Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.
Bayesian Modelling of Short Fatigue Crack Growth and Coalescence

Cathal Dominic Walsh

Thesis submitted for the degree of Doctor of Philosophy

Trinity College Dublin
Department of Statistics
October 1999
Declaration

This thesis is entirely my own work. The contents of this thesis have not been submitted as an exercise for a degree at this or any other University. The contents of this thesis may be lent or copied by the Library at Trinity College Dublin.

Cathal Dominic Walsh
Summary

Failure of metal structures is caused by cracks appearing and growing in the material until the strength of the structure is compromised. The way in which such cracks grow in metal has been researched extensively; the great majority of this work concentrates on long cracks, that is cracks of the order of 1mm or longer.

This research explores new ways of modelling reliability by developing a model for short crack propagation and interactions in steel. Short crack propagation is affected by the microstructural properties of the material, and differs significantly from the growth of long cracks. Interactions consist of cracks growing into one other to form longer cracks. This is an important mechanism for damage accumulation, and is termed coalescence. During a substantial portion of useful structural lifetime, short cracks are present in the material. These form and grow under everyday stresses that are experienced by the structure. For this reason, predictions as to when the length of the cracks will compromise the strength of the structure are of interest.

Data on the growth and coalescence of short cracks has been provided by the Department of Mechanical and Manufacturing Engineering, Trinity College Dublin.

A model is developed for short crack propagation which takes into account
the microstructural features of the material. This gives reasonable reliability predictions when coalescence is not a dominant feature. A further model is developed for coalescence. This is combined with the model for growth, using auxiliary variables for unobserved parameters of interest.

The analysis is carried through in a Bayesian framework. Markov Chain Monte Carlo (MCMC) techniques are used to sample from the posterior distribution of the parameters of the model.

Recommendations are made as to the nature of the spatial data that may be easily recorded in the future, in order that reliability predictions may be improved.
Acknowledgements

This research has been carried out with the support of a Forbairt Basic Research Grant SC/96/626, under the supervision of Dr Simon P. Wilson of the Department of Statistics and with the collaboration of Professor David Taylor from the Department of Mechanical and Manufacturing Engineering, Trinity College Dublin.

I would like to express my thanks to Dr Simon Wilson, for his guidance and encouragement throughout this research. I am grateful to Professor David Taylor, Dr Brendan McCormack and Dr Patrick Prendergast for their collaboration and discussion of the engineering principles behind damage accumulation in steel and PMMA. The academic staff in the Department of Statistics have been of tremendous assistance and support over the years; especially Professor John Haslett who has contributed ideas from the earliest stages of the research, and Eamonn Mullins whose broad outlook on applied statistical inference has been invaluable. A special mention goes to Aideen, Michael, Carl, Cassian, Clare, Dina, Rakhi and Ronan who have influenced my work in various ways. I am thankful to my family for all their support over the years; especially my Dad, Bernard for the proof reading and comments. Finally, my deepest gratitude goes to my wife, Sam, for her encouragement and support, at every stage of this work.
Contents

1 Introduction 1
   1.1 Review of Current Models .............................................................. 2
   1.2 Statistical Methodology ................................................................. 3
   1.3 Growth Model .................................................................................... 3
   1.4 Coalescence ........................................................................................ 4
   1.5 Joint Model ........................................................................................ 4
   1.6 Overall Framework and Major Contributions ..................................... 5

2 Modelling Fatigue and Crack Propagation ........................................... 6
   2.1 Factors Important to Modelling Fatigue ........................................ 7
      2.1.1 Specimen Geometry ............................................................ 7
      2.1.2 Material Properties ............................................................... 7
      2.1.3 Nature of Loading ............................................................... 8
      2.1.4 Statistics of Interest ............................................................ 9
      2.1.5 Manufacturing Issues ............................................................ 10
   2.2 Deterministic Models for Crack Growth ........................................ 11
   2.3 Wöhler or S–N Curve ....................................................................... 12
   2.4 Empirical Models .............................................................................. 12
      2.4.1 Paris-Erdogan Equation ..................................................... 13
2.4.2 Forman Equation ................................................. 13

2.5 Theoretical Models .................................................. 15
2.5.1 Cumulative Damage Theories ................................. 15
2.5.2 Elastic Fatigue Fracture ......................................... 16
2.5.3 J Integral Methods .............................................. 17
2.5.4 Empirical Stochastic Parameterisation ...................... 17
2.5.5 Stochastic Models ............................................. 18

2.6 Fatigue Crack Growth Process .................................. 19

2.7 Current Research Questions ..................................... 20
2.7.1 Growth Aspect .................................................. 21
2.7.2 Coalescence Aspect .......................................... 22
2.7.3 Other Considerations .......................................... 22

3 Statistical Methodology ............................................ 23
3.1 Statistics .............................................................. 23
3.2 Bayesian Approach ................................................ 24
3.2.1 Formal Bayesian Methodology .............................. 25
3.2.2 Predictive Distribution ....................................... 29
3.2.3 Kernel Density Estimation ................................. 29
3.2.4 A Simple Example – \( N(\mu, \frac{1}{\sigma^2}) \) .................. 30
3.2.5 Prior Elicitation and Non-informative Priors .......... 32

3.3 Sampling from the Posterior Distribution .................... 33
3.3.1 Stratified Sampling .......................................... 34
3.3.2 Importance Sampling ....................................... 35
3.3.3 Monte Carlo Method ....................................... 36
3.3.4 Markov Chains ............................................. 37
3.3.5 Markov Chain Monte Carlo ................................................... 39
3.3.6 Metropolis Hastings .......................................................... 40
3.3.7 Gibbs Sampling ................................................................. 43

3.4 Issues of Convergence ......................................................... 43
3.5 Nature of Multidimensional Posterior ....................................... 45
3.6 Graphical Representation of a Model ......................................... 47

4 Growth Model ........................................................................ 50
4.1 Stages of Crack Growth .......................................................... 50
4.2 Characteristics of Microcracks .................................................. 52
4.3 Experimental Details ............................................................... 54
4.4 Data ..................................................................................... 56
  4.4.1 Treatment of Raw Data - Coalescence ............................... 60
4.5 Model .................................................................................. 63
  4.5.1 Hierarchical Population Model ........................................... 64
  4.5.2 Directed Graph ................................................................. 67
4.6 Analysis ................................................................................ 69
  4.6.1 Likelihood ....................................................................... 69
  4.6.2 Priors for the Parameters .................................................. 70
  4.6.3 Evaluation of the Posterior distribution ............................... 71
  4.6.4 Evaluation of Other Quantities ............................................ 71
4.7 Practicalities ......................................................................... 72
  4.7.1 Evaluation of $a_i$ ............................................................... 73
  4.7.2 Sampling from Posterior .................................................... 73
  4.7.3 Algorithm ........................................................................ 74
  4.7.4 Programming ................................................................... 76
4.8 Results ................................................................. 77
4.9 Conclusions ....................................................... 81

5 Coalescence ........................................................ 82
  5.1 Motivation ....................................................... 82
  5.2 Fatigue in PMMA .............................................. 83
  5.3 Data ................................................................. 84
     5.3.1 Preliminary Analysis ...................................... 89
  5.4 Model ............................................................... 90
  5.5 Analysis ........................................................... 91
     5.5.1 Likelihood .................................................... 92
     5.5.2 Priors ......................................................... 92
     5.5.3 Evaluation of Posterior ................................. 94
  5.6 Results ............................................................. 94
  5.7 Conclusions ..................................................... 97

6 Combining Coalescence and Growth .................. 99
  6.1 Background ..................................................... 99
  6.2 Data ............................................................... 100
  6.3 Model ............................................................. 101
  6.4 Analysis .......................................................... 104
     6.4.1 Likelihood .................................................. 104
     6.4.2 Priors ....................................................... 106
     6.4.3 Evaluation of Posterior ............................... 106
  6.5 Results ........................................................... 107
  6.6 Conclusions ..................................................... 110
7 Summary & Conclusions

