circ \bar{z}(t)]; \\
\lambda_B(t; \bar{z}(t), z_1 = 1) = \lambda_0(t) \cdot \exp[y_1 \cdot 1] \cdot \exp[\gamma \circ \bar{z}(t)].
\]

It is easy to see that if a line is composed solely of one type of segment or the other that the line’s hazard rate will also be proportional with the value \(\exp[y_1 \cdot z_1 + y_2 \cdot z_2]\). As an example, suppose that Line 1 consists solely of \(n\) material A segments. Then its failure rate will be

\[
\lambda_{\text{Line 1}}(t; \bar{z}(t), z_1 = 0) = \sum_{i=1}^{n} \lambda_A(t; \bar{z}(t), z_1 = 0) \\
= n \cdot \lambda_0(t) \cdot \exp[\gamma \circ \bar{z}(t)],
\]

while the failure rate for Line 2 consisting solely of \(m\) material B segments will be

\[
\lambda_{\text{Line 2}}(t; \bar{z}(t), z_1 = 1) = \sum_{i=1}^{m} \lambda_B(t; \bar{z}(t), z_1 = 1) \\
= m \cdot \lambda_0(t) \cdot \exp[\gamma \circ \bar{z}(t)] \cdot \exp[y_1],
\]
and we can see that the proportionality found in the individual segments directly applies to the lines as a whole. This type of model has been successfully used to estimate the risk of failure of gas distribution lines [17][19], water lines [18], and electrical lines.

However, if a line is not composed solely of segments of one type of material, or some other characteristic, the use of binary covariates is not as straightforward. Consider an example where the line is composed of \( n \) segments of type A and \( m \) of type B. The line’s failure intensity is (with the simplifying assumption that all other covariates are the same):

\[
\lambda_{\text{Line}}(t; \tilde{z}(t)) = n \cdot \lambda_0(t) \cdot \exp[\tilde{y} \circ \tilde{z}(t)] + n \cdot \lambda_0(t) \cdot \exp[\tilde{y} \circ \tilde{z}(t)]
\]

\[
= (n + m) \cdot \lambda_0(t) \cdot \exp[\tilde{y} \circ \tilde{z}(t)] \cdot \left[ \frac{n}{n + m} + \frac{m \cdot \exp[y_1]}{n + m} \right].
\]

One natural approximation of this would be to use fractional values for the binary covariates; in the above example, \( z_I \) would be given a value of \( m/(n+m) \) and we might estimate the line’s risk of failure with the following approximate model:

\[
\lambda_{\text{Line}}(t; \tilde{z}(t)) = (n + m) \cdot \lambda_0(t) \cdot \exp[\tilde{y} \circ \tilde{z}(t)] \cdot \exp \left[ 0 + \frac{m}{n + m} \cdot y_1 \right].
\]

Taking the Taylor Series of (4-28)\(^3\) and (4-29)\(^4\), we see that the approximation is identical to the first order term only; this relation holds with more than two material types (or some other characteristic). Thus, it may be accurate in some cases, especially where the covariate parameters (e.g. \( y_1 \) and \( y_2 \)) are small and/or there are small differences between them, to use binary covariates with fractional values to model the presence of different segment types within a linear asset.

\(^3\) The Taylor series of the latter term of (4-28) is \( \left[ \frac{n}{n+m} + \frac{m}{n+m} \cdot (1 + y_1 + y_1^2 + y_1^3 + \ldots) \right] \), which simplifies to \( \left[ 1 + \left( \frac{m}{n+m} \right) \cdot y_1 + \left( \frac{m}{n+m} \right) \cdot y_1^2 + \left( \frac{m}{n+m} \right) \cdot y_1^3 + \ldots \right] \).

\(^4\) The Taylor series of the latter exponential term of (4-29) is \( \left[ 1 + \left( \frac{m}{n+m} \right) \cdot y_1 + \left( \frac{m}{n+m} \right)^2 \cdot y_1^2 + \left( \frac{m}{n+m} \right)^3 \cdot y_1^3 + \ldots \right] \), which may be rearranged as \( \left[ 1 + \left( \frac{m}{n+m} \right) \cdot y_1 + \left( \frac{m}{n+m} \right)^2 \cdot y_1^2 + \left( \frac{m}{n+m} \right)^3 \cdot y_1^3 + \ldots \right] \).
However, there is a significant discrepancy between the explicit and approximate formulae when the covariate parameters are far apart and the fractional binary covariates are close to each other. Thus, it may be beneficial to model separately portions of a line in which the segments of one type are more concentrated, i.e. when deciding on the [virtual [16]] sectioning a line for reliability analysis purposes, the static conditions and method of construction of the line’s segments and component parts should be strongly considered.

Based on the preceding discussion, we can make our next statement about modeling the failure behaviour of linear assets.

**Statement 4**—Where a population of linear asset is composed of lines with segments of differing characteristics (e.g. different construction methods, right-of-way types, or materials), it is often useful to model their differing failure rates using a PHM/PIM with a set of binary covariates. If these characteristics are not uniform throughout a single line, then the use of fractional values in place of the binary covariates, with the fractions as the proportion of each line with the appropriate characteristic, may be effective in modelling their differing reliability with some accuracy.

**Inclusion of Process, Condition Monitoring, and External Condition Measurements in a Reliability Model**

While it is common to use the PHM or PIM to estimate the reliability of assets of a similar but not identical type using the same statistical model, their greatest success in the reliability field is their ability to account for the effect of operating conditions on the risk of failure of assets [6], by including measurements in the PHM/PIM 5. These models have great utility in reliability and maintenance engineering in condition-based maintenance (CBM) and other proactive maintenance philosophies, such as total productive maintenance (TPM) [20]. As the proportional hazards/intensities models have already been applied to linear assets of various types (e.g. gas distribution lines [17][19], telecommunications lines [21], and water lines [18]), we will only briefly discuss their application to linear assets.

---

5 As noted earlier, "condition measurements" may include process measurements, external condition measurements such as weather, or inspection reports, among others.
As in the preceding discussion, if a condition variable applies uniformly to the segments that make up an asset, the proportionality in segment failure rate also applies to the line as a whole. Similarly, if there is some difference in the conditions that apply to various segments of the asset, some average of the conditions, as in (4-29), may be an effective approximation. As found in previous studies (e.g. [10] [21]), weather-related covariates—which will obviously not be uniform over the large area that any linear asset of non-trivial size will occupy—can be incorporated into a reliability model for linear assets. This experience suggests that it is practicable with many other external covariates or process and condition monitoring information, and we make our penultimate statement about reliability modeling of linear assets.

Statement 5—A reliability model of a population of linear assets should attempt to incorporate the effect on failure rate of internal process measurements (e.g. rate of flow, current, etc.) and conditions (e.g. inspection data), as well as external conditions such as weather. The PHM/PIM offers a useful method for doing so, though there are alternative models for estimating risk of failure that might be effective. As long as these conditions do not vary significantly over the length of the asset, they may be accurately included; if they do vary significantly, Statement 4 may be useful.

Finally, based on the discussion in this subsection, we make our last statement about the virtual sectioning of a linear asset to facilitate analysis of its failure behaviour.

Statement 6—When deciding on the virtual sectioning of a linear asset for the purposes of reliability analysis, the static and dynamic conditions and method of construction of the line’s segments and component parts should be considered. It may also be beneficial to attempt to model separately portions of a line in which the segments of one type are more concentrated than another, or certain external conditions such as calm or poor weather predominate.
### 4.3 A Linear Asset as a System of Three-State Line Sections

Having discussed how an analysis of a series of simple (i.e. binary state) components correctly models many of the reliability features of linear assets, we direct our attention to complications and extensions of this model. This is done to see if we can make useful predictions about the reliability of linear assets and whether examination of this theoretical model may provide a useful guide to conducting reliability analyses of linear assets in the future.

Let us begin by exploring the implications of modeling a linear asset as a line of segments that may be in a “damaged” state, but still functioning, along with potential “failed” and “operating” states; this extension is illustrated in Figure 4-3. To a user or the system operator, this system would still appear, functionally, to have the two original states: failed and operating. But to the reliability engineer and line maintainers, the damage to a line’s segment(s) may represent some defect that may affect the asset’s failure rate.

*Figure 4-3: (Left) Standard model of segment failure behaviour; (Right) Proposed three-state model of segment failure behaviour.*
In the following discussion, we will suppose that, by some process, a segment that has been damaged may return to the “normal” state or that it may fail; it may also transition from “normal” to “failed”. (As we are attempting to model the failure rate of a linear asset, and not the repair rate, we will not consider transitions from the “failed” state to “normal”.)

We will label the states as follows (see Figure 4-4): direct transitions of a segment’s reliability from “Normal” to “Failed” states occur with rate $\varepsilon(t; \tilde{z}(t))$, transitions from the “Normal” to the “Damaged / High Risk” states occur with rate $\alpha(t; \tilde{z}(t))$, and transitions from “Damaged / High Risk” to “Normal” or “Failed” occur with rates $\mu(t; \tilde{z}(t))$ and $\zeta(t; \tilde{z}(t))$ respectively.

Figure 4-4: Labelled three-state model of segment failure behaviour.

---

6 This could be as a result of some maintenance action, or some natural phenomenon. An example of the latter might be a tree that has fallen on a segment of an electrical line, resting on it and causing that segment to be at greater risk of failing; it might return to “normal” if the tree simply slides off of the line without breaking it.
Chapter 4: Modelling an Ideal Linear Asset

The observed failure rate of a line of \( n \) such segments will be

\[
\lambda(t; \bar{z}(t)) = (n - N(t)) \cdot e(t; \bar{z}(t)) + N(t) \cdot \zeta(t; \bar{z}(t)),
\]

which is the sum of the failures from the “Normal” segments and the failures caused by the \( N(t) \) segments that are damaged, where \( N(t) \) is governed by the following differential equation:

\[
dN(t) = n \cdot \alpha(t; \bar{z}(t))dt - N(t) \left( \mu(t; \bar{z}(t)) + \zeta(t; \bar{z}(t)) + \alpha(t; \bar{z}(t)) \right) dt.
\]

Such a model could be estimated for a line, but with great difficulty. As this may seem somewhat intractable, we will continue the discussion by making some simplifying assumptions. If the number of undamaged segments is significantly larger than the number of damaged ones (i.e. \( n \gg N(t) \)), we can approximate (4-30) as

\[
\lambda(t; \bar{z}(t)) = n \cdot \alpha(t; \bar{z}(t)) + N(t) \cdot \zeta(t; \bar{z}(t)),
\]

and approximate the rate of change in \( N(t) \) with the differential equation:

\[
dN(t) = n \cdot \alpha(t; \bar{z}(t))dt - N(t) \left( \mu(t; \bar{z}(t)) + \zeta(t; \bar{z}(t)) \right) dt.
\]

Let us further suppose, for the sake of simplicity, that segments may only be damaged when some component \( z_P(t) \) of the covariate vector is greater than some threshold level \( Z_P \), and zero otherwise (we will call the period when \( z_P(t) > Z_P \) a “damaging event”); this significantly simplifies (4-31) and (4-33) as the \( n \cdot \alpha(t; \bar{z}(t))dt \) term takes on non-zero values for particular times only, thereby facilitating our analysis by eliminating the need to include \( \alpha(t; \bar{z}(t)) \) for most times. This results in the decay-type equation

\[
dN(t) = -N(t) \left( \mu(t; \bar{z}(t)) + \zeta(t; \bar{z}(t)) \right) dt,
\]

for most \( t \), which has as its solution (if we assume there are \( N_0 \) damaged components at some time \( T_0 \), and none become damaged before some time \( T_1 > T_0 \)):
Chapter 4: Modelling an Ideal Linear Asset

\[ N(t|t > T_0, t < T_1) = N_0 \cdot \exp \left[ - \int_{T_0}^{t} (\mu(t; \tilde{z}(t)) + \zeta(t; \tilde{z}(t))) \, dt \right]. \]  

(4-35)