7.1 Summary of Growth Model .............................................................. 112
7.2 Summary of Coalescence Model ..................................................... 113
7.3 Summary of Joint Model .................................................................. 113
7.4 Conclusions ....................................................................................... 113
7.5 Further Work ...................................................................................... 114
  7.5.1 Spatial Aspect ............................................................................... 114
  7.5.2 Methodological Issues ................................................................. 115
  7.5.3 Factors Influencing Coalescence ............................................ 115
  7.5.4 General Process of Failure ....................................................... 115
  7.5.5 Alternative Growth Models .................................................. 116
7.6 Closing Remarks ................................................................................. 116

A Calculations

A.1 Bayes Theorem .............................................................................. 118
A.2 Inverse Gamma ............................................................................... 118
A.3 Constant Variance Growth Model .............................................. 119
A.4 Multiplicative Variance Growth Model ....................................... 120
A.5 Acceptance Probabilities .............................................................. 121
  A.5.1 Growth Model for $\phi_i$ with Multiplicative Variance ........ 121
  A.5.2 Growth Model for $\phi_i$ Constant Variance .................. 122
List of Figures

2.1 Forman Equation................................................................. 14
3.1 Posterior Density in x and y.............................................. 46
3.2 Transformed Density for Better Mixing.............................. 47
4.1 Growth at Microstructural Level......................................... 52
4.2 Plot of Rates of Growth for 140kN Data............................ 55
4.3 140kN Data................................................................. 57
4.4 160kN Data................................................................. 58
4.5 200kN Data................................................................. 59
4.6 Effect of \( \phi \) on Rate of Growth................................. 66
4.7 Effect of \( m \) on Rate of Growth................................. 66
4.8 DAG for Growth Model..................................................... 68
4.9 Typical Trace for \( \phi \)....................................................... 77
4.10 Typical Trace for \( m \)...................................................... 78
4.11 Typical Trace for \( D \)...................................................... 78
4.12 Estimate of \( R(N) \) for 140kN.......................................... 79
4.13 Estimate of \( R(N) \) for 160kN.......................................... 80
5.1 Data on Mechanism of Coalescence................................. 85
List of Tables

4.1 Raw Engineering Data................................................. 60
4.2 Transformed Data for Growth Model......................... 62
5.1 Raw Data Before Coalescence................................. 87
5.2 Data Used in Coalescence Model......................... 89
5.3 Coalescence Data.................................................. 90
Chapter 1

Introduction

Failure of metal structures is caused by cracks appearing in the material and growing until the strength of the structure is compromised. Such cracks are formed and grow under stresses far lower than the stress at which unstable fracture occurs. It is useful for designers and engineers to be able to predict how long a material will last under certain operating conditions, so that safe structures may be manufactured. Such structures include suspension bridges, ships, bicycles and aircraft.

The way in which cracks grow in metal has been researched extensively, for different types of metal, and for different operating conditions. The great majority of this work has concentrated on long cracks, that is cracks of the order of 1mm or longer. Short cracks do not adhere well to the models which work for the long cracks. In addition, there tend to be many short cracks within a specimen which grow into one another. This interaction between short cracks is termed coalescence, and results in the formation of a dominant long crack.

Reliability modelling is concerned with making probability statements
about how long an object will survive. This dissertation is concerned with modelling reliability by developing a model for short crack propagation and coalescence in steel.

The models have been developed in collaboration with, and the data analysed comes from material experiments carried out at the Department of Mechanical and Manufacturing Engineering, Trinity College Dublin.

1.1 Review of Current Models

Chapter 2 consists of a review of the relevant engineering background. A brief overview of the various issues which have to be considered is given, followed by a more detailed review of the various methods which have been used in the past. These methods are very diverse, and range from plots of lifetime versus stress to explicit modelling of stresses and crack elongation. One of the models reviewed is adapted to model short crack growth.

It is clear from the review in Chapter 2 that while a large number of models exist, they can generally be described as targeted at a particular area of the problem of reliability. This may be developing a deterministic model, developing statistics which correlate with reliability or fitting parameters to empirical data. In the context of such a diverse range of models, this work focuses on detailed examination and adaptation of a single model, rather than a comparative review of how one may adapt the broad range of models in the literature. The main focus is on modelling the underlying growth, rather than an examination of alternative damage functions, which are common in the literature.
1.2 Statistical Methodology

The modelling is carried out in a Bayesian framework. The background to the Bayesian method is detailed in Chapter 3. There has been much interest in Bayesian methods in the recent past. Much of this interest stems from the fact that the computational power is now readily available to make real Bayesian models tractable. In order to sample from the posterior distributions of random variables Markov Chain Monte Carlo (MCMC) is used. Some of the important issues surrounding MCMC are discussed. These issues were of a real practical concern when implementing the models described later in the thesis.

1.3 Growth Model

A model for the rate of growth of short cracks is proposed. This is adapted from a model commonly used for describing the behaviour of long cracks. The physical properties of the microstructure of the material which influence short crack behaviour are taken into account. The model differs from other methods which have been used to model short cracks, in that it focuses on the behaviour of a population of individual cracks; that is the growth of each crack is considered. Alternative methods have used overall damage criteria, such as microcrack density, to describe the behaviour of the population of short cracks.
1.4 Coalescence

Coalescence is a phenomenon observed where a large family of cracks is present in a specimen. It is very different from a pure growth situation since the coalescence of multiple cracks will produce a long crack far quicker than it would take for a short crack to develop to such a length through the mechanism of growth alone. For this reason it is useful to model the phenomenon of crack coalescence. This is done with a view to extending the physical insight which is the foundation for the growth model. A statistical analysis carried out on data of crack formation in orthopaedic hip replacements forms the inspiration for progress in this part of the research.

1.5 Joint Model

A joint model for growth and coalescence is proposed. It allows the growth of short cracks to be modelled explicitly and individually, while taking into account the role of coalescence in forming long cracks. This model is flexible and allows spatial data to be taken into account should such become available. It is recognised that the experimental process is difficult and tedious and with this in mind a recommendation on the type of spatial data which might be routinely recorded is made.
1.6 Overall Framework and Major Contributions

This research is framed as a Bayesian analysis of a reliability problem. As such, techniques are drawn from various disciplines. The main contribution is that insight into the engineering process is obtained through consideration of the process and data from a statistical perspective. Specifically, the following are the main contributions made by this research:

- A hierarchical model is developed to model short crack growth in steel. This constitutes a new application of the methodology in an engineering context;

- A model for rate of coalescence is developed, which is motivated by work on an analysis of damage accumulation in orthopaedic hip replacement fixation;

- A unification of the model for growth and the model for coalescence is carried out, using auxiliary variables for unobserved parameters. Practical conclusions as to what spatial data would be useful and appropriate are drawn, in order to guide future engineering research.

Thus this work develops new statistical models for an engineering problem, and draws practical conclusions in order to focus future research.
Chapter 2

Modelling Fatigue and Crack Propagation

Fatigue is the deterioration of a structure’s ability to sustain normal operating stresses. This deterioration occurs gradually over time, and is caused by cracks growing within the material of the structure. Examples of such failures in an engineering context include cracking in the skin of an aircraft, engine blocks on cars, cranks and forks in bicycles, ships propellers and bridges. In all these cases, the cause of structural failure is the growth of one dominant crack.

While it is possible to model lifetime in more general terms, the key to structural failure is cracking within the material under stress, and when modelling fatigue, this is a good place to start.
2.1 Factors Important to Modelling Fatigue

Fatigue is an inevitable feature of most material structures subjected to cyclic stresses, under quite normal loading. There are a number of factors to consider when trying to determine the useful life of a structure. These aspects should be considered in any overall model for structural failure caused by fatigue. Of particular relevance are specimen geometry, material properties, nature of loading and manufacturing issues.

2.1.1 Specimen Geometry

The shape of the structure under stress is important in determining how it responds to loading. For example, it is well known that sharp edges within structures, and rivet holes, for example, can act as stress concentrators, in such a fashion that very high stress is felt at that point, even though the stress applied to the structure is quite small.

There are many models which may be used to analyse the effect of specimen geometry. Such models include finite element analysis. However, for the purposes of the models here, it is worth just noting that structural geometry is an important issue in any model; an issue that will not be addressed in depth further.

2.1.2 Material Properties

It is well appreciated that materials act differently from one another. For example, some metals are brittle, while others are quite flexible. Even at a microscopic level, there are differences which are important. For example
plastics can form long thin layers, whereas metals often have a more granular structure.

Some materials crack easily, whereas others resist cracking for a long time, until suddenly they shatter. This research is concerned with the rôle of the material in the whole fatigue process. Specifically, the material can affect the way in which cracks propagate, with such propagation dependent on a combination of external and internal material effects. In particular, some of the parameters of the model take into account such features as temperature, crack geometry and strength of grain boundary.

2.1.3 Nature of Loading

A well studied area is modelling of loading. Depending on the structure of interest, different types of loading will be prevalent. For example, a structure such as a bridge will experience a loading for each vehicle that passes over it, with the magnitude of the loading dependent upon factors such as weight of vehicles, and their time of arrival at the bridge. An aeroplane on the other hand, experiences stresses during flight, as a continuous function, coupled with spikes on take off and landing.

Measurement of the stresses at a particular point in a particular structure yields a trace which can be denoted $S(t)$ representing the stress as a function of time. For different applications, the characteristic shape of this function will differ. It is useful to be able to classify the various stresses that will be experienced by structures, so that these may be taken into account during the design process. These functions may be summarised with statistics, which it is assumed yield a classification of the load experienced.
2.1.4 Statistics of Interest

Fractional occupation time is a quantity $l_u$ which is the proportion of time for which the continuous stress, $S(t)$ is above some level, $u$. This is one property of experienced stress.

$N_u$ is the number of times $S(t)$ crosses the stress level, $u$. A crossing is the transition from a stress level less than $u$ to one in excess of $u$ or vice-versa.

The properties of the local maxima of $S(t)$ can be important. The number of times a maximum stress is reached within a period $\tau$, $M_\tau$, together with the size of the maxima, is a useful statistic. In particular, the distribution of the height of the peaks, $f_{\text{max}}(z)$, is of interest.

The relationship of one peak to the next is of interest and an idea of the range of loading is fundamental to any fatigue modelling. In some circumstances, it may be possible to model the loading in the form of a wave, in which case the quantities of interest are amplitude and wavelength.

Where a process has many high frequency oscillations, but there is an overall structure to the magnitude of such oscillations, it may be possible to create an envelope, that is, a function which bounds $S(t)$. In [14], [42], they discuss details of how to do this for a random signal. The loading can then be considered as represented by the properties of the envelope.

In the case of a complicated loading history, it is sometimes possible to approximate this by a sequence of constant amplitude loadings. Cycle counting is one way of doing this [15].

It is possible to use these statistics, that is, $l_u, N_u, M_\tau, f_{\text{max}}(z)$ and the envelope functions, to summarise the loading experienced by a system, and then use data from similar loading situations to make reliability predictions.
Indeed, it is because of this that such methods are important.

While the type of loading plays a vital part in the fatigue process, this research concentrates on the study of material properties examined under laboratory conditions. In this case, the data come from sinusoidal loading with a well-defined maximum, minimum and a constant controlled stress range. Time may thus be represented in terms of number of cycles, which is well defined, and is denoted $N$. It is clear that using $N$ as time brings with it the implicit assumption that the frequency of stressing is unimportant, and in any case can be taken account of. In practice this is a reasonable assumption. Increasing the frequency of the stressing under laboratory conditions allows simulation of loading which would be experienced in real conditions over many years.

2.1.5 Manufacturing Issues

A chain is only as strong as its weakest link. This is a somewhat overused expression, but goes unappreciated by some of the people modelling fatigue. It is possible to be very careful about all aspects of the structural design, determine the material characteristics precisely, yet not appreciate the degree to which the actual structure is susceptible to error during construction.

An example of an error which resulted in catastrophic failure was the case involving a light aircraft [1] in 1996. A weld was improperly completed on the nose wheel, which led it to shear off on contact with the runway. The failure was due to fatigue cracking which had been initiated when the weld was carried out.

For this reason, it is most important that the overall picture is concerned
with the very real part that the manufacturing process plays in the determination of system reliability.

2.2 Deterministic Models for Crack Growth

A deterministic model of fatigue may be considered to be a mathematical system which will allow one to make accurate predictions about the lifetime of a material or structure, given information about the material properties, details of the geometry, and the actions to which it is subjected. A deterministic model is one such that if the various parameters are specified exactly, then an exact prediction for the lifetime is obtained.

In practice there is a large amount of scatter in observed lifetime data for similar materials under similar conditions. This could be because the true parameters are not known exactly, or because there is a natural variability in the physical system. Such scatter will be addressed in the probabilistic adaptations of some of the deterministic models.

The different types of model may be roughly divided into two categories; namely those which are empirical – based on observed data, and constructed to fit the data – and those that are more theoretical – based on some physical reasoning, or mechanism which is known to affect the lifetime. Details of both empirical and theoretical deterministic models for crack growth are discussed in what follows.
2.3 Wöhler or S–N Curve

A graph of stress versus lifetime constitutes one of the first attempts to quantitatively examine fatigue lifetimes. This graph, introduced by August Wöhler in 1858, has since become known as the S–N curve. Although the S–N plot provides an estimate of lifetime for a given stress level, it does not take into account any estimate of uncertainty. S–N–P plots are a probabilistic modification which give an estimate of probability of failure, conditional on number of cycles and stress level.

It should be noted that this is very much an empirical method in the sense that the only variables are time and stress. Calibration of such curves are done by experiment, using test specimens.

Another common method of estimating lifetime consists of fitting the parameters of a Weibull distribution from test data, and this can yield good predictions in practice.

2.4 Empirical Models

Laboratory tests on materials and observation of structures indicate that the length of the largest crack is dependent on factors such as the stresses, material properties, temperature and chemical agents to which the sample or structure is subjected. Two such equations which describe crack growth, and fit well with the observed data are the Paris-Erdogan and Forman equations.


2.4.1 Paris-Erdogan Equation

The Paris–Erdogan equation is derived from empirical considerations, and has no real theoretical basis. The equation models the relationship between the velocity of crack propagation and an abstract quantity called the range of stress intensity, which describes the magnitude of the stress at the crack tip. This range is denoted $\Delta K$ and is usually defined as $\Delta K = Q\Delta \sigma \sqrt{a\pi}$, where the constant $Q$ reflects the crack geometry, $\Delta \sigma$ is the stress range, and $a$ is the length of the crack.

The form of the Paris–Erdogan equation is

$$\frac{da}{dN} = C(\Delta K)^m,$$

(2.1)

where $C$ and $m$ are regarded as material constants that depend upon factors such as frequency, temperature and stress ratio. The stress ratio, which is defined to be $R = S_{\text{min}}/S_{\text{max}}$ has an important effect on crack growth, according to [2], but does not explicitly appear in Paris–Erdogan.

The Paris–Erdogan equation gives good results for long cracks when the material constants are known, but a large effort is required to determine them, since they are functions of many variables. $R$ is an observable and where it appears in the model explicitly, the effort in determining the remaining material constants is much reduced. An empirical equation which incorporates the stress ratio, $R$, is the Forman equation.

2.4.2 Forman Equation

The Forman equation accounts for the stress ratio, and can be written in the following form
Figure 2.1: Graph of $\Delta K$ versus $\frac{da}{dN}$. Forman Equation.

\[
\frac{da}{dN} = \frac{C(\Delta K - \Delta K_0)^m}{((1-R)K_c - \Delta K)},
\]

where $K_c$ is a critical level for the Stress Intensity $K$, corresponding to unstable fracture. This is called the fracture toughness. $\Delta K_0$ is a level below which no damage occurs; for some materials this is zero, but for many ferrous materials it is positive. The graph in Figure 2.1 of this function shows that there exist two limits such that

\[
\lim_{\Delta K \downarrow \Delta K_0} \left( \frac{da}{dN} \right) = 0,
\]

and

\[
\lim_{\Delta K \uparrow \Delta K_c} \left( \frac{da}{dN} \right) = \infty.
\]

It is clear from this that there is a lower limit to $\Delta K$, below which no crack growth takes place, and an upper limit, around which growth is explosive.
Finally, with regard to the Forman Equation, it has been shown to give good results for crack growth in aluminium alloys and steels, but not in many other materials.

Other empirical models appear in the literature; for example [44] via the “Famous Robder Law” demonstrate, partly in jest, how common practice it is to propose new models to fit data within this field. There are very many adaptations to well known laws which are not mentioned here.