As an example, if we suppose that \( z_P(t) \) is greater than the threshold level \( Z_P \) for some period \( \Delta t \) starting at time \( T_0 \), then the expected number of segments of the line damaged in time \( \Delta t \), \( N_0 \), will be approximately

\[ N_0 = \alpha(T_0; \tilde{z}(T_0)) \cdot \Delta t \cdot n, \]  

(4-36)

if we make the assumption that there were no damaged components at time \( T_0 \); if there were damaged components we could add the above value to our current estimate of \( N(t) \). If we attempt to model the failure hazard rate of a line, rather than the failure intensity, \( N(t) \) is simplified further as the \( \zeta(t; \tilde{z}(t)) \) term may be removed because we are conditioning the failure rate of the line on it not experiencing a failure up to time \( t \). Therefore, we have:

\[ N(t|t > T_0, t < T_1) = N_0 \cdot \exp \left[ - \int_{T_0}^{t} \mu(t; \tilde{z}(t)) \, dt \right], \]  

and following \( T_0 \) the observed failure hazard rate of the line as a whole will be given by

\[ \lambda(t; \tilde{z}(t)|t > T_0, t < T_1) = n \cdot \epsilon(t; \tilde{z}(t)) + N_0 \cdot \exp \left[ - \int_{T_0}^{t} \mu(t; \tilde{z}(t)) \, dt \right] \cdot \zeta(t; \tilde{z}(t)). \]  

(4-37)

If we suppose that \( \zeta(t; \tilde{z}(t)) \) is proportional to \( \epsilon(t; \tilde{z}(t)) \) with some parameter \( q \), and that \( \mu(t; \tilde{z}(t)) \) is a constant \( \mu \), we may express the hazard function as:

\[ \lambda(t; \tilde{z}(t)|t > T_0, t < T_1) = \left[ 1 + \frac{q \cdot N_0}{n} \cdot \exp[-\mu(t - T_0)] \right] \cdot n \cdot \epsilon(t; \tilde{z}(t)). \]  

(4-38)

If we then assume that the “normal” failure rate of a line of \( n \) segments is described by a standard PHM designed as suggested in 4.2, i.e.

\[ \lambda_{Normal}(t) = n \cdot \epsilon(t; \tilde{z}(t)) = n \cdot \lambda_0(t) \cdot \exp[\tilde{y} \circ \tilde{z}(t)], \]  

we may rewrite the hazard as
Chapter 4: Modelling an Ideal Linear Asset

\[ \lambda(t; \tilde{z}(t) | t > T_0, t < T_1) = \lambda_0(t) \cdot \exp \left[ \tilde{z} \cdot \tilde{z}(t) + \ln[n] + \ln \left[ 1 + \frac{q \cdot N_0 \cdot \exp[-\mu(t - T_0)]}{n} \right] \right], \]

which is a more tractably estimated function than (4-30) and (4-31) together, and might be more straightforward to estimate.

In any case, such a model as the simplified one in (4-41) or the more general case in (4-30) and (4-31) may be useful in modelling the failure bunching phenomenon noted in electric power transmission and distribution systems [21][22]. A model like the simplified one in (4-41) could be placed easily within the PHM/PIM framework by using a time-dependent hazard ratio for lines recently experiencing a damaging event. Methods for estimating such models have been developed in epidemiology [23][24] and could be transferred to the reliability and maintenance engineering field. This model could effectively model failure bunching, while maintaining the use of widely-applied reliability models such as the PHM.

We may also explore variations on this model. For instance, in what sort of system might \( \mu(t; \tilde{z}(t)) \) be zero? That is, in what sort of system would segments or component parts not return to a “normal” operating state after damage? In such a system \( N(t) \) will continue growing as the asset experiences damaging events; similarly aged lines that have experienced more damaging events will be at greater risk of failure due to damaged components. If damaging events commonly occur at the same time that a line typically experiences failures—a reasonable approximation, in many cases—this model could provide a partial explanation for the association of failure history (e.g. number of failures that a line has experienced) with increases in failure intensity [17][19].

Thus, the extension of our simple initial model from Section 4.2 can shed light on the failure behaviour of linear assets, leading to better suggestions for modeling their failure rates. Such improved models could be useful in justifying or optimizing maintenance decisions and supporting reliability strategies.
4.4 Development of Maintenance Strategies

With improved reliability models, we may be able to develop improved maintenance and reliability engineering practices. In the following section, we briefly discuss some maintenance decisions that may be supported with the reliability models discussed earlier. These are only two of the many possibilities; a myriad of decision-support tools could potentially use reliability models as inputs; for further examples interested readers are directed to such works as Dekker [25], Van Horenbeek et al. [26], or Jardine and Tsang [6].

Selection of Segment/Component Types

One common maintenance decision that may be supported with a reliability model is the selection of one type of component among several. Let us suppose that we are managing a distribution line that is currently in operation, and need to decide whether to use material A or material B for a certain component distributed through the \( n \) segments of the line. Both materials provide equivalent operational characteristics, except that a segment with the material A component is more reliable given the conditions in which the line will be operating, with a hazard ratio of \( R \) between the failure rates of A and B. The new components will be installed as the current ones on the line experience failure. The cost to purchase and install a component of material A is \( m \) times that of material B, which has cost \( C_B \). Upon failure, the failed component is removed from service and replaced. The cost of experiencing a failure is \( C_F \). The line segments with components made out of these materials follow hazard rates with a Weibull baseline, viz.

\[
\begin{align*}
\lambda_A(t; z_1 = A) &= \frac{\beta (t)}{\eta} \left( \frac{t}{\eta} \right)^{\beta - 1}, \\
\lambda_B(t; z_1 = B) &= \frac{\beta (t)}{\eta} \left( \frac{t}{\eta} \right)^{\beta - 1} \cdot R.
\end{align*}
\]
As the line is a distribution line, and we are interested in long run costs, we assume that we may make use of Statement 1. We make our decision based on the long-run cost rate of selecting each component. The cost rate may be calculated as:

\[
\text{Cost Rate} = \text{Failure Rate} \cdot \text{Failure Cost},
\]

where the failure cost is the total cost experienced at each failure, namely the cost of installing the new component and experiencing the failure. For the line with component type A, the cost rate is

\[
\text{Cost Rate}_A = \frac{n}{MTBF_A} \cdot (m \cdot C_B + C_F),
\]

making use of equation (4-3); meanwhile, for the line with component type B

\[
\text{Cost Rate}_B = \frac{n}{MTBF_B} \cdot (C_B + C_F).
\]

We now calculate the mean time between failures for each to make a decision. For the Weibull baseline and a given hazard ratio, \( R \), these are [27]

\[
MTBF_A = \eta \cdot \Gamma \left(1 + \frac{1}{\beta}\right),
\]

\[
MTBF_B = \eta \cdot R^{1/\beta} \cdot \Gamma \left(1 + \frac{1}{\beta}\right),
\]

and the ratio of the cost rates is:

\[
\frac{\text{Cost Rate}_A}{\text{Cost Rate}_B} = \frac{(m \cdot C_B + C_F)}{(C_B + C_F)} R^{1/\beta}.
\]

So we should select component A for our line if the above ratio is less than 1, and select component B otherwise.
Determining Whether to Inspect a Linear Asset after it Experiences a Damaging Event

Let us examine a decision model that uses the three-state model suggested above. Suppose that the hazard for a distribution line follows equation (4-41), and we wish to determine whether or not to inspect a line following a damaging event. The inspection will detect and correct any components that are in a damaged state; the cost of inspection is $C_i$. The cost of experiencing a failure and repairing the line is $C_f$; this includes the work of inspection that must be done to find the failed component—thus when a failure is experienced, the cost is $C_i + C_f$. (We also assume that any damaged components are corrected during this inspection, as with the standard inspection.)

As the distribution line has been in service for a sufficient period, we may make use of Statement 1. Let us simplify our analysis further by making the approximation that the line's condition measurements are constant and that it will not experience any more damaging events over our time horizon of 90 days. With this case, the failure hazard rate of the line following will be a function of time only:

\[
\lambda(t, T_0 | t < T) = \bar{\lambda} \cdot \left[ 1 + \frac{q \cdot N_0}{n} \cdot \exp[-\mu(t - T_0)] \right],
\]

where $\bar{\lambda}$ is the average failure rate from undamaged components on the line, estimated to be 0.01/day.

If we fit such a model and have reason to believe that a typical damaging event affects 4% of the line (i.e. $N_0/n = 0.04$), that damaged components are 50 times as likely as undamaged ones to experience a failure (i.e. $q=50$), and that the damaged components return to normal at a rate of 0.1/day, then (4-46) becomes

\[
\lambda(t|t < T, T_0) = 0.01 \cdot [1 + 2 \cdot \exp[-0.1 \cdot (t - T_0)]].
\]

If we inspect the line, the expected failure rate over the time horizon will simply be a constant $\bar{\lambda}$ (0.01/day, as noted above), while if it is not inspected, the hazard rate will
follow equation (4-47). To compare costs, we first consider the alternative cumulative hazards at the end of the time horizon:

\[
\Lambda_{\text{Inspect}} = \int_{t=T_0}^{T_0+90} \lambda(t; t < T, T_0) \, dt = \int_{t=T_0}^{T_0+90} \bar{\lambda} \, dt = 90 \cdot 0.01 = 0.9,
\]

\[
\Lambda_{\text{Don't Inspect}} = \int_{t=T_0}^{T_0+90} 0.01 \cdot [1 + 2 \cdot \exp[-0.1 \cdot (t - T_0)]) \, dt = 1.1.
\]

For the case where we choose to inspect our line, the cumulative hazard is equivalent to the expected number of failures over the interval, as the failure rate is constant. Thus, the expected cost of the choice to inspect is

\[
(4-49)
\]

\[
\text{Total Cost if Inspect} = C_i + \Lambda_{\text{Inspect}} \cdot C = C_i + 0.9 \cdot C.
\]

For the case where we choose not to inspect the line, the calculations are somewhat more difficult. The cost over our time horizon if the line does not fail is zero, as we have incurred no inspection cost and there is no failure cost. However, the line’s reliability at the end of our time horizon is rather poor:

\[
(4-50)
\]

\[
R_{\text{Don't Inspect}} = \exp(-\Lambda_{\text{Don't Inspect}}) = 0.333,
\]

so we must determine the expected cost when the line does fail. We may find this by determining the mean time of failure and the mean remaining time in which the line is at its average failure rate $\bar{\lambda}$. We first need to calculate the failure time probability distribution using (2-5); this is plotted graphically below in Figure 4-5.
The cumulative hazard following the failure must be calculated to determine the expected number of failures that occur after the first:

\[
\Lambda_{\text{After Failure}} = \bar{\lambda} \cdot \left[ \int_{t=T_0}^{T_0+90} (90 - t) f(t) dt \right] = 0.5.
\]  \hspace{1cm} (4-51)

The expected total cost for our choice not to inspect is:

\[
\text{Total Cost if Don't Inspect} = F_{\text{Don't Inspect}} \cdot \left[ (C_1 + C) + \Lambda_{\text{After Failure}} \cdot C \right] \\
= 0.667 \cdot [C_1 + C + 0.5C] \\
= 0.667 \cdot C_1 + C
\]  \hspace{1cm} (4-52)

The difference in costs between the two choices is:

\[
\text{Total Cost if Don't Inspect} - \text{Total Cost if Inspect} = 0.333 \cdot C_1 - 0.1C.
\]  \hspace{1cm} (4-53)

We see that the economically optimal choice is to inspect the line after a damaging event if the cost of inspection is less than 30% of the cost of experiencing a failure, and to not inspect otherwise.
4.5 Summary and Potential Areas of Future Exploration

The simple model of an idealised linear asset presented in Section 4.2 can provide useful explanations of the failure behaviour of linear assets, and hence suggest better ways to model their failure rates. Extending this model with a three-state component/segment model, as shown in Section 4.3, could build a theoretical base for studies of more complicated linear asset behaviour, even with restrictive assumptions. Future work could focus on advanced study of our three-state model and on determining how to present and estimate such a model for use in reliability analysis and decision-making.