2.5 Theoretical Models

While the Forman Equation in particular might yield some nice physical insight, one can tackle the whole problem of structure lifetime directly, using the stress intensity factor and so called fracture toughness. This requires application of fracture mechanics, a detailed explanation of which is beyond the scope of this dissertation, and which, in any case, is well covered in numerous texts such as [45].

It is sufficient to note that application of fracture mechanics, with the appropriate examination of the crack geometry, nature of stress applied, and length of crack to date, along with experimental data for materials, can give an idea of when failure may be likely to occur.

2.5.1 Cumulative Damage Theories

In order to estimate the reliability of real components which undergo varying conditions and stresses at different times in their lifetime, cumulative damage theories were introduced. The idea behind these is that failure is due to a
structure accumulating damage at different stress levels until finally fracture occurs. Initially, it was proposed that such accumulation was just a linear combination of damages, but this was subsequently modified to give a non-linear accumulation rule.

A continuous version of the accumulation rule, a damage function, was proposed in the 1950s by Kachanov [22]. This whole area is called continuum damage mechanics. The use of the damage function allows one to take into account the many and varied mechanisms that affect useful lifetime. As well as the actual crack growth process, one may be interested in the effect that such things as weather plays in the accumulation of damage. Indeed, it is possible to consider damage at the initial stages of the damage process, by determining a general “pattern” of damage, from experiment, rather than trying to observe the initiation process directly. Such consideration may be made in the specification of the damage function. Indeed much of the previous work on short cracks has concentrated on an examination of damage functions such as total length of cracks per unit area, $L_{\text{tot}}$ [4].

2.5.2 Elastic Fatigue Fracture

Under certain conditions it is possible to calculate the stress experienced at any point within a material, using simple engineering theory. Specifically, if the stress is not too great, the material acts under what is called the elastic regime and elastic fatigue fracture occurs. In such a situation, the stress felt at the tip of a crack can be calculated, and the crack opening force deduced.

Using these methods it is possible to determine the rate of crack propagation. However, the material will not always follow elastic behaviour, and
plastic effects must be taken into account. J-integral methods deal with such a situation.

2.5.3 J Integral Methods

The path independent J integral proposed by Rice (1968) [41] describes the stress–strain field at the tip of the crack under elastic–plastic loading. Since the integral can be taken far from the crack tip and then information about the crack tip may be deduced, this allows an extension of fracture mechanics associated with elastic fatigue fracture from linear elastic to elastic–plastic behaviour.

From experiment, under certain theoretical conditions, it has been shown that the governing equation for crack growth can be written in a form similar to Paris-Erdogan,

$$\frac{da}{dN} = C(\Delta J)^m,$$

where $\Delta J$ is the J–integral range, and $m$ is determined from experiment. It should be remembered that the J integral methods have strong analytical basis, which differs from the totally empirical Paris–Erdogan.

2.5.4 Empirical Stochastic Parameterisation

There are many ways to incorporate a probabilistic aspect into a so-called deterministic model. One way to do this is to add random noise to the model. Another way is to associate a random distribution with the model parameters. Yang et. al. (1983) [58] demonstrate this for a hyperbolic crack growth rate function. A quick overview of the method demonstrates the idea.

Given the crack growth equation;
\[ Y = C_1 \sinh[C_2(X + C_3)] + C_4, \quad (2.3) \]

where \[ Y = \log\left(\frac{da}{dN}\right) \quad \text{and} \quad X = \log(\Delta K). \]

one can add in a random element in two ways. The first is to add a Gaussian random process, \( Z(X) \). The crack growth equation then becomes

\[ Y = C_1 \sinh[C_2(X + C_3)] + C_4 + Z(X). \quad (2.4) \]

The task then remains to evaluate the precise form that \( Z(X) \) might take.

The next is to consider \( C_2, C_3, \) and \( C_4 \) as random variables. In this case the equation would take the same form as Equation 2.3, where the coefficients are considered to be random quantities.

### 2.5.5 Stochastic Models

A number of different stochastic models exist in the literature, and a review demonstrates the diversity of such models.

Markov chains can be used to model fatigue in materials. The assumption that damage is a function of independent parameters, combined with damage accumulated to date is consistent with the Markov property, and hence such methods are employed in a natural fashion. Of specific interest to this research is work on short cracks by Cox and Morris [12], [13]. The continuous version, the Markov diffusion, has also been examined.

The differential equation approach assumes that cracks grow continuously. In reality, crack growth can be a discontinuous process. In order to model this, it may help to consider growth as a combination of a growth event,
together with a certain growth magnitude attributable to that event. Such has been modelled in the cumulative jump models [49] [23] [48].

2.6 Fatigue Crack Growth Process

As mentioned above, the agent which causes fatigue is crack growth and propagation through a material.

The process of fatigue can be considered to consist of five steps;

1. Dormant. The structure is devoid of cracking.
2. Initiation / Nucleation. This consists of new cracks forming.
3. Short or microcrack growth. Microcrack growth is the term given to the growth phase during which the crack is microstructurally small.
4. (Large) crack propagation. This is the phase of crack growth which has been extensively modelled in the literature.
5. Failure. This is when the structure finally fails. The crack grows very quickly in this phase and so it is relatively short.

Much work has previously been done on modelling crack growth behaviour, the great majority of this having been done for long cracks. The general cause of failure is a single crack growing exponentially. However, a substantial portion of the total time to failure can be spent in the short crack phase of development, and this needs to be taken into account in any estimation of reliability.

It is with this in mind that a study of microcrack propagation was undertaken. A Bayesian framework provides many practical advantages for this
investigation. No previous Bayesian modelling of this problem has appeared in the literature.

2.7 Current Research Questions

Aim 1 The primary aim of this research is to develop a suitable model or models in order to make good reliability predictions for the material of interest.

In order to clarify this aim the following definitions are useful.

Definition 2.1 Specimen. A specimen is a piece of material tested by engineers in an experimental context, under laboratory conditions.

Often a specimen will consist of a test structure, or a model of a structure which will be put into operation in a real project. In the context of this research, the specimens are designed so that the fundamental properties of the material, rather than any individual structure, may be assessed.

Definition 2.2 Failure. Complete failure shall be defined to be that time such that the specimen contains a crack of the same dimensions as the specimen itself.

Definition 2.3 Threshold length. \( A_{th} \) will be termed the threshold length, and is a length such that the time to complete failure for a specimen exhibiting a crack of this length will be known (to within a given degree of accuracy), and in general will be short (in comparison to the total lifetime for the specimen.)
**Definition 2.4** Reliability. Reliability, $R(t)$ is the probability that the system does not fail by time $t$. For the purpose of this research, the reliability is defined to be the probability that no crack within the specimen exceeds the threshold length.

The purpose of the research, therefore, is to develop models which yield a prediction for how long a specimen will last, together with an estimate of the uncertainty attached to this prediction. In particular, an estimate for $R(t)$ will serve this purpose.

**2.7.1 Growth Aspect**

In order to make statements about the probability that no crack exceeds a particular threshold length, $A_{th}$, it is necessary to examine how cracks grow within the material.

The area of primary research, was in the area of modelling crack growth. To this end various models were examined, leading to closer scrutiny of the Paris–Erdogan model. An adaptation of Paris–Erdogan is developed. Consideration is given to a random variable parameterisation of the model, and a hierarchical population model is developed for the family of microcracks within specimens [57].

As was seen above, models for long crack growth are plentiful in the literature. It was decided to examine other aspects of the problem rather than choosing to compare variations on other standard models in the literature.
2.7.2 Coalescence Aspect

Since there is more than one crack in each specimen, and since this is the case in general in structures, there is the possibility of interactions between cracks. These interactions take the form of cracks coalescing, which leads to longer cracks becoming established. While there is ongoing debate as to how these interactions are significant within the context of overall reliability, in general, in this case we felt that it would be useful to model this effect directly.

A study of reliability in a very different situation, on the application to orthopaedic hip replacements, led to some insight on other methods of damage accumulation. In particular, it was clear that a large family of short cracks cause as much damage as a single long crack.

For the specimen data, a number of approaches were made at modelling the effect of coalescence. The effect was modelled in two distinct stages. Firstly the rate of coalescence was modelled and then the growth and coalescence models were combined. Again the analysis was carried through in a Bayesian framework.

2.7.3 Other Considerations

There are a number of other aspects which one may validly examine in order to answer the research question. Indeed the whole question of crack initiation and initiation length is not dealt with directly. The spatial aspect of the problem was not addressed explicitly, since there was no relevant data available. These, and some of the other aspects are mentioned in the section on future research.
Chapter 3

Statistical Methodology

The methods used in analysing data and drawing inferences are termed the statistical methodology. Underlying any of the models discussed in detail in this work, there is a significant amount of statistical methodology, which is introduced and discussed in this chapter.

3.1 Statistics

It is mentioned in an introductory text [10], that the term statistics refers to a collection of numerical facts and estimates, the purpose of statistics being to enable correct decisions to be taken. Elsewhere [33], it is noted that one of the functions of statistics is the provision of techniques for making inductive inferences based upon data. It is important also to have an estimate of the uncertainty attaching to those inferences.

In real life situations, information can often be usefully summarised numerically. For example, percentage unemployment, mortality rate for males aged 65, or maximum stress level below which a structure is likely to survive.
Statistics have long been used to estimate such quantities based on observed data. For example a random survey of cars aged 8 to 10 years in a particular country, may show that, say, 20 out of 100 examined were structurally unsound due to rust damage. From this it may be inferred that the proportion of cars in the country of that particular age which had severe rust is in the region of 20%. Of course, there is some uncertainty attached to this estimate, and if 100 different cars were surveyed then a different answer may have been obtained, and there are ways of estimating the uncertainty. In classical statistical inference what one is doing is making an estimate of the true (but unknown) proportion, based on data. The assumption is that the proportion of the total population of cars which experience severe rust is a fixed unknown, and that data is being used to estimate it.

In the context of this research, statistics may be defined to be concerned with the analysis of data collected under uncertainty. Specifically, the aim is to develop suitable models, in order to make reliability predictions based upon recorded test data. Classical, or frequentist methodology in statistics concentrates on making inferences about the true situation having observed certain data, whereas the Bayesian approach is concerned with updating subjective knowledge in the light of data.

3.2 Bayesian Approach

Bayesian inference is different from classical inference, in that one is concerned with answering the following question, “What should a rational person believe, after collecting the data, and given what was believed before the data was collected?”
Essentially, this question differs from what a classical statistician asks in a number of different ways;

- The question is unapologetically subjective.
- Previous information is important.
- The focus is rational belief based on current knowledge, rather than on obtaining an estimate of any "true" value.

The Bayesian framework has attractions for a number of reasons [5]. Bayesian statistics has a strong axiomatic foundation, it incorporates prior information directly into the analysis, and it has a naturally formulated decision structure. Bayesian inference has not been as commonly used as frequentist methods in the past, in part due to computational complexities [27]. Since about 1960 there has been a revival of interest [38] to the extent that it is now well established as an alternative to classical methods.

As to the question of why one might choose to undertake a Bayesian analysis of a situation, rather than an appropriate classical analysis, the answer is simple. Apart from the philosophical reasons, for a number of real problems the answer is that the methodology works [46].

3.2.1 Formal Bayesian Methodology

More formally, the following is the method employed. As mentioned above, statistics is concerned with the estimation of numerical quantities. In the Bayesian context, the quantities of interest will be random variables, and could, for example, be the proportion of rusted cars as referred to above.
Before an experiment or survey, the prior knowledge about the quantities of interest are summarised in the form of a probability statement.

Denote the parameter or parameters of interest \( \theta \) and the state of current experiences to date is denoted \( H \). Such experience might be to do with knowledge of the material properties of cars, the nature of the roads, the weather and indeed knowledge of previous studies. The probability statement about initial beliefs is denoted \( p(\theta|H) \) and is termed the prior belief. Since this is a probability statement it takes the form of a probability distribution and is often referred to as the prior distribution, or more simply the prior.

**Prior Knowledge**

There are a number of philosophical issues raised in any discussion on prior probabilities. For further information on such discussion see [38], [27], [5]. It is essential, when considering \( \theta \) as a random variable, to assign prior probabilities, simply because such must exist. In the case where prior knowledge shows that no particular value or values of \( \theta \) are more likely than any others, then \( \theta \) will be uniformly distributed. That is to say, \( p(\theta|H) \propto 1 \). It is important to note that such a statement of initial belief is saying that at the outset, it is believed that for example 100% of cars being rusty is as equally likely to be prevalent as 0%, or indeed any other intermediate value.

A more reasonable situation would be one where cars are being surveyed in the light of previous work and with some knowledge of the materials involved. Then the prior might take the form of a normal distribution with some mean and (perhaps large) variance.

For notational simplicity \( \pi(\theta) \) is written and taken to mean \( p(\theta|H) \) from
here onwards.

**Model or Likelihood**

The idea of likelihood is common to all statistical inference, and is well understood by frequentist and Bayesian statisticians alike.

The relationship between the parameters of a model and the observables is fundamental to the process of updating knowledge of parameters based upon the data. The likelihood is sometimes termed the model, and takes the form of a probability statement $p(X|\theta)$, where $X$ are the observable data in the system.

Note that the likelihood is a conditional probability statement as to how likely it is for $X$ to be observed if the parameters take the value $\theta$. In a statistical analysis, it is the knowledge of $\theta$ which is of interest, that is to say, the distribution of $\theta$ given that $X$ is observed. This is termed the posterior, and is dealt with below.

Other methods of inference concentrate on the likelihood in their analysis, in which case the focus is $p(X|\theta)$ as a function of $\theta$ for fixed $X$. Of course while $\int_X p(X|\theta)dX = 1$ the same is not true of the integral with respect to $\theta$. For this reason, and to avoid confusion, the likelihood is sometimes written $l(\theta|X)$.

**An Example**

O'Hagan [38] gives a somewhat contrived example of why it is important to consider the prior as well as the likelihood. Let $G$ be the event of seeing a big green structure, with blob like attachments outside a window. Let $T$ be the
hypothesis that a tree is outside the window, and let $C$ be the hypothesis that a cardboard model is outside the window. Since $C$ and $T$ are equally consistent with the observation, $G$, one shouldn't have any reason for believing one over the other. That is $l(C|G) = l(T|G)$. However, the probability that $C$ is in fact outside the window, conditional on the observation, is $p(C|G)$, which depends on $p(C)$, the prior probability of cardboard structures being outside windows, and is likely to be much less than $p(T|G)$. Incorporation of prior knowledge is an essential part of the inference.

**Posterior Distribution**

Of interest to the modeller, then, is the conditional distribution of the parameters, given the data, that is $p(\theta|X)$. Bayes Theorem for random variables [27] yields

$$p(\theta|X) \propto \pi(\theta) \cdot p(X|\theta).$$

The distribution $p(\theta|X)$ is termed the posterior distribution and describes the current state of knowledge about $\theta$, given the initial knowledge of $\theta$, together with the model, such knowledge having been updated by information. The constant of proportionality in the above is just $\frac{1}{p(X)}$ where $p(X)$ can be obtained from $p(X) = \int p(X|\theta)\pi(\theta)d\theta$.

The Bayesian method, is then, quite straightforward [16]:

- construct a model, obtaining a likelihood $p(X|\theta)$;
- elicit a prior distribution $\pi(\theta)$;
- derive the posterior density $p(\theta|X)$ as above.
In practice these tasks can be difficult to implement, and more is said about the details later.

3.2.2 Predictive Distribution

In the case where one is interested in making a probability statement about the distribution of the random variable of interest, given that one has observed realisations $\mathcal{D} = \{x_1, \ldots, x_n\}$ one can use the marginal distribution

$$f(X|\mathcal{D}) = \int_{\Theta} f(X|\mathcal{D}, \Theta)f(\Theta|\mathcal{D})d\Theta$$

(3.1)

which is termed the predictive distribution. In practice, this integral can not generally be calculated, since the analytical form of $f(\Theta|\mathcal{D})$ is not known. However, samples may be drawn from $f(\Theta|\mathcal{D})$, in which case the predictive distribution, together with any other distributions may be estimated using the kernel density estimate.