Additionally, expanded use of time-dependent hazard ratios (TDHRs) in modeling complex phenomena in system failure behaviour could benefit reliability analysts, whether they are modeling linear assets or point assets. The use of TDHRs permits continued use of the already-advanced PHM/PIM modelling framework while giving the ability to handle more complex system behaviour.

Finally, the improved integration of physically-based wear and reliability models with statistically-based models would be of great utility to managers of linear assets. For instance, the baseline failure rate of a pipeline may be strongly based on the expected wear rate of such a pipe under certain flow conditions. The inspection of this pipe may permit a more accurate gauge of its true “age”, resulting in a more accurate reliability estimate. Because such linkages between observation-based and statistically-based reliability models have the potential to significantly improve reliability forecasts, they should be explored.
Chapter 4: Modelling an Ideal Linear Asset

References


4.A Appendix

In this derivation, we have selected the shape parameter $\beta_{VA}$ for the virtual age baseline to be the same as the shape parameter for the line’s segments, and the virtual age baseline’s scale parameter $\eta_{VA}$ will be $\eta/(1/\beta)$ as we wish to exactly match the line’s failure intensity before its first failure.

For the virtual age models, the failure intensity for the line following the first failure of the line at $t = T_1$ will be given as a function of the virtual age, which will be

$$v(t|t > T_1, t < T_2) = t - \rho \cdot T_1,$$

where $\rho$ is the repair effectiveness, set to $1/n$, and the modelled system failure intensity will be

$$\lambda_{\text{Line } V.A.}(v(t|t > T_1, t < T_2); \tilde{z}(t)) = \frac{\beta_{VA}}{\eta_{VA}} \left(\frac{t - \rho \cdot T_1}{\eta_{VA}}\right)^{\beta_{VA}-1} \cdot \exp[\bar{y} \circ \tilde{z}(t)].$$

After subbing in our values $\beta_{VA} = \beta$, $\eta_{VA} = \eta/(1/\beta)$, and $\rho = 1/n$ this becomes

$$\lambda_{\text{Line } V.A.}(v(t|t > T_1, t < T_2); \tilde{z}(t)) = \frac{\beta}{\eta/(1/\beta)} \left(\frac{t - (1/n) \cdot T_1}{\eta/(1/\beta)}\right)^{\beta-1} \cdot \exp[\bar{y} \circ \tilde{z}(t)],$$

which simplifies to

$$\lambda_{\text{Line } V.A.}(v(t|t > T_1, t < T_2); \tilde{z}(t)) = \frac{n\beta}{\eta} \left(t - (1/n) \cdot T_1\right)^{\beta-1} \cdot \exp[\bar{y} \circ \tilde{z}(t)].$$

The explicit model gives a failure intensity following the first failure as follows, after some rearrangement

$$\lambda_{\text{Line}}(t; \tilde{z}(t)|t \geq T_1, t < T_2) = \frac{\beta}{\eta^\beta} \cdot [(n - 1) \cdot (t)^{\beta-1} + (t - T_1)^{\beta-1}] \cdot \exp[\bar{y} \circ \tilde{z}(t)].$$

This gives us a ratio of the virtual age intensity to the explicitly modelled one of

$$\frac{\lambda_{\text{Line } V.A.}(v(t|t > T_1, t < T_2); \tilde{z}(t))}{\lambda_{\text{Line}}(t; \tilde{z}(t)|t \geq T_1, t < T_2)} = \frac{n\beta}{\eta^\beta} \left(t - (1/n) \cdot T_1\right)^{\beta-1} \cdot \exp[\bar{y} \circ \tilde{z}(t)] / \frac{\beta}{\eta^\beta} \cdot [(n - 1) \cdot (t)^{\beta-1} + (t - T_1)^{\beta-1}] \cdot \exp[\bar{y} \circ \tilde{z}(t)].$$
which can be simplified significantly to

\[
\frac{\lambda_{LineVA}(v(t|t > T_1, t < T_2); \bar{z}(t))}{\lambda_{Line}(t; \bar{z}(t)|t \geq T_1, t < T_2,)} = \frac{n(t - (1/n) \cdot T_1)^{\beta - 1}}{[(n - 1) \cdot (t)^{\beta - 1} + (t - T_1)^{\beta - 1}]}
\]

For \( \beta = 1 \), the above ratio is:

\[
\frac{\lambda_{LineVA}(v(t|t > T_1, t < T_2); \bar{z}(t))}{\lambda_{Line}(t; \bar{z}(t)|t \geq T_1, t < T_2,)} = \frac{n(t - (1/n) \cdot T_1)^0}{[(n - 1) \cdot (t)^0 + (t - T_1)^0]} = \frac{n}{n - 1} = 1
\]

For \( \beta = 2 \), the above ratio is:

\[
\frac{\lambda_{LineVA}(v(t|t > T_1, t < T_2); \bar{z}(t))}{\lambda_{Line}(t; \bar{z}(t)|t \geq T_1, t < T_2,)} = \frac{n(t - (1/n) \cdot T_1)^1}{[(n - 1) \cdot (t)^1 + (t - T_1)^1]} = \frac{n \cdot t - T_1}{(n - 1) \cdot (t) + n \cdot t - T_1} = 1
\]

For \( \beta = 3 \), the above ratio is:

\[
\frac{\lambda_{LineVA}(v(t|t > T_1, t < T_2); \bar{z}(t))}{\lambda_{Line}(t; \bar{z}(t)|t \geq T_1, t < T_2,)} = \frac{n(t - (1/n) \cdot T_1)^2}{[(n - 1) \cdot (t)^2 + (t - T_1)^2]} = \frac{n \cdot t^2 - 2 \cdot t \cdot T_1 + (1/n) \cdot T_1^2}{n \cdot t^2 - 2 \cdot t \cdot T_1 + T_1^2}
\]

and its fractional difference, or the difference of the above ratio from unity is:

\[
\frac{\left(\frac{n - 1}{n}\right) \cdot T_1^2}{n \cdot t^2 - 2 \cdot t \cdot T_1 + T_1^2}
\]

and since \( T_1 \) is necessarily less than \( t \) we can say that the difference is \( \leq \frac{1}{n} \).
# Chapter 5

## Case Study with Telephone Company

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>78</td>
</tr>
<tr>
<td>5.2 Project Motivation</td>
<td>80</td>
</tr>
<tr>
<td>5.3 Current state of Reliability Modelling at TC</td>
<td>82</td>
</tr>
<tr>
<td>5.4 Data Sets Used for Case Study</td>
<td>84</td>
</tr>
<tr>
<td>5.5 Initial Data Analysis</td>
<td>87</td>
</tr>
<tr>
<td>Local Area Data Set Analyses</td>
<td>87</td>
</tr>
<tr>
<td>Operating Area Data Set Analyses</td>
<td>90</td>
</tr>
<tr>
<td>5.6 A New Model</td>
<td>94</td>
</tr>
<tr>
<td>Basic Considerations</td>
<td>94</td>
</tr>
<tr>
<td>Identifying a New Model</td>
<td>100</td>
</tr>
<tr>
<td>New Operating Area Analysis</td>
<td>103</td>
</tr>
<tr>
<td>Square Function Results</td>
<td>103</td>
</tr>
<tr>
<td>logExp Results</td>
<td>106</td>
</tr>
<tr>
<td>New Local Area Analysis</td>
<td>109</td>
</tr>
<tr>
<td>5.7 Discussion</td>
<td>111</td>
</tr>
<tr>
<td>5.8 Decision Model</td>
<td>112</td>
</tr>
<tr>
<td>5.9 Conclusions and Future Work</td>
<td>114</td>
</tr>
<tr>
<td>References</td>
<td>116</td>
</tr>
</tbody>
</table>

Note: This case study was conducted with a network operator, Telephone Company, that wishes to remain anonymous. The information in this case study has been sanitized to protect TC’s identity. As the network operator also provided the data for a previous case study [1] conducted under the author’s academic supervisor, every effort has been made to make the terminology used in this paper consistent with that used in the previous one.
5.1 Introduction

In this chapter, we outline and summarize a project to provide Telephone Company with a statistical reliability analysis of the failure histories from part of its network. TC operates within several adjacent regions and maintains both trunk service (i.e. transmission) and local home and business services (i.e. distribution). In this work, we focus exclusively on the local service network; the circuits of this network are frequently above ground and, thus, especially prone to the problems brought on by a variety of weather conditions.

Because of its local service network’s vulnerability to the elements, TC wanted to understand and quantify the risk associated with various weather phenomena. Thus, in 2002 it initiated a project with researchers at the University of Toronto. The project was to develop a procedure for analysing the failure histories of TC’s local lines using the condition-based maintenance software EXAKT \(^1\), with each line’s local weather data being considered as the set of condition measurements. The project resulted in a thesis [1], which developed a method for preparing data for analysis in EXAKT. This involved preprocessing the weather data for efficient analysis and extracting the relevant failure histories from TC’s various maintenance and process databases in a format useable by EXAKT.

Since that time, the procedures developed in the aforementioned thesis have been used by TC for a variety of purposes and are a major part of TC’s reliability modelling and monitoring programme. Some issues raised in the course of the previous project were not fully resolved, however, and others have arisen over the past few years. Another project was initiated in 2010 after several of these unresolved problems were presented to the author by one of TC’s analysts; the project led to the case study described herein.

This chapter is organized as follows. First, we discuss the motivation for the project, the assets being analysed, and the work done by TC on the problems. For comparative

---

\(^1\) This software was developed by the University of Toronto’s Centre for Maintenance Optimization & Reliability Engineering (C-MORE), under the direction of Prof. AKS Jardine.
purposes, before presenting our analysis, we provide an overview of the current state of reliability modelling at TC. The analysis section begins with a description of the data set used in this study, including a review of the shortcomings of the set. Several analyses are performed, including a model based on work described earlier in the thesis. Finally, a benefit of the new model is demonstrated through a decision tool that uses its results as an input. The chapter concludes with a discussion of the results of the case study, potential areas for improvement, and future work.
5.2 Project Motivation

The reliable operation of its network is TC’s fundamental goal. Service interruptions are a major source of dissatisfaction to customers, and minimizing both the number and duration of such interruptions is of the utmost importance to TC. There are several causes of failure in TC’s network, among them equipment wearout and interference by animals or people, etc.; several of these failure causes are at least partially driven by weather conditions.

TC’s network has a radial structure, which means that a break in one of its local lines results in loss of service to all customers beyond the break relative to their local telephone exchange. Each local exchange is, in turn, connected a to main exchange office by a smaller trunk line. Each main exchange is connected with other main exchanges in TC’s network – as well as international exchanges and those of other telecommunications networks – by main trunk lines. Most lines in the network have protective devices that can help prevent damage to the lines in the event of dangerous behaviour somewhere on them. The network has a strongly hierarchal structure, whereby the consequences of a line’s failure increase significantly – almost geometrically – depending on where that line is in the hierarchy.

Because of the major consequences of failure for main trunk lines, TC has ensured that these lines are largely protected from the elements and thus have a high level of reliability. Not surprisingly, because the failure of local lines is less consequential, fewer capital resources have been dedicated to making them resistant to natural phenomena. But when study [1] led TC to conclude that approximately 70% of its local line interruptions were weather-related, an effort was made to model and quantify the effect of weather measurements on the risk of failure of the local lines.

This effort resulted in a 2003 thesis by Tang [1] which developed a method for incorporating weather condition measurements – provided by a government meteorological agency – into a proportional hazards model for TC’s local lines. As the data provided by the meteorological agency are given in hourly records for most measurements,
a significant amount of effort in the thesis was devoted to data reduction and determining how to process the weather data in a way that would allow efficient reliability analysis.

The data processing procedure computed daily values for statistics believed to be well correlated, or having the potential to be well correlated, with risk of failure of the circuits. These statistics included the daily average, maximum, and minimum of a measurement, as well as the daily maximum increase or decrease in a measurement. The use of daily statistics was found to be more useful than the hourly measurements provided by the meteorological agency. As a result, a tool was developed to automate the weather data processing to allow easy analysis of TC’s network histories.