3.2.3 Kernel Density Estimation

Kernel density estimation consists of estimating a posterior density for a function of interest, using samples from the posterior, often drawn using one of the numerical techniques. Let $\vartheta_1, \ldots, \vartheta_n$ be samples from the posterior distribution $f(\Theta|\mathcal{D})$. If one is interested in the properties of the posterior density function $g(X|\mathcal{D})$, where conditional on $\Theta$, $X$ is independent of $\mathcal{D}$, that is $g(X|\mathcal{D}, \Theta) = g(X|\Theta)$, the following result is useful;

$$g(X|\mathcal{D}) = \int_{\Theta} g(X|\mathcal{D}, \Theta)f(\Theta|\mathcal{D})d\Theta$$

$$= \int_{\Theta} g(X|\Theta)f(\Theta|\mathcal{D})d\Theta$$

$$= \mathbb{E}_{\Theta|\mathcal{D}}[g(X|\Theta)].$$

(3.2)
This expected value may be approximated in the usual fashion, as a simple numerical average of the values of the function at each of the sample points. That is using \( \hat{g} \) given by

\[
\hat{g}(X|D) = \frac{1}{n} \sum_{i=1}^{n} g(X|\theta_i)
\]

(3.3)

The fact that \( \hat{g} \) is a density function follows from the fact that each of the \( g(X|\theta_i) \) is a density function. Kernel density estimation is a standard method of examining posterior distributions, and properties of functions of the parameters.

### 3.2.4 A Simple Example - \( N(\mu, \frac{1}{\tau}) \)

Consider the case of drawing from a population of unknown mean, \( \mu \), but known variance \( \frac{1}{\tau} \). (\( \tau \) is termed precision, and is just the reciprocal of variance.)

The model is that the data, \( X \), will be normally distributed with unknown mean but given variance. Thus, in terms of a single observation, \( x \), we can write down the likelihood;

\[
p(x|\mu) = \sqrt{\frac{\tau}{2\pi}} \exp \left( -\frac{\tau}{2}(x - \mu)^2 \right).
\]

The next step is to elicit a prior for \( \mu \). It may be reasonable to assume that the prior beliefs about \( \mu \) can be expressed as a normal distribution, that is

\[
\mu \sim N(\nu_{prior}, \frac{1}{\rho_{prior}}).
\]

where both \( \nu_{prior} \) and \( \rho_{prior} \) are specified. Typically \( \nu_{prior} \) is the expected location of \( \mu \), and \( \rho_{prior} \) is an expression of how precise that estimate is. In general, \( \rho_{prior} \) will be small.
Thus, having collected data, it is possible to derive the posterior for $\mu$ according to Bayes theorem for random variables;

$$p(\mu|x) \propto p(x|\mu) \cdot \pi(\mu)$$
$$= \sqrt{\frac{\tau}{2\pi}} \exp \left( -\frac{\tau}{2} (x - \mu)^2 \right) \sqrt{\frac{\rho_{\text{prior}}}{2\pi}} \exp \left( -\frac{\rho_{\text{prior}}}{2} (\mu - \nu_{\text{prior}})^2 \right)$$
$$\propto h(\nu_{\text{prior}}, \rho_{\text{prior}}, x) \cdot \exp \left( -\frac{\mu^2}{2} (\rho_{\text{prior}} + \tau) + \mu(\nu_{\text{prior}} \rho_{\text{prior}} + x\tau) \right),$$

where $h(\cdot)$ is independent of $\mu$. Defining

$$\rho_{\text{post}} = \rho_{\text{prior}} + \tau \quad \text{and} \quad \nu_{\text{post}} = \frac{\tau}{\rho_{\text{post}}} x + \frac{\rho_{\text{prior}}}{\rho_{\text{post}}} \nu_{\text{prior}},$$

and multiplying by $\exp \left( -\frac{1}{2} \rho_{\text{post}} \nu_{\text{post}}^2 \right)$ which is independent of $\mu$, the above is

$$\exp \left( -\frac{1}{2} (\mu^2 \rho_{\text{post}} - 2\mu(\nu_{\text{post}} \rho_{\text{post}} + \rho_{\text{post}} \nu_{\text{post}})) \right),$$

which reduces to

$$\exp \left( -\frac{\rho_{\text{post}}}{2} (\mu - \nu_{\text{post}})^2 \right),$$

which is the form of the normal density with mean $\nu_{\text{post}}$ and precision $\rho_{\text{post}}$. Thus, in the case of inference for the unknown mean, with normal prior, the posterior is normal. This simple form of the posterior depends on the choice of the prior, given the likelihood. The choice of prior that leads to the simple posterior, is called a conjugate prior; more formally, given a likelihood, $l(\theta|X)$, then a prior chosen from a family of densities, such that the posterior is also from that family, is said to be conjugate.

As can be seen from the above, in the case of conjugate densities, the problem of obtaining a posterior is simplified [5]. However, this is only appropriate where the chosen prior distribution, with suitable parameters can
accurately represent the prior knowledge. The alternative is to use numerical techniques to obtain the properties of interest from the posterior distribution.

The question of prior elicitation is one that needs mentioning also. Apart from the philosophical difficulties that many have with prior probabilities, there are practical problems which need addressing.

### 3.2.5 Prior Elicitation and Non-informative Priors

Difficulties have arisen with specifying a prior in the situation where there is, in fact, no actual prior information. While it was possible to specify a uniform prior for the example of determination of the proportion of rusted vehicles (i.e. \( \pi(\theta) = 1 \)) this is not possible where the possible range for \( \theta \) is infinite and the prior being a proper distribution. A prior \( \propto 1 \) for the range \((0, \infty)\) is a solution, as an improper prior, but even then issues arise as to transformations of the parameters of interest. Clearly, if \( \pi(\theta) = 1 \) then all values of \( \theta \) in the range \([0, 1]\) are equally likely. This is not prior ignorance as maintained in [38] but is in fact a concrete and active statement of prior belief that all values of \( \theta \) are as likely as each other, and that belief will quite properly correspond with a non-uniform prior for transformations of \( \theta \). For example, if we have \( N \) competitors each running in a race, with 1 from country A and \( N - 1 \) from country B, and prior information tells us that each is equally likely to win the race, then this does not correspond to prior information that country A and country B are equally likely to have winners. It is important, therefore to ensure that it is clear as to what prior information is being elicited.

Prior elicitation is the process of specifying, in the form of a probability
distribution, prior information about the parameters of interest. The practical issues detailing methods of obtaining an informative prior are dealt with in [37]. Examples in practice are mentioned in [46] and [55]. It is the assertion of this author that all priors are informative and that for this reason, due consideration should be given in every circumstance to the elicitation process.

In including an informative prior, the statistical analysis is not objective. It has been mentioned above that the Bayesian framework is unapologetically subjective, and this is emphasised once again here.

In the past there have been attempts to "objectify" Bayesian techniques. Notably we have work by Jeffreys [21], but this depends on the form of the data. Subjective scientific inquiry seems a contradiction in terms, but is quite acceptable, provided that we realise that we have subjective inputs, and are careful about such things. For this reason, Bayesian statisticians are interested in concepts of sensitivity and robustness [3].

### 3.3 Sampling from the Posterior Distribution

In any Bayesian analysis, the aim is to obtain posterior estimates for some parameters, or functions of parameters. In a limited number of cases, such estimates may be directly obtained e.g. in the case of conjugate priors. However, in general, this is not the case, and one has to resort to more indirect methods.

Before the advent of modern numerical techniques, and computing power,
the necessary calculations were in practical terms impossible. However, be­cause of the advances of technology, and due to the development of powerful numerical methods in a range of disciplines, infeasible problems of the past have become tractable.

The most important of these techniques in Bayesian statistics has been Markov chain Monte Carlo and in particular Gibbs sampling and Metropolis Hastings.

3.3.1 Stratified Sampling

Consider a set of $N$ types of job within an organisation, which has a total of $M$ employees. Let $J_j$ where $1 \leq j \leq N$ be the number of people who have a job of type $j$ with all people doing the same type of job getting paid the same salary. Then, clearly;

$$\sum_{j=1}^{N} J_j = M.$$ 

If interested in the average salary paid and if $M$ is very large the average may be approximated as follows;

$$\mu_X \approx \bar{X} = \frac{1}{m} \sum_{i=1}^{m} X_i,$$

where we sample a total of $m$ people from the organisation and $X_i$ is the salary paid to the $i^{th}$ person we sampled. Ordinary random sampling would involve picking the $m$ people uniformly from the total population of $M$ people in the organisation. However, another method would be to ensure that the probability of choosing a person from job type $j$ is the number of people doing job type $j$ divided by the total number of people, $M$. This
latter idea is just stratified sampling and is an important and well known sampling technique.

3.3.2 Importance Sampling

Importance sampling is a technique for numerically approximating an integral. It is mentioned here as a basis for the numerical concepts which follow. It is similar to stratified sampling in that the fundamental idea is that the sampling process is distorted, to take into account the weighting of the underlying distribution.

An example of importance sampling in a Monte-Carlo context, is detailed in Section 3.3.3, but the basic principle is as follows:

In wanting to estimate

\[ I = \int_{-\infty}^{\infty} g(x)f(x)dx, \]

where \( f(x) \) is a density function, one could sample \( n \) values of \( x \) from \( f(x) \) and then approximate with

\[ \hat{I} = \frac{1}{n} \sum_{i=1}^{n} g(x_i). \]

Alternatively, \( m \) values of \( x \) could be sampled from another density \( h(x) \) and then \( I \) could be estimated using

\[ \hat{I} = \frac{1}{m} \sum_{i=1}^{m} \frac{g(x_i)f(x_i)}{h(x_i)}. \]

Consideration can then be made as to how \( h(x) \) may be chosen so that the estimator is most efficient. It turns out that the most efficient form for \( h(x) \) samples from areas where \( g(x) \) is large, provided that \( f(x) \) is not small, [25]. Such ideas are important in any method when simulating from the posterior.
3.3.3 Monte Carlo Method

Markov chain Monte Carlo is an important technique used by Bayesian practitioners to sample from the posterior distribution. The Monte Carlo method is, in general terms, any technique used for obtaining solutions to deterministic problems using random numbers. The term Monte Carlo was coined by von Neumann and Ulam in the 1940’s in the context of such problems [34].

A simple example of this [25] is the evaluation of the following integral;

\[ I = \int_{y}^{\infty} \frac{1}{x} \lambda \exp(-\lambda x) dx. \]

Analytical solution of the above is difficult, but Monte Carlo simulation proposes the following;

1. Let \( i = 0 \); Let \( N \) be some large number.

2. Sample \( x_i \) from the exponential so \( f(x) = \lambda \exp(-\lambda x) \)

3. Let \( g(x_i) = \frac{1}{x_i} \) if \( x_i > y \) and 0 otherwise

4. Let \( i = i + 1 \). If \( i < N \) return to step 2.

5. Then \( I \) is estimated by \( \hat{I} = \frac{1}{N} \sum_{i=0}^{N-1} g(x_i) \).

Observe that the above is the standard estimator for \( E[\frac{1}{x} | x > y] \). In practice, many of the values of interest are expected values. To obtain posterior expectations of a function of our parameter, \( f(\theta) \), we need to calculate integrals of the type

\[ E[f(\theta)|X] = \frac{\int f(\theta)p(X|\theta)p(\theta)d\theta}{p(X)} \]
It is possible to use the above idea of Monte Carlo methods, importance sampling, together with some Markov Chain theory, to efficiently approximate such expressions. Some theory is outlined below.

### 3.3.4 Markov Chains

**Definition 3.1** *Stochastic process.* A stochastic process is a collection of random variables, $X_i$ where $i \in I$ for some indexing set $I$, each $X_i$ taking values in a state space, $S$.

**Definition 3.2** *Markov chain.* A Markov chain is a stochastic process with a discrete indexing set, $I$ such that the conditional distribution of $X_{t+1}$ is independent of all other previous states given $X_t$, that is $p(X_{t+1}|X_1, X_2, \ldots, X_t) = p(X_{t+1}|X_t)$.

For simplicity, theory and details are given for a discrete state space, $S$.

**Definition 3.3** *Stationary (in time).* A Markov Chain is said to be stationary $\iff \forall j, k \in S$, and $\forall i \in \{1, 2, 3, \ldots\},$

$$P(X_i = j|X_{i-1} = k) = P(X_1 = j|X_0 = k).$$

A stationary Markov chain is sometimes referred to as homogeneous in time, since, by definition, the probability of moving between two states remains constant in time.

**Definition 3.4** *Markov Matrix.* For a stationary Markov chain, the matrix of probabilities,

$$M_j^k = P(X_n = j|X_{n-1} = k)$$

is called the Markov Matrix.
Note that this definition is independent of $n$ (stationarity), that the entries in the Matrix are $\in [0, 1]$ (probabilities) and that $\sum_{j} \mathcal{M}_{j}^{k} = 1$, since the chain must move to some state, $j$. This is sometimes called a transition matrix, and the associated probabilities transition probabilities. It is also worth noting that the matrix $[\mathcal{M}^{m}]_{j}^{i}$ is the matrix of probabilities $P(X_{n+m} = k|X_{n} = j)$.

**Definition 3.5** Connected. A Markov chain is said to be connected or irreducible, if $\forall j, k \in S$, there exists a sequence $i_{1}, \ldots, i_{n}$ such that

$$\mathcal{M}_{j}^{i_{n}} \mathcal{M}_{i_{n-1}}^{i_{n-1}} \cdots \mathcal{M}_{i_{1}}^{i_{1}} \neq 0.$$ 

That is, there is a non-zero probability of going to state $j$ from state $k$ in $n$ steps, for some $n$.

**Definition 3.6** Recurrent. A state $j$ is said to be recurrent $\iff \sum_{n=1}^{\infty} [\mathcal{M}^{n}]_{j}^{j} = \infty$ else it is said to be transient.

**Definition 3.7** Aperiodic. The period $d(j)$ of a state $j$ is that integer such that $[\mathcal{M}^{n}]_{j}^{j} \neq 0, \forall n$ such that $d$ divides $n$. A state with $d(j) = 1$ is said to be aperiodic.

**Definition 3.8** Limiting Distribution. If

$$l_{j} = \lim_{n \to \infty} [\mathcal{M}^{n}]_{j}^{i}$$

exists $\forall j$ (independent of $i$), then this is called the limiting distribution of the Markov chain.

**Definition 3.9** Stationary distribution. A stationary distribution for a Markov chain is a distribution $\pi$ such that $\pi_{j} \geq 0$ for all $j$, $\sum_{j} \pi_{j} = 1$ and

$$\pi = \mathcal{M} \pi$$

38
The stationary distribution is also referred to as the invariant distribution or equilibrium distribution of a Markov chain.

**Theorem 1 (Ergodic)** For an irreducible, aperiodic, positively recurrent Markov chain, a unique limiting distribution exists, which is the invariant distribution for the chain.

Recall the discussion above regarding stratified and importance sampling. If it were possible to construct a Markov chain that would visit each category the ‘correct’ number of times, then this method could be used to sample from the distribution of interest. In practice, what ‘correct’ means here, is that the equilibrium distribution of the Markov chain is the same as the distribution of interest. In a sense this is the reverse of the theory above, since the distribution of interest is known, and the Markov chain needs to be constructed.

It is possible to do this, under certain conditions, and there are a number of ways of doing it. Of primary interest will be the approach of Metropolis-Hastings.

### 3.3.5 Markov Chain Monte Carlo

Let $\phi_j$ be the distribution of interest. Let $M_j^i$ be the Markov matrix to be constructed. Now, what is needed is a method of constructing $M_j^i$ so that it is indeed a Markov Matrix, and that the stationary distribution of this Matrix is $\phi_j$, the distribution of interest.

**Definition 3.10** Detailed Balance. If $\phi$ is some probability distribution then
(\mathcal{M}, \phi) \text{ satisfies detailed balance } \iff 

\mathcal{M}_j^i \phi_i = \mathcal{M}_i^j \phi_j.

This property yields a method of constructing a suitable matrix, by using the result of the following theorem.

**Theorem 2** \((\mathcal{M}, \phi)\) satisfies detailed balance \(\Rightarrow \phi \text{ is the stationary distribution for } \mathcal{M}.