In addition to the work on measurement processing, Tang’s thesis conducted several analyses of TC’s network. Using the EXAKT software package, the analyses generated proportional hazards models for the failure of lines in several operating areas of TC’s network. Analyses were conducted with a variety of combinations of failure modes tracked in TC’s maintenance database. One finding was that there were many instances of short histories in the data set – a single line failing multiple times within a short period of time – and that in most cases, these “bunches” of failures had a common initial cause, e.g. a strong storm damaging the system and causing an initial failure, with later failures due to the initial damage.

These short histories created an artefact in the PHM baseline function, causing the model to erroneously predict that the risk of failure of a line would continually decrease in time as it aged. It was found that removing from the data set all failure histories with a length of 30 days or less generally corrected this problem, and so this removal was made part of the analysis procedure proposed to TC. While the removal of short histories corrected this artefact, it also resulted in a less close fit of the data, according to Appendix B in Tang [1], but TC found this to be an acceptable trade-off.
5.3 Current state of Reliability Modelling at TC

Currently, to model the reliability of its circuits, TC follows the procedure described in Tang [1]. It receives weather data from a government meteorological agency; those data are processed with specialized software and combined with the circuit failure histories contained in TC’s maintenance records. For its analyses, TC uses EXAKT to generate PHM models for sets of local circuit lines; each set of lines is served by a common local exchange. While it is able to analyse the reliability of lines with a common main exchange, it does not do so. Rather, TC focuses on sets of lines on local exchanges in order to identify which areas of its local network may require additional capital investment or changes in maintenance procedure, as well as to identify potential improvements in methods of line construction.

TC also has data on the physical characteristics of its lines, such as their length, type of construction, and other features. In addition, it has data on the maintenance actions conducted on each line, e.g. when it was last inspected. While both of these additional sets of data would likely improve the accuracy of its models, TC has not yet incorporated them into any PHM analyses with EXAKT.

In an attempt to avoid the bunching issue identified by Tang [1], TC removes all failures from the circuit histories that occur within 30 days after a failure on a particular line, as outlined above. In several areas that have been analysed, this procedure results in an estimate of significantly less than 1.0 for the shape parameter ($\beta$) in the Weibull baseline hazard of the fitted PHM. This indicates that the lines exhibit a “break-in” effect whereby they become more reliable with age, or that the effect of a strong storm may persist for longer than the 30 day cut-off, or that the risk of failure is not being modelled correctly.

As the components on the lines are believed to exhibit wearout effects, the first possibility has been discounted. While plausible, the second possibility would require an arbitrary cut-off point to be imposed in the data processing for the model, i.e. an arbitrary removal of data. We are left with the final possibility, namely, that a better model of the time-varying behaviour of the failure risk of the lines is needed.
This bunching problem was brought to the author’s attention in a 2010 presentation. As described in the presentation, the case had the potential to use the results described earlier in this thesis. Therefore, a project was begun to analyse the failure histories of TC’s lines, with the aim of finding a better model of the time-varying aspect of the risk of failure. This offered the possibility of improving TC’s analyses by allowing it to model the true risk of service interruption to its customers, not just the risk of an initial service interruption, thereby providing a clearer picture of its customers’ experience. Additionally, a model that did not require the removal of short histories from the data set could reduce the amount of effort required by TC to produce its analyses and improve the accuracy of the estimation of the effects of weather on its lines.
Chapter 5: Case Study with Telephone Company

5.4 Data Sets Used for Case Study

This study analyses two sets of failure histories from a single operating area of the Telephone Company network. This operating area is in a region that has strong adverse weather conditions from time to time. Within the overall TC network, it is an area where the bunching problem is most apparent, and EXAKT analyses of its reliability exhibit the artefact described above. The first series of failure histories under analysis is one of the smaller local sets that TC has analysed; the second series constitutes the set of failure histories for all lines in the operating area. For the first series, one version contains all recorded events; in another version, all histories with times between failures shorter than 30 days are removed. The period under study is from 31 December 2004 to 30 December 2008.

The failure histories in the data sets consist of several events for each line under study. Every history contains at least one start time, i.e. when the unit was put in service or when the study began. Other events indicate when a unit was removed from or re-entered into service; each of these includes the time of the event. Other events include: failures, indicating both the cause and the time of the event; censor events indicating when the unit was removed from service for reasons other than failure (e.g. maintenance) or when the study ended; and “begin” events indicating when a unit was returned to service after being removed. The various failure causes found in the histories are shown below in Table 5-1.

<table>
<thead>
<tr>
<th>Cause Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled Maintenance</td>
</tr>
<tr>
<td>Lightning</td>
</tr>
<tr>
<td>Weather</td>
</tr>
<tr>
<td>Abnormal System Behaviour</td>
</tr>
<tr>
<td>Equipment Overcapacity</td>
</tr>
<tr>
<td>Tree growth</td>
</tr>
<tr>
<td>Tree fallen</td>
</tr>
<tr>
<td>Tree fallen (weather)</td>
</tr>
<tr>
<td>Vehicles</td>
</tr>
<tr>
<td>Vandalism</td>
</tr>
<tr>
<td>Customer Equipment</td>
</tr>
<tr>
<td>Other - foreign interference</td>
</tr>
<tr>
<td>Forced Outage</td>
</tr>
<tr>
<td>Human error</td>
</tr>
<tr>
<td>Equipment failure</td>
</tr>
<tr>
<td>Other (other-forced)</td>
</tr>
<tr>
<td>Wildlife</td>
</tr>
</tbody>
</table>

Table 5-1: Failure event cause codes for Telephone Company datasets.
The failure events considered in this study are those with Lightning, Weather, and Tree growth failure modes and the “Equipment Failure” mode, as TC believes these are most closely related to weather conditions. TC expects that some of these failures may be prevented with changes in maintenance strategies or construction methods. The total number of lines analysed and the number of events from the lines under study for each data set are shown below in Table 5-2.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Local Set (Including Short Histories)</th>
<th>Local Set (Excluding Short Histories)</th>
<th>Full Area Set (Including Short Histories)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Units</td>
<td>6</td>
<td>6</td>
<td>126</td>
</tr>
<tr>
<td>Number of Failures</td>
<td>14</td>
<td>12</td>
<td>104</td>
</tr>
<tr>
<td>Number of non-failure Events</td>
<td>12</td>
<td>12</td>
<td>252</td>
</tr>
<tr>
<td>Total Number of Events</td>
<td>26</td>
<td>24</td>
<td>356</td>
</tr>
</tbody>
</table>

*Table 5-2: Summary statistics for data sets under study, for events of interest.*

In addition to the failure histories, condition measurements generated from government meteorological association measurements are processed according to the procedure developed in the project [1] outlined above. Each unit history contains a daily weather measurement – treated as an inspection in EXAKT– consisting of several computed values each for pressure, temperature, wind speed and direction, etc. The values computed use statistics such as the daily maximum, the maximum increase in a measurement over the course of a day, the daily mean, etc. Table 5-3 outlines the contents of the weather records.
Chapter 5: Case Study with Telephone Company

<table>
<thead>
<tr>
<th>Element Code</th>
<th>Element Description (Variable Measured)</th>
<th>Max</th>
<th>Min</th>
<th>Avg</th>
<th>Max Jump</th>
<th>Max Drop</th>
<th>Hours Of Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>Sea Level Pressure (kilopascals)</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>Wind Speed (km/h)</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Dry Bulb Temperature (deg C)</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Thunder Storm</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Rain</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>Freezing Rain</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>Freezing Drizzle</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Snow</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>156</td>
<td>Wind Direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lightning Strikes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-3: Variables included in each weather record/inspection; a checkmark indicates that the computed statistic for each element of the record was in the data set.

The above are the only data used in this study. In addition to weather data, TC also maintains databases of process measurements on each line, and records outlining the physical parameters of each circuit. Both of these would be invaluable for this study, but TC does not have these data available in a format usable for analysis.
5.5 Initial Data Analysis

Local Area Data Set Analyses

The first analysis conducted for this study followed the procedures of Telephone Company\(^2\) in analysing the local circuit dataset to provide a base case for comparison. This analysis entailed using the “Marginal Analysis” feature of EXAKT, which simplifies the statistical reliability analysis of systems that experience failures of different components or different modes [2].

Directed by TC’s own analysis on this particular set of circuits, we produced several proportional hazards models with the standard Weibull baseline, as well as a constant baseline (i.e. \(\beta=1\) in the Weibull model) hazard ratio. This was done with short histories both included and excluded. For the data set without short histories, the estimated parameters of models with pressure, temperature, and westerly wind direction statistics as covariates are shown below in Table 5-4 (standard Weibull model) and Table 5-5 (\(\beta=1\)); the goodness of fit test results for the models are included as well.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>9822</td>
<td>1.666E+04</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>0.8352</td>
<td>0.1897</td>
<td>No</td>
<td>0.7541</td>
<td>0.3852</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.391</td>
<td>0.4781</td>
<td>Yes</td>
<td>8.468</td>
<td>0.003615</td>
</tr>
<tr>
<td>E78_MAX</td>
<td>0.4447</td>
<td>0.1018</td>
<td>Yes</td>
<td>19.08</td>
<td>0</td>
</tr>
<tr>
<td>E78_MAX_JUM</td>
<td>-1.455</td>
<td>0.5019</td>
<td>Yes</td>
<td>8.406</td>
<td>0.003741</td>
</tr>
<tr>
<td>E78_MIN</td>
<td>-0.4489</td>
<td>0.1030</td>
<td>Yes</td>
<td>18.97</td>
<td>0</td>
</tr>
<tr>
<td>HOURS_OF_W</td>
<td>-0.4493</td>
<td>0.2049</td>
<td>Yes</td>
<td>4.809</td>
<td>0.02831</td>
</tr>
</tbody>
</table>

Goodness of Fit Test

| Kolmogorov-Smirnov Statistic | 0.115063 | p-Value | 0.960683 |

Table 5-4: Estimated parameters for Weibull baseline model on local area data (short histories excluded).

---

\(^2\) These procedures are contained in an internal document of Telephone Company which is not referenced or produced for reasons of confidentiality.
Chapter 5: Case Study with Telephone Company

Parameter | Estimate | Standard Error | Significant | Wald Statistic | p-Value
---|---|---|---|---|---
Scale | 5692 | 7065 | - | - | -
Shape | 1 (Fixed) | - | - | - | -
E73_MAX_DROP | -1.351 | 0.4741 | Yes | 8.117 | 0.004386
E78_MAX | 0.4448 | 0.1011 | Yes | 19.35 | 0
E78_MAX_JUM | -1.483 | 0.5017 | Yes | 8.736 | 0.003121
E78_MIN | -0.4535 | 0.1016 | Yes | 19.94 | 0
HOURS_OF_W | -0.4505 | 0.2043 | Yes | 4.864 | 0.02742

### Goodness of Fit Test

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov Statistic</th>
<th>0.165856</th>
<th>p-Value</th>
<th>0.664476</th>
</tr>
</thead>
</table>

PHM Fits Data: Not Rejected at 5% significance level

*Table 5-5: Estimated parameters for constant baseline model on local area data (short histories excluded).*

The residuals for the model with the Weibull baseline and the transformed residuals against expectation are shown below in Figure 5-1.

*Figure 5-1: Residuals of PHM on local area data with Weibull Baseline (left); transformed residuals against expectation (right).*

The same analysis was conducted with the short histories included. As can be seen, the shape parameter is still not significant when the two short histories are included in the model, and in the Weibull model, the wind direction covariate is not significant. The estimated parameters for these models are shown below in Table 5-6 (standard Weibull model) and Table 5-7 (β=1).
Table 5-6: Estimated parameters for Weibull baseline model on local area data (short histories included).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>7305</td>
<td>1.269E+04</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>0.7207</td>
<td>0.1474</td>
<td>No</td>
<td>3.59</td>
<td>0.05812</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.039</td>
<td>0.469</td>
<td>Yes</td>
<td>4.909</td>
<td>0.02672</td>
</tr>
<tr>
<td>E78_MAX</td>
<td>0.3999</td>
<td>0.09985</td>
<td>Yes</td>
<td>16.04</td>
<td>0</td>
</tr>
<tr>
<td>E78_MAX_JUM</td>
<td>-1.263</td>
<td>0.484</td>
<td>Yes</td>
<td>6.814</td>
<td>0.009042</td>
</tr>
<tr>
<td>E78_MIN</td>
<td>-0.4065</td>
<td>0.1015</td>
<td>Yes</td>
<td>16.05</td>
<td>0</td>
</tr>
<tr>
<td>HOURS_OF_W</td>
<td>-0.2775</td>
<td>0.1434</td>
<td>No</td>
<td>3.744</td>
<td>0.05298</td>
</tr>
</tbody>
</table>

Goodness of Fit Test

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0956002</td>
<td>0.989912</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level

Table 5-7: Estimated parameters for constant baseline model on local area data (short histories included).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>3131</td>
<td>3564</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>1(Fixed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-0.9909</td>
<td>0.4653</td>
<td>Yes</td>
<td>4.535</td>
<td>0.03321</td>
</tr>
<tr>
<td>E78_MAX</td>
<td>0.4025</td>
<td>0.09781</td>
<td>Yes</td>
<td>16.93</td>
<td>0</td>
</tr>
<tr>
<td>E78_MAX_JUM</td>
<td>-1.321</td>
<td>0.4818</td>
<td>Yes</td>
<td>7.521</td>
<td>0.006097</td>
</tr>
<tr>
<td>E78_MIN</td>
<td>-0.4177</td>
<td>0.09837</td>
<td>Yes</td>
<td>18.03</td>
<td>0</td>
</tr>
<tr>
<td>HOURS_OF_W</td>
<td>-0.2827</td>
<td>0.1431</td>
<td>Yes</td>
<td>3.903</td>
<td>0.04821</td>
</tr>
</tbody>
</table>

Goodness of Fit Test

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.225708</td>
<td>0.227639</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level

Looking at the estimates in the above tables, we see that the covariate parameters do not change appreciably with the change in the baseline or with the inclusion or exclusion of short histories. While these analyses may be useful in identifying which weather conditions are correlated with the failure of lines in a local area, there is a danger of the model being overspecified, given the small number of failures.
Operating Area Data Set Analyses

The larger data set for the full operating area was analysed generally following the procedure described in TC’s manual, using the marginal analysis feature of EXAKT. Without a TC analysis to suggest which covariates would be significant for the full area data set, a procedure for determining which covariates to include in the model needed to be devised. The simple algorithm used for this purpose is as follows.

0. Fit a PHM to the data set with all covariates included.
1. Analyse PHM estimates and determine which covariate parameter estimate has the lowest Wald Test Statistic and is not significant.
2. Fit a PHM to the data set removing the covariate parameter determined in step 1.
3. Return to step 1 if there are still parameter estimates that are not significant; otherwise, stop modeling.

Using this procedure, we estimated several PHMs for the full operating area data set and fitted models with the standard Weibull baseline and a constant baseline. The covariates found to have significant estimated parameters were related to pressure, wind speed, and snow. The estimated parameters for these models are shown below in Table 5-8 (standard Weibull model) and Table 5-9 (β=1).
### Table 5-8: Estimated parameters for Weibull baseline model on operating area data (short histories included).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>5.613</td>
<td>9.126</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>0.79</td>
<td>0.07365</td>
<td>Yes</td>
<td>8.13</td>
<td>0.004353</td>
</tr>
<tr>
<td>E73_AVG</td>
<td>0.8745</td>
<td>0.16</td>
<td>Yes</td>
<td>29.89</td>
<td>0.000000</td>
</tr>
<tr>
<td>E73_MAX</td>
<td>-0.6483</td>
<td>0.1078</td>
<td>Yes</td>
<td>36.15</td>
<td>0.000000</td>
</tr>
<tr>
<td>E73_MIN</td>
<td>-0.2336</td>
<td>0.06346</td>
<td>Yes</td>
<td>13.55</td>
<td>0.0002318</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.217</td>
<td>0.2005</td>
<td>Yes</td>
<td>36.82</td>
<td>0.000000</td>
</tr>
<tr>
<td>E76_MAX_DROP</td>
<td>-0.1113</td>
<td>0.02936</td>
<td>Yes</td>
<td>14.38</td>
<td>0.0001494</td>
</tr>
<tr>
<td>E91_AVG</td>
<td>-4.878</td>
<td>1.473</td>
<td>Yes</td>
<td>10.96</td>
<td>0.0009297</td>
</tr>
<tr>
<td>E91_HOURS_OF_MAX</td>
<td>0.1761</td>
<td>0.06246</td>
<td>Yes</td>
<td>7.953</td>
<td>0.004802</td>
</tr>
<tr>
<td>E91_MAX</td>
<td>3.686</td>
<td>0.6387</td>
<td>Yes</td>
<td>33.31</td>
<td>0.000000</td>
</tr>
<tr>
<td>E91_MAX_JUMP</td>
<td>-2.172</td>
<td>0.3625</td>
<td>Yes</td>
<td>35.89</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

#### Goodness of Fit Test

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0682312</td>
<td>0.222102</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level

### Table 5-9: Estimated parameters for constant baseline model on operating area data (short histories included).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>2.153</td>
<td>2.382</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>1 (Fixed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E73_AVG</td>
<td>0.8568</td>
<td>0.1602</td>
<td>Yes</td>
<td>28.62</td>
<td>0</td>
</tr>
<tr>
<td>E73_MAX</td>
<td>-0.6379</td>
<td>0.1075</td>
<td>Yes</td>
<td>35.23</td>
<td>0</td>
</tr>
<tr>
<td>E73_MIN</td>
<td>-0.2279</td>
<td>0.06392</td>
<td>Yes</td>
<td>12.72</td>
<td>0.0003627</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.219</td>
<td>0.1989</td>
<td>Yes</td>
<td>37.59</td>
<td>0</td>
</tr>
<tr>
<td>E76_MAX_DROP</td>
<td>-0.1031</td>
<td>0.02916</td>
<td>Yes</td>
<td>12.51</td>
<td>0.0004046</td>
</tr>
<tr>
<td>E91_AVG</td>
<td>3.785</td>
<td>1.443</td>
<td>Yes</td>
<td>6.884</td>
<td>0.008713</td>
</tr>
<tr>
<td>E91_HOURS_OF_MAX</td>
<td>-0.1298</td>
<td>0.06238</td>
<td>Yes</td>
<td>4.326</td>
<td>0.03753</td>
</tr>
<tr>
<td>E91_MAX</td>
<td>3.555</td>
<td>0.6432</td>
<td>Yes</td>
<td>27.22</td>
<td>0</td>
</tr>
<tr>
<td>E91_MAX_JUMP</td>
<td>2.236</td>
<td>0.3597</td>
<td>Yes</td>
<td>38.64</td>
<td>0</td>
</tr>
</tbody>
</table>

#### Goodness of Fit Test

<table>
<thead>
<tr>
<th>Kolmogorov-Smirnov Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0855249</td>
<td>0.0634346</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level
The residuals for the operating area PHM with the Weibull baseline and the transformed residuals against expectation are given below in Figure 5-2.

![Residuals in order of appearance](image1.png) ![Transformed Ordered Residuals](image2.png)

**Figure 5-2: Residuals of PHM on operating area data with Weibull Baseline (left); transformed residuals against expectation (right).**

Having more failures to draw on, these models are able to produce more precise estimates of the parameters of interest. As in the local area case, the covariate parameters do not change appreciably with the change in the baseline hazard function from the Weibull to the constant. However, it is readily apparent that the Weibull model provides a better fit than the constant baseline: the PHM with $\beta=1$ has a p-value of less than 10%, whereas the Weibull model has a p-value of 0.222102.

A feature that is apparent in the operating area model, as compared to the local area model, is the lack of wind direction as a significant covariate. This is likely because a local area set will contain lines which generally have a common direction; over a whole area, however, the various lines go in many directions. Wind direction could likely be incorporated into an area-level analysis if the physical parameters of each line were available, but given the unavailability of these data, it is not possible to do so in the present project.

After performing analyses of the operating area data set, we determined that it would be more useful to focus modeling efforts on that data set due to the higher number of histories in the set. As the main goal of the case study was to provide an improved model of the time-varying nature of the circuits’ risk of failure, a larger number of failures would likely
help to determine the appropriate model. Although there are, on average, approximately as many failures per line in the local set as in the whole operating area set, there are many more circuits in the latter. Thus, the likelihood of finding spurious correlations or relationships would be decreased by using the larger data set.
5.6 A New Model

Basic Considerations

To produce a better model of the time varying behaviour of the failure risk of TC’s lines, we must first consider their physical structure. The lines are circuit wires strung above ground between poles, with protective equipment along each line as necessary. These lines run beside roadways or through farmland, etc., with or without trees within the lines’ rights of way. Each line connects to customers with branches that are split off in junction devices, where appropriate. The failure of any of these subsystems will cause the failure of the asset.

As shown in Section 4.2, the long run failure intensity for a linear asset that is repaired will tend to a constant average value. Given their large number of components and the long period in service of most lines in the operating area, it is reasonable to assume that the underlying age-related baseline failure intensity for the assets in the TC data set is constant. This assumption is supported by many of TC’s analyses in other operating areas, where the exclusion of short failure histories results in an estimated baseline failure rate that is not significantly different from constant. However, it is apparent from TC’s previous studies that there is a time-varying component of the failure intensity of its lines beyond that imposed by system operating conditions such as weather; thus, we must find a way of modelling this additional risk component instead of ignoring it, as occurs when short histories are excluded.

The removal of short histories also removes sources of data from the analysis. In addition to any time-varying effect that may be revealed, their inclusion may affect the estimated hazard coefficients in some way. It is therefore important to incorporate as much data as possible into the new model.

One possible set of new models is described in Section 4.3. The fundamental assumption is that the system and its subcomponents have not two possible reliability states – operating and failed – but three: operating, damaged but operating, and failed. These alternatives are
illustrated in Figure 5-3. If some event, such as adverse weather or dangerous operating conditions, damages a line, then it is reasonable to expect a greater risk of failure following that event. Estimating the time-varying behaviour of this greater, or excess, risk of failure is the aim of this new model.

One method to model the risk-of-failure behaviour of TC’s lines is using a time-dependent hazard ratio (TDHR)[3] in a proportional hazards model, with the ratio being a parametric function of the time since a major weather event was experienced by a line. Since we are assuming that the underlying failure baseline is constant and that in the long run, lines’ failure intensity returns to some “normal” level, only functions that tend to zero in the long run will be considered. (That is, only functions that result in a long run hazard ratio of one will be considered.) The form of the PHM being considered is then:

\[
\lambda(t; \bar{z}(t), \bar{T}) = \lambda_0(t) \cdot \exp[\bar{\gamma} \cdot \bar{z}(t) + \gamma_F \cdot F(T)],
\]

where \(\lambda_0(t)\), \(\bar{\gamma}\), and \(\bar{z}(t)\) have their standard definitions, and \(F(T)\) is the function used to artificially introduce a TDHR which is some function of the time \(T\) since the last major
weather event\(^3\); \(\gamma_F\) is the PHM parameter associated with the TDHR covariate. Alternatively, and more correctly given our stated purpose, we may propose that all or some covariate coefficients have time-dependent hazard ratios, \textit{viz.:}

\[
\lambda(t; \tilde{z}(t), \vec{T}) = \lambda_0(t) \cdot \exp[\tilde{\nu}(\vec{T}) \circ \tilde{z}(t)],
\]

where \(\vec{T}\) is the vector of times since events of interest. While this style of presentation makes the general principle more obvious, it does not provide a clearer interpretation of the goal of this study than the previously mentioned (5-1). To effect a practical analysis of TC’s failure data and continue to use EXAKT, several simplifying assumptions must be made. First, as noted earlier, the baseline hazard function is assumed to be constant; i.e. \(\lambda(t) = 1/\eta\), and \(\beta = 1\). Second, due to limitations in EXAKT and the difficulty of specifying such an event otherwise, a major weather event is defined as occurring at the time of any line failure. A further simplifying assumption is that a major weather event only occurs on the line that fails, even though several lines are collocated for large portions of their length. Thus, in practical terms, we are looking for a new baseline function in EXAKT of the form:

\[
\lambda_0(t) = \exp[F(T)],
\]

where \(T\) is the time since the last significant weather-related failure; it is equal to the time in the standard Weibull model used in EXAKT. While a constant baseline failure rate simplifies the work necessary in this study, it should be noted that incorporating a time-dependent hazard ratio may be useful in modelling the risk of failure of equipment in cases where the underlying baseline hazard function is non-constant.

Following the work done in [3], functional forms for \(F(T)\) we considered were the triangular function, the square (or step or delta) function, and the exponential function. In addition, we devised a form of \(F(T)\) that resulted in a hazard function of unity plus an exponential decay term and included it in the new analyses. These functions are shown mathematically below in Table 5-10, along with a small graph of their form; the resulting

\[\text{For simplicity we assume that for a line that has not encountered a major adverse weather event } T \rightarrow \infty.\]
hazard ratio \( \exp[F(T)] \) is given in Table 5-11. The two sets of plots are on the same arbitrary horizontal axes; however, the vertical axes differ, as the hazard ratio has a long run value of unity.

\[
F(T) = \begin{cases} 
1 - T/a, & T < a \\
0, & T \geq a 
\end{cases}
\]

\[
F(T) = \delta(T - a) = \begin{cases} 
1, & T < a \\
0, & T \geq a 
\end{cases}
\]

\[
F(T) = e^{-T/a}
\]

\[
F(T) = \ln \left[ 1 + b \cdot e^{-T/a} \right]
\]

*Table 5-10: Functions used to describe the time-dependent hazard ratio.*
To implement these functions in EXAKT, we used the OutputVars modelling feature [2], as it allows the failure and inspection (i.e. weather) data to be processed, output, and used as an input variable to the statistical modelling algorithms in EXAKT. The OutputVars feature outputs a new variable into EXAKT's inspections table; this variable can be used as a covariate in a PHM model. The code used to calculate the value of the function $F(T) = \exp[-T/a]$ where $a = 200$ days is shown in Figure 5-4.
//OutputVarScript
HWAge = WorkingAge - First(WorkingAge);
TDHR_Exp200 = (HN>0)*Exp(-HWAge/200)/200;

Figure 5-4: OutputVars code used to calculate TDHR function.

In the above code, the variable HWAge is the time since the line whose history is being reviewed experienced its last failure – “First(WorkingAge)” in the above code – i.e. the “History[‘s] Working Age [Since Failure]”. The next line of code calculates the TDHR variable, “TDHR_Exp200”; in this case it has been normalised so that its integral is unity. The leading term “(HN>0)” expresses the assumption that a line that has not encountered a major adverse weather event will have a hazard ratio of unity, or a value of \( F(T) \) of 0. For instance, a line experiencing a failure 200 days prior to the day of study will have a value of TDHR_Exp200 of \( 1/e \), while one that has not experienced a failure will have a value of TDHR_Exp200 of 0.
Identifying a New Model

The first stage of our new operating area analysis was to determine which functional form offers the best description of the time varying nature of the hazard/failure intensity of TC's lines following a major adverse weather event. To provide guidance on which functions were good candidates, each of the calculated TDHR variables was used singly in a PHM of the reliability of TC's lines. For each function, we estimated a PHM with the $a$ in the equations in Table 5-10 and Table 5-11 set to values between 5 days and 1000 days, with 25-day or 50-day increments between each point. For the Logarithmic function, the value of $b$ was set to 1 or 2 in this first stage of analysis.

We used the Kolmogorov-Smirnov statistic as a goodness of fit test to evaluate which functions offered an improved model. The EXAKT-calculated K-S statistics for the various TDHR functions on TC's operating area data are shown in Figure 5-5 below; the corresponding p-values appear in Figure 5-6.

![K-S Statistic for Various TDHR Functions](image)

*Figure 5-5: K-S Statistic for various TDHR functions, calculated by EXAKT from TC operating centre data.*
In a test similar to that using a standard Weibull baseline function, the Kolmogorov-Smirnov statistic was 0.0584004 with a corresponding p-value of 0.399616. All TDHR functions had a minimal K-S statistic lower than that of the Weibull baseline fit of the data. Thus, any of the TDHR functions provides a better model of the system behaviour than the current Weibull model. A decision must now be made on which functions to explore further.

Of the several functions used initially, the simplest is the square function; it is a single-variable single-operation function when implemented in EXAKT and has two simple interpretations within the PHM framework. First, it may be interpreted as a multiplicative (i.e. proportional) factor increasing the risk of failure of TC’s lines when they have recently experienced a failure, \( \text{viz.} \):

\[
\lambda(t; \bar{z}(t), \bar{T}) = \frac{1}{\eta} \cdot \exp[\gamma \cdot \delta(T - a)] \cdot \exp[\bar{y} \circ \bar{z}(t)].
\]
Second, it may be interpreted as an “excess” risk of failure due to the effects of some previous adverse weather event, with the excess risk proportional to weather in the same manner as the baseline risk of failure:

\[
\lambda(t; \bar{z}(t), \bar{T}) = \frac{1}{\eta} \cdot [1 + \delta(T - a) \cdot (e^{y \bar{T}} - 1)] \cdot \exp[\bar{y} \circ \bar{z}(t)].
\]

These interpretations are equivalent but may be useful in different ways.

The most interesting function from a theoretical perspective, as described in Section 4.3, is the logarithm of the sum of unity and an exponential term (from here on referred to as the “logExp” function). If this is expanded, we see that it naturally has an interpretation similar to the second one for the square function above:

\[
\lambda(t; \bar{z}(t), \bar{T}) = \frac{1}{\eta} \cdot \left[1 + b \cdot e^{-T/a}\right]^{\gamma} \cdot \exp[\bar{y} \circ \bar{z}(t)],
\]

where the exponential term describes the “excess” risk of failure due to weather-caused damage.

The two additional possible candidates for describing the time-varying nature of the failure risk of TC’s lines, the exponential function and the triangular function, offer the possibility of better model than the current Weibull. However, we continued our investigation using the square function due to its simplicity and using the logExp function because of its interesting theoretical properties. The other candidates offer no compelling alternative advantages.
New Operating Area Analysis

Using the standard EXAKT analysis procedure described earlier, we conducted a new operating area analysis with the TDHR covariates used to model the after-effects of adverse weather phenomena. We generated a full PHM of the data, including short histories, with the weather covariates found to be significant in the analysis described in Section 5.5. To determine the optimal parameters for the TDHR functions, as measured by goodness of fit test, we used a grid search. While it would have been more accurate to use a log-likelihood test to determine the optimal parameter estimates, it is not practical to do so with EXAKT.

For the square function, we estimated PHMs with the value of $a$ varied in increments of 25 days, over a range of 100 to 500 days, in order to find the optimum. Near the optimum, the value of $a$ was again varied, but with smaller increments of 5 days. For the logExp function, we estimated PHMs with the values of $a$ varied in the same manner; the value of $b$ was set to 1.0, 1.5, 2.0, 2.5, and 3.0 in order to find an appropriate value for that parameter. While this method of determining the appropriate values of $a$ and $b$ is not very precise, it is straightforward to implement and thus can be easily applied by TC.

Square Function Results

Using the procedure outlined above, the PHM that best fit the TC failure data was found to be a square function with a value of $a$ of 350 days. The parameters for this model are shown below in Table 5-12, and the residuals for the PHM and the transformed residuals against expectation appear in Figure 5-7. The new model provides a significantly better fit to the data than does the Weibull baseline function, with a K-S statistic of 0.0381209 and a corresponding p-value of 0.884441 compared to 0.0682312 and 0.222102 respectively for the Weibull. Additionally, the Wald statistic for the TDHR covariate is 24.36, suggesting it is more significant than the shape parameter in the Weibull function (which has a value of 8.13). If we consider the residuals, those for the new model are closer to expectations than the old one, again suggesting an improved fit.
Parameter | Estimate | Standard Error | Significant | Wald Statistic | p-Value
---|---|---|---|---|---
Scale | 6.84 | 8.41 | - | - | -
Shape | 1 (Fixed) | - | - | - | -
E73_AVG | 0.8375 | 0.1587 | Yes | 27.85 | 0
E73_MAX | -0.6275 | 0.107 | Yes | 34.4 | 0
E73_MIN | -0.2185 | 0.06345 | Yes | 11.86 | 0.000574
E73_MAX_DROP | -1.233 | 0.2027 | Yes | 36.99 | 0
E76_MAX_DROP | -0.1101 | 0.02902 | Yes | 14.38 | 0.0001495
E91_AVG | -4.385 | 1.414 | Yes | 9.62 | 0.001925
E91_HOURS_OF_MAX | 0.1607 | 0.06078 | Yes | 6.994 | 0.008178
E91_MAX | 3.563 | 0.6259 | Yes | 32.4 | 0
E91_MAX_JUMP | -2.22 | 0.3638 | Yes | 37.25 | 0
TDHR Parameter
TDHR_sq350 | 0.9964 | 0.2019 | Yes | 24.36 | 0

Goodness of Fit Test

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov Statistic</td>
<td>0.0381209</td>
</tr>
<tr>
<td>p-Value</td>
<td>0.884441</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level

Table 5-12: Estimated parameters for model on operating area data with square function (a=350 days) as TDHR function (short histories included).

Figure 5-7: Residuals of PHM on operating area data with square function (a=350 days) as TDHR function (left); transformed residuals against expectation (right).

We performed an additional test to determine the effectiveness of the TDHR approach to modelling the time-varying nature of the risk of failure of TC's lines: more specifically, we estimated a PHM that included both the TDHR function and the Weibull baseline. The results of this procedure are shown below in Table 5-13, and the residuals displayed in
Chapter 5: Case Study with Telephone Company

Figure 5-8. The estimated shape parameter is very nearly unity, suggesting that the TDHR function provides a good model of the time-varying behaviour of the risk of failure.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>6.785</td>
<td>8.23</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>1.003</td>
<td>0.1024</td>
<td>No</td>
<td>0.0008676</td>
<td>0.9765</td>
</tr>
<tr>
<td>E73_AVG</td>
<td>0.8383</td>
<td>0.159</td>
<td>Yes</td>
<td>27.79</td>
<td>0</td>
</tr>
<tr>
<td>E73_MAX</td>
<td>-0.628</td>
<td>0.1072</td>
<td>Yes</td>
<td>34.34</td>
<td>0</td>
</tr>
<tr>
<td>E73_MIN</td>
<td>-0.2188</td>
<td>0.06353</td>
<td>Yes</td>
<td>11.86</td>
<td>0.0005727</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.233</td>
<td>0.2027</td>
<td>Yes</td>
<td>36.98</td>
<td>0</td>
</tr>
<tr>
<td>E76_MAX_DROP</td>
<td>-0.1099</td>
<td>0.02912</td>
<td>Yes</td>
<td>14.26</td>
<td>0.0001592</td>
</tr>
<tr>
<td>E91_AVG</td>
<td>-4.372</td>
<td>1.468</td>
<td>Yes</td>
<td>8.871</td>
<td>0.002898</td>
</tr>
<tr>
<td>E91_HOURS_OF_MAX</td>
<td>0.1602</td>
<td>0.06302</td>
<td>Yes</td>
<td>6.459</td>
<td>0.01104</td>
</tr>
<tr>
<td>E91_MAX</td>
<td>3.559</td>
<td>0.6361</td>
<td>Yes</td>
<td>31.3</td>
<td>0</td>
</tr>
<tr>
<td>E91_MAX_JUMP</td>
<td>-2.221</td>
<td>0.365</td>
<td>Yes</td>
<td>37.02</td>
<td>0</td>
</tr>
<tr>
<td>TDHR Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDHR_sq350</td>
<td>1</td>
<td>0.2592</td>
<td>Yes</td>
<td>14.9</td>
<td>0.0001136</td>
</tr>
</tbody>
</table>

Goodness of Fit Test

<table>
<thead>
<tr>
<th></th>
<th>Kolmogorov-Smirnov Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHM</td>
<td>0.0372621</td>
<td>0.900112</td>
</tr>
</tbody>
</table>

PHM Fits Data: Not Rejected at 5% significance level

Table 5-13: Estimated parameters for model on operating area data with square function (α=350 days) as TDHR function and Weibull function as baseline (short histories included).

Figure 5-8: Residuals of PHM on operating area data with square function (α=350 days) as TDHR function and Weibull function as baseline (left); transformed residuals against expectation (right).
logExp Results

Using the procedure outlined above, we found the PHM with the logExp function that best fit the TC failure data had an $a$ value of 275 days and a $b$ value of 2.5. The parameters for this model are given below in Table 5-14, and the residuals for the PHM and the transformed residuals against expectation appear in Figure 5-9. This version of the new model also provides a significantly better fit to the data than does the Weibull baseline function, with a K-S statistic of 0.0500976 and a corresponding p-value of 0.597314 for the former and 0.0682312 and 0.222102 respectively for the latter. Additionally, the Wald statistic for the TDHR covariate is 18.06, suggesting it is more significant than the shape parameter in the Weibull function (which has a value of 8.13, as indicated in Table 5-8).

If we consider the residuals, those for the new model are closer to expectations than the old one, again suggesting an improved fit.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>9.103</td>
<td>11.18</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>1 (Fixed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E73_AVG</td>
<td>0.8608</td>
<td>0.159</td>
<td>Yes</td>
<td>29.31</td>
<td>0</td>
</tr>
<tr>
<td>E73_MAX</td>
<td>-0.6408</td>
<td>0.1072</td>
<td>Yes</td>
<td>35.75</td>
<td>0</td>
</tr>
<tr>
<td>E73_MIN</td>
<td>-0.2284</td>
<td>0.06329</td>
<td>Yes</td>
<td>13.02</td>
<td>0.0003082</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.229</td>
<td>0.2003</td>
<td>Yes</td>
<td>37.66</td>
<td>0</td>
</tr>
<tr>
<td>E76_MAX_DROP</td>
<td>-0.1086</td>
<td>0.02913</td>
<td>Yes</td>
<td>13.9</td>
<td>0.0001925</td>
</tr>
<tr>
<td>E91_AVG</td>
<td>-4.681</td>
<td>1.438</td>
<td>Yes</td>
<td>10.6</td>
<td>0.001132</td>
</tr>
<tr>
<td>E91_HOURS_OF_MAX</td>
<td>0.1731</td>
<td>0.06157</td>
<td>Yes</td>
<td>7.9</td>
<td>0.004943</td>
</tr>
<tr>
<td>E91_MAX</td>
<td>3.638</td>
<td>0.6302</td>
<td>Yes</td>
<td>33.33</td>
<td>0</td>
</tr>
<tr>
<td>E91_MAX_JUMP</td>
<td>-2.178</td>
<td>0.3611</td>
<td>Yes</td>
<td>36.39</td>
<td>0</td>
</tr>
<tr>
<td>TDHR Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDHR_In25Exp275</td>
<td>0.9937</td>
<td>0.2338</td>
<td>Yes</td>
<td>18.06</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 5-14: Estimated parameters for model on operating area data with logExp function ($a=275$ days, $b=2.5$) as TDHR function as TDHR function (short histories included).**
Chapter 5: Case Study with Telephone Company

Figure 5-9: Residuals of PHM on operating area data with logExp function (a=275 days, b=2.5) as TDHR function (left); transformed residuals against expectation (right).

Again, an additional test was done to determine the effectiveness of the TDHR approach in modelling the time-varying nature of the risk of failure of TC’s lines, estimating a PHM that included both the logExp TDHR function and the Weibull baseline. The results are shown below in Table 5-15 and the residuals displayed in Figure 5-10. Once again, the estimated shape parameter is very nearly unity.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>9.764</td>
<td>12.65</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>0.9692</td>
<td>0.1163</td>
<td>No</td>
<td>0.07039</td>
<td>0.7908</td>
</tr>
<tr>
<td>E73_AVG</td>
<td>0.8729</td>
<td>0.1602</td>
<td>Yes</td>
<td>29.69</td>
<td>0</td>
</tr>
<tr>
<td>E73_MAX</td>
<td>-0.6488</td>
<td>0.108</td>
<td>Yes</td>
<td>36.12</td>
<td>0</td>
</tr>
<tr>
<td>E73_MIN</td>
<td>-0.2321</td>
<td>0.06366</td>
<td>Yes</td>
<td>13.3</td>
<td>0.0002656</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.22</td>
<td>0.2009</td>
<td>Yes</td>
<td>36.89</td>
<td>0</td>
</tr>
<tr>
<td>E76_MAX_DROP</td>
<td>-0.1098</td>
<td>0.0293</td>
<td>Yes</td>
<td>14.03</td>
<td>0.0001798</td>
</tr>
<tr>
<td>E91_AVG</td>
<td>-4.795</td>
<td>1.464</td>
<td>Yes</td>
<td>10.72</td>
<td>0.00106</td>
</tr>
<tr>
<td>E91_HOURS_OF_MAX</td>
<td>0.1755</td>
<td>0.06236</td>
<td>Yes</td>
<td>7.92</td>
<td>0.004889</td>
</tr>
<tr>
<td>E91_MAX</td>
<td>3.657</td>
<td>0.6331</td>
<td>Yes</td>
<td>33.36</td>
<td>0</td>
</tr>
<tr>
<td>E91_MAX_JUMP</td>
<td>-2.17</td>
<td>0.3611</td>
<td>Yes</td>
<td>36.11</td>
<td>0</td>
</tr>
<tr>
<td>TDHR Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDHR_In25Exp275</td>
<td>1.152</td>
<td>0.504</td>
<td>Yes</td>
<td>5.222</td>
<td>0.0223</td>
</tr>
</tbody>
</table>

Goodness of Fit Test

| Kolmogorov-Smirnov Statistic | 0.0576259 | p-Value | 0.416382 |

Table 5-15: Estimated parameters for model on operating area data with logExp function (a=275 days, b=2.5) as TDHR function and Weibull function as baseline (short histories included).
Figure 5-10: Residuals of PHM on operating area data with logExp function (a=275 days, b=2.5) as TDHR function and Weibull function as baseline (left); transformed residuals against expectation (right).

However, the model with the logExp function does not offer as good a fit as the model with the square function. This is clear when we compare the K-S statistic difference; it can be also be seen graphically by comparing the ordered residuals from the two model fits in Figure 5-9 and Figure 5-7.
New Local Area Analysis

While there are few failures in the local area data set and, hence, concerns about model overspecification, we thought it would be a valuable exercise to compare the results of one of the new models with results produced by the current procedure used by TC. Therefore, based on the results from the full operating area modeling, we worked on estimating a new model for the local data set using the logExp function to model the time-dependent hazard ratio between lines with and without a recent weather-related failure.

The initial model used the covariates found to be significant using TC’s procedure. The grid search procedure outlined in the previous section was used to determine the best-fitting model, or at least a local optimum. The estimated parameters of the best model fit to the data are shown below in Table 5-16; the residuals appear in Figure 5-11. The optimum logExp function calculated used $a=170$ days and $b=1$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significant</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>6281</td>
<td>7436</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shape</td>
<td>1 (Fixed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E73_MAX_DROP</td>
<td>-1.045</td>
<td>0.4612</td>
<td>Yes</td>
<td>5.137</td>
<td>0.02342</td>
</tr>
<tr>
<td>E78_MAX</td>
<td>0.4038</td>
<td>0.09588</td>
<td>Yes</td>
<td>17.74</td>
<td>0</td>
</tr>
<tr>
<td>E78_MAX_JUM</td>
<td>-1.291</td>
<td>0.4711</td>
<td>Yes</td>
<td>7.512</td>
<td>0.006128</td>
</tr>
<tr>
<td>E78_MIN</td>
<td>-0.4222</td>
<td>0.09804</td>
<td>Yes</td>
<td>18.55</td>
<td>0</td>
</tr>
<tr>
<td>HOURS_OF_W</td>
<td>-0.3009</td>
<td>0.142</td>
<td>Yes</td>
<td>4.492</td>
<td>0.03405</td>
</tr>
<tr>
<td>TDHR_lnExp170</td>
<td>3.344</td>
<td>1.009</td>
<td>Yes</td>
<td>10.99</td>
<td>0.0009136</td>
</tr>
</tbody>
</table>

**Goodness of Fit Test**

| Kolmogorov-Smirnov Statistic | 0.0932691 | p-Value | 0.992495 |

**Table 5-16: Estimated parameters for model with new baseline on operating area data (short histories included).**
Compared with the standard analysis of the data which excludes short histories, the new model provides a lower K-S statistic of 0.0932691 vs. 0.115063, and a correspondingly higher p-value of 0.992495 compared to 0.960683. When compared to the constant-baseline model without short histories, the new model provides a significant improvement in the K-S statistic (0.165856 for the old vs. 0.0932691 for the new) and p-value (0.664476 vs. 0.992495). However, given the small number of events and relatively large number of parameters in the model, these comparisons do not make a strong case for using one model over another.
5.7 Discussion

The new EXAKT model that approximates the one recommended earlier in this thesis provides a significantly better predication of the risk of failure of TC's local circuits, as measured by goodness of fit tests. For the full operating area models, when both the artificial hazard ratio covariate and the shape parameter were included, the estimated shape parameter in the model was not significantly different than unity – suggesting that the time-varying nature of the line failure intensity was captured better by the new model than by a Weibull baseline. The new model can also incorporate all of the failure histories for a set of lines and correctly model the long run risk of failure of the lines—i.e. the underlying baseline risk of failure should tend to a constant average in the long run, and the hazard of a circuit should only depend on the conditions the line experiences.

Furthermore, as the model incorporates TC's short failure histories, they are correctly modelled as a result of excess hazard due to prior events—failures that could potentially be prevented by a change in maintenance strategies—and not simply as an extension of a single interruption. As this more accurately represents the experience of TC's customers, it may improve customer-focused reliability and maintenance strategies.

Finally, we observe that one of the improved new models including a TDHR, specifically the one with the logExp function, models the circuit failure behaviour as approximately Markovian. This observation may be useful in future explorations of the failure behaviour of TC's circuits following storms or other adverse weather conditions.
5.8 Decision Model

With the new reliability model for the circuits, more information is known about the system’s behaviour following a weather-related failure. Using this additional information, it may be possible to justify certain maintenance decisions. As a simple example, let us assume there is a perfect inspection procedure available for the lines being studied. This inspection procedure could detect any system components that are in a high-risk or damaged state and prevent future failure due to this damage. Thus, if we inspect a circuit after it has experienced a weather-related failure, we can eliminate these future potential failures.

Based on our new model, the cumulative “excess” failure intensity\(^4\) from subcomponents in a damaged state on a line immediately after a failure as compared to a line that has not had a recent failure is

\[
\int_{t=t_E}^{\infty} \frac{1}{\eta} e^{\bar{y}_o(t)} \delta(t - a) \cdot (e^{y_F} - 1) \, dt
\]

in the case of the square function, where \(t_E\) is the time of the last weather-related failure event, and approximately

\[
\int_{t=t_E}^{\infty} \frac{1}{\eta} e^{\bar{y}_o(t)} a \cdot e^{-(t-t_E)/b} \, dt
\]

in the case of the logExp function. Using the median value of the covariates, and the parameters for \(\bar{y} \), \(a\) and \(b\) found with EXAKT for the operating area model, the above integrals predict an estimated 0.24 excess failures (in the case of the square function) or 0.26 excess failures (in the case of the logExp function) would be associated with damage due to earlier adverse weather incidents. These failures might be avoided if a perfect

\[^{4}\text{Here we make the assumption that the hazard function for a line estimated by EXAKT is approximately equal to its failure intensity function.}\]
inspection occurred after a weather-related failure. The line, therefore, should be inspected if it costs less to do so than approximately one quarter of the cost of experiencing another failure.

Improved calculation of the above integrals can be made by incorporating EXAKT's probabilistic predictions of the time evolution of the covariates following a failure. However, for reasons of brevity, these calculations are not included here.

In the above cost model, we see that with more information incorporated into our model of the risk of failure of the circuits, more evidence can be brought to bear on maintenance and asset management decisions. This provides managers and maintainers with an improved ability to decide on maintenance strategy.
5.9 Conclusions and Future Work

This case study has found a model that gives a better fit to TC's failure data and provides more useful information about its lines than the current Weibull model. It should allow the simplification of the reliability analyses performed by Telephone Company by obviating the need to clean the data of short histories. Additionally, this study demonstrates that considering the physical nature of a system can lead to insights into the creation of more useful reliability models.

To extend this project, EXAKT's full transition model could be used to calculate a more correct estimate of the total “Excess Risk” to a line after a weather-related failure. This would likely improve the accuracy of the risk of failure estimate, as the calculation using the average values of the covariates probably underestimates the expected number of failures; generally speaking, shortly after a storm, the composite covariate slowly tends towards the average. Alternatively, using weather forecasts to make the calculation instead of the transition probability model in EXAKT may be beneficial.

In the longer term, it may be beneficial to modify EXAKT. While the results are useful, the process used with EXAKT to optimize the new model is tedious and potentially inaccurate. The Weibull model is highly flexible, though not always appropriate; it would be useful to be able to specify a different model or perhaps use the standard nonparametric Cox model. Giving EXAKT the flexibility of permitting data from individual histories to affect system-wide calculated covariates would also be valuable.

This study also raises questions about the proposed models in and of themselves. For instance:

- Could the time parameter $a$ in the new time-dependent hazard ratio covariate be related to inspection procedures? Discussing this issue with TC's maintenance managers, the tentative answer would be “no”, as lines are inspected every 3 years in urban locations and every 6 years in rural ones. This question may require more detailed exploration.
- Could this \( a \) parameter depend on weather covariates? Could the parameter perhaps depend on physical system parameters, e.g. length or right-of-way type?
- Could the scale parameter \( b \) in the logExp model be related to the number of subcomponents damaged in a storm? Or to system parameters?

There are questions about the assumptions involved in the model as well. For instance, are the “damaged” systems’ failure rates actually proportional to those of the “normal” subcomponents? Additional statistical analysis of the model, which is not practicable with EXAKT, would be important in answering these questions.

Finally, the model could be improved by including physical information about the lines, even though without this improvement the model may still provide a useable decision model. Future work should be done to incorporate physical parameters of the lines under study – especially their length – into the reliability model. This would likely improve the accuracy of TC’s hazard / failure intensity estimates for its lines.

Furthermore, with careful use of categorical and interaction variables to describe all of the system’s physical and operational parameters, it may be possible to produce a comprehensive model of the reliability of all lines in TC’s local service networks. This would allow a much larger pool of system histories to be used and could improve the precision of the various model parameter estimates. This type of analysis could more accurately estimate the risk of failure of lines of alternative construction or those lines with differing operating parameters. Information of this nature could allow for improved evaluation of the benefits of additional capital expenditure or changes in maintenance strategies or methods and, consequently, permit further economic optimization of TC’s network.
References


Chapter 6
Summary & Conclusions

6.1 Thesis Contributions........................................................................................................120

Guidelines for Building Simple Reliability Models of Linear Assets............................121

6.2 Suggestions for Future Work.............................................................................................123

Suggestions for Reliability Engineering Theorists.........................................................123

Suggestions for TC’s Reliability Engineers.....................................................................124

Suggestions for the EXAKT Software’s Developers.......................................................125
6.1 Thesis Contributions

The safe and reliable operation of linear assets is of the utmost importance to managers, maintainers, and engineers. This is greatly facilitated by knowledge of such assets’ failure behaviour and accurate estimates of their reliability. This thesis has suggested an idealised model of a linear asset and with it explains several reliability features of linear assets. Study of this ideal model permits us to make several suggestions on how to estimate the reliability of linear assets accurately using the proportional hazards model or proportional intensity model and how the use of this model may support a variety of maintenance decisions. These guidelines are repeated below for reference.

The simple model presented in this paper may be extended to permit analysis of more complicated phenomena observed in the failure behaviour of linear assets. Such features may be modelled within the PHM/PIM framework by the inclusion of time-dependent hazard ratios or other modelling techniques. The utility of one of these model extensions has been demonstrated in a case study with TC; as shown, an approximation of this new model in the EXAKT software package provides a better estimate the risk of failure of TC’s lines than does the present method of analysis. The new method may also benefit TC by simplifying its reliability analysis procedures and providing more useful information about the failure behaviour of its lines.
Guidelines for Building Simple Reliability Models of Linear Assets

Statement 1: In the long run, a linear asset which has its components renewed upon failure will experience a failure intensity that tends to a constant average. Thus the asset’s failure rate may be modelled as the sum of the average failure rate of all of its components.

Statement 2: Component ageing significantly affects the failure intensity of a linear asset, and its effect must be included in any reliability model of a system that is not very aged (i.e. an asset where Statement 1 does not apply). The effect of component replacement—as a result of a failure or otherwise—on the failure intensity of a linear asset depends on several factors, in particular the time-in-service of the asset, the number of failures it has previously experienced, and the number of components in the asset. Often the assumption of minimal repair may be made without great loss of accuracy for systems that experience very few failures over their service lives. It may be useful, however, to use different models in estimating the failure intensity of a linear asset over certain parts of its life; certain reliability models (e.g. the virtual age model) may be considered for use over the whole life of the asset.

Statement 3: The failure rate of a linear asset is strongly dependent on the number of components that make up the asset; the number of components in a linear asset is typically strongly affected by its length. When modeling the reliability of a population of linear assets, it is necessary to model the proportionality of the failure rate to the number of components; the lengths of the lines are often useful proxy measurements for the number of certain components within the linear assets; however, it is most accurate to explicitly use the number of components rather than the length.

Statement 4: Where a population of linear asset is composed of lines with segments of differing characteristics (e.g. different construction methods, right-of-way types, or materials), it is often useful to model their differing failure rates using a PHM/PIM with a set of binary covariates. If these characteristics are not uniform throughout a single line being modelled, then the use of fractional values in place of the binary covariates, with the
fractions the proportion of each line with the appropriate characteristic, may be effective in modelling their differing reliability with some accuracy.

Statement 5: A reliability model of a population of linear assets should attempt to incorporate the effect on failure rate of internal process measurements (e.g. rate of flow, current, etc.) and conditions (e.g. inspection data), as well as external conditions such as weather. The PHM/PIM offers a useful method for doing so, though there are alternative models that might be effectively used. As long as these conditions do not vary significantly over the length of the asset, they may be practicable; if they do vary significantly, Statement 4 may be useful.

Statement 6: When deciding on the virtual sectioning of a linear asset for reliability analysis purposes, the static conditions and method of construction of the line’s segments and component parts should be considered. It may also be beneficial to model separately portions of a line in which segments of one type are more concentrated than another type, or certain external conditions, such as weather, predominate.
6.2 Suggestions for Future Work

Suggestions for Reliability Engineering Theorists

The simple model of an idealised linear asset presented in Section 4.2 can provide useful insight into the failure behaviour of linear assets, leading to better suggestions for modeling their failure rates. As shown in Section 4.3, extending this model with a three-state component/segment model can provide a theoretic base with which to understand more complicated linear asset behaviour, even with restrictive assumptions. Future work could focus on more advanced study of the three-state model and on determining how to present and estimate such a model for use in reliability analysis and decision-making.

Additionally, expanded use of time-dependent hazard ratios in modeling complex phenomena in system failure behaviour could benefit many reliability analysts, whether they are modeling linear assets or point assets. The use of TDHRs permits continued use of the already-advanced PHM/PIM modelling framework while giving analysts the ability to handle more complex system behaviour. The use of such methods should be explored by reliability engineers for modeling complex system failure behaviour.

As was noted briefly in section 4.2, a linear asset shares many of the reliability features of fleets of equipment. Thus, it may be worthwhile to explore the possible application of analysis methods developed for fleets to linear assets, and *vice versa*.

Finally, the improved integration of physics-based wear and reliability models with statistically-based models would be very useful to managers of linear assets and reliability engineers in general. For instance, the baseline failure rate of a pipeline may be strongly based on the expected wear rate of such a pipe under certain flow conditions. The inspection of this pipe may permit a more accurate gauge of its true “age” and hence allow for a more accurate reliability estimate. Such linkages between observation-based and statistically-based reliability models offer great potential to improve reliability forecasts and should be explored.
Chapter 6: Summary & Conclusions

Suggestions for TC’s Reliability Engineers

The TC case study found a model with a better fit to TC’s failure data, one that provides more useful information about its lines than the current Weibull model. It should simplify the reliability analyses performed by TC by obviating the need to clean the data of short histories. There are several potential extensions of this project.

For one thing, EXAKT’s full transition model could be used to calculate a more correct estimate of the total “Excess Risk” to a line after a weather-related failure. This would likely improve the accuracy of the risk of failure estimate, as the calculation using the average values of the covariates probably underestimates the expected number of failures. Alternatively, using weather forecasts to make the calculation instead of the transition probability model in EXAKT may improve the usefulness of TC’s model, by e.g. improving maintenance workforce needs forecasts.

The case study also raises questions about the proposed models in and of themselves. For instance:

- Could the time parameter $a$ in the new time-dependent hazard ratio covariate be related to inspection procedures? After discussing this issue with TC’s maintenance managers, our tentative answer is “no”, as lines are inspected every 3 years in urban locations and every 6 years in rural ones. This question may require more detailed exploration.

- Could this $a$ parameter depend on weather covariates? Could the parameter perhaps depend on physical system parameters, e.g. length or right-of-way type?

- Could the scale parameter $b$ in the logExp model be related to the number of subcomponents damaged in a storm? Or to system parameters?

There are questions about the model’s assumptions as well. For instance, are the “damaged” systems’ failure rates actually proportional to those of the “normal” subcomponents? Additional statistical analysis of the model, which is not practicable with EXAKT, would be important in answering these questions.
Furthermore, the model could be improved by including physical information about the lines, though even without this improvement, the model may still provide a useable decision model. Future work should incorporate physical parameters of the lines under study—especially their length—into the reliability model. This would likely improve the accuracy of TC's hazard / failure intensity estimates for its lines.

Finally, with careful use of categorical and interaction variables to describe all of the system's physical and operational parameters, it may be possible to produce a single comprehensive model of the reliability of all lines in TC's local service networks. This would allow a much larger pool of system histories to be used and could improve the precision of the various model parameter estimates. Arguably, this type of analysis would more accurately estimate the risk of failure of lines of alternative construction or lines with differing operating parameters. Information of this nature could allow for improved evaluation of the benefits of additional capital expenditure or changes in maintenance strategies or methods and, consequently, permit the further economic optimization of TC's network.

**Suggestions for the EXAKT Software’s Developers**

In the longer term, it may be necessary to modify EXAKT. While the results obtained in the case study with TC are useful, the process used with EXAKT to optimize the new model is tedious and potentially inaccurate. The Weibull model is highly flexible, though not always appropriate for an asset; it would be useful to many of the software’s users to be able to specify a different baseline model, or perhaps use the standard nonparametric Cox model. Additionally, giving EXAKT the flexibility of permitting data from individual histories to affect system-wide calculated covariates would be invaluable.
Bibliography