Proof: since

\[ \mathcal{M}_j^i \phi_i = \mathcal{M}_i^j \phi_j \Rightarrow \sum_j \mathcal{M}_j^i \phi_j = \sum_j \mathcal{M}_i^j \phi_i = \phi_i \sum_j \mathcal{M}_j^i = \phi_i \cdot 1, \]

and this is true \(\forall i\) thus \(\mathcal{M} \phi = \phi\), that is \(\phi\) is the stationary distribution for \(\mathcal{M}\). So, given a distribution, \(\pi_j\), it is possible to construct a Markov matrix with \(\pi_j\) as the stationary distribution, by imposing the condition of detailed balance.

That is, if \(\mathcal{M}_j^i\) are chosen so that \(\mathcal{M}_j^i \pi_i = \mathcal{M}_i^j \pi_j\), and of course subject to the constraints that \(\mathcal{M}_j^i \in [0, 1]\) and \(\sum_i \mathcal{M}_j^i = 1\), and that the matrix is aperiodic irreducible, then \(\mathcal{M}\) is a transition matrix for a Markov chain whose equilibrium distribution is \(\pi\). The details of how one might go about such a construction are given in the Metropolis–Hastings Algorithm [30] [19].

### 3.3.6 Metropolis Hastings

The Metropolis–Hastings algorithm is a Markov chain Monte Carlo method as described previously. The algorithm sets about constructing a Markov matrix which has as its equilibrium distribution some target density \(\phi\), of interest to the operator. The algorithm requires the specification of a proposal
density, $q_j^i$, which is a probability density for $j$ and may depend upon $i$. This is then used in order to propose transitions from $i$. The condition of detailed balance is then imposed in the following fashion.

Construct $\alpha_j^i$ by imposing detailed balance, so that the matrix with entries given by $M_j^i = q_j^i \cdot \alpha_j^i$ is a Markov matrix. This is done as follows:

If $q_j^i \phi_i = q_j^i \phi_j$, then $\alpha_j^i = \alpha_j^i = 1$.

Otherwise, assume (without loss of generality) that $q_j^i \phi_i > q_j^i \phi_j$, then setting $\alpha_j^i = 1$, and constructing

$$\alpha_j^i = \frac{q_j^i \phi_j}{q_j^i \phi_i},$$

detailed balance holds.

Thus, by defining in general

$$\alpha_j^i = \min \left[ \frac{q_j^i \phi_j}{q_j^i \phi_i}, 1 \right],$$
detailed balance is satisfied.

In order that what has been constructed is a Markov matrix which will generate a chain having $\phi$ as the invariant distribution, it remains to show that $M$ is indeed Markov. This imposes conditions on the form of $q$ which is related in turn to $\phi$. The conditions are as referred to before, aperiodicity and connectedness. These are indeed satisfied for quite a large family of densities [30], [54].

The algorithm then, works as follows;

1. Set $i = 0$; Set $N = $ some large value; Choose an initial state $x_0$.  

41
2. Propose \( y \) from \( q_y^{x_i} \).

3. Accept the proposal with probability \( \alpha_j^{x_i} \).

4. If accepted, set \( x_{i+1} = y \), else set \( x_{i+1} = x_i \).

5. If \( i < N \) set \( i = i + 1 \); and back to step 2.

Although theory demonstrates that a chain constructed using this algorithm has a limiting distribution which is the target distribution, the question of the rate at which the limiting distribution is attained is still open.

Note that the samples \( x_0, x_1, \ldots, x_j, \ldots \) generated by the chain will depend upon the choice of \( x_0 \) and only when close to the limiting distribution are the samples to be considered as having come from the target distribution.

What size should \( N \) be, and for what minimum \( j \) should \( x_j \) be considered as a sample from the target? A number of methods have been proposed in order to answer these questions. Diagnostic methods of Gelman and Rubin [18] and others are reviewed by Cowles and Carlin [11]. Murdoch and Green have developed methods of demonstrating convergence [35], but these methods are far less practical than the heuristic diagnostics described elsewhere. A review of methods to date including those of Murdoch and Green is provided by Brooks and Roberts [7].

The Metropolis–Hastings algorithm is valid for sampling from the \( \phi(x) \), for \( x \in \mathbb{R}^n \), that is for a general vector, \( x \). However, in practice it can be more natural to consider \( x \) as the combination of subvectors \( x = \{x_1, x_2\} \). It turns out [9] that a transition matrix for a chain which converges to the target \( \phi(x) \) may be constructed by considering matrices for a chain which samples from \( \phi(x_1|x_2) \) and \( \phi(x_2|x_1) \).
3.3.7 Gibbs Sampling

Efficiency of proposal density is an issue, but where the form of the full conditional distributions is known, these may be used to obtain proposals for the above algorithm.

The special case of the Metropolis-Hastings algorithm, where the proposal density, \( q \) is the product of full conditional distributions is called the Gibbs sampler. For example, consider the case of sampling from a target \( \phi(X, Y) \), with the knowledge of the conditional distributions, \( \phi(X|Y) \), and \( \phi(Y|X) \). Now, since \( \phi(X, Y) = \phi(X|Y)\phi(Y) \), detailed balance holds and the proposal is always accepted. In practice, it is possible that \( \phi(X|Y) \) is known, but that \( \phi(Y|X) \) has to be sampled using more general methods. In this case Gibbs sampling is combined with for example Metropolis-Hastings techniques. Such a sampling method is sometimes referred to as Metropolis-Hastings within Gibbs; although since Gibbs sampling is a special case of Metropolis-Hastings, this terminology is incorrect [9].

3.4 Issues of Convergence

For any of the sampling schemes outlined above, it should be remembered that although the target distribution is the invariant distribution, and that the sequence generated by the algorithms will tend in distribution to the invariant distribution, issues of rate of convergence will be important.

Specifically there are two main important considerations:

1. When will the samples be independent of the initial value, \( x_0 \)?

2. What number of samples, \( N \) are needed?
The first question refers to the fact that $x_0$ is just some (operator chosen), possible value for $X$ and is unlikely to come from the target distribution. Indeed, it may be some time before $x_j$ is from $\phi$, (call this time $J$), only after which time the samples may be used. This time is called burn-in. The chain is said to have converged after time $J$.

While it is possible to determine burn-in exactly in principle, certainly for a limited number of cases [35], analytical methods of determining $J$ are tedious if not wholly impractical. Even in such cases the question then arises as to whether one should use the outputs of multiple chains or a single long chain [36].

For practical applications, time series plots of the chain can give an idea of $J$. In the literature, a review of a number of diagnostic tools is provided in [11] and [7] to assess convergence.

The second question is as to how many samples should be taken. This depends on what the samples are being used for, that is, what is being estimated, and how accurate the estimator needs to be. Of course, $N$ depends on $J$ also, since only $N - J$ samples come from the target distribution.

Again, diagnostics exist for determining how many samples are needed. A comparison of estimates based on two different chains started at different points is one method of checking the variance of the estimators used.

The choice of the proposal distribution is fundamental to the rate of convergence. Common choices for the proposal density include the normal, centred on $x_{old}$, choice of variance to be decided; uniform, centred on $x_{old}$; normal centred on $x_0$; uniform centred on $x_0$. In the case of the last two of these, the proposal, $x_{i+1}$ is independent of $x_i$, and hence they are known as
independence samplers [54].

As well as the question of when the chain has converged, of interest is the rate of mixing of the chain. Mixing is the speed at which the chain explores the target distribution. If the chain mixes slowly, then it requires very many samples to explore the whole support of the target. In the case of the first proposal mentioned, mixing depends upon the variance. The acceptance rate is the number of times a move is made divided by the total number of steps in the chain. If the acceptance rate is too high, this indicates that the chain does not have the opportunity to sample from the tails of the distribution. If the acceptance rate is too low, this indicates that the chain is too stationary, and thus does not move around much. Both these cases would indicate insufficient mixing. Experience has shown that an optimum acceptance rate is between 0.25 and 0.5 for the case of normal target and proposals, with lower rates acceptable for higher dimensions [9].

3.5 Nature of Multidimensional Posterior

Another factor which can seriously affect the rate of convergence, and the mixing of the chain is the shape of the posterior distribution. Jarner [20] has considered in detail the theoretical speed of convergence for different shapes of the posterior, his work building on theoretical work of Roberts [43]. This phenomenon was observed in practice, in the context of the sampling in the models presented later on.

An example demonstrates what occurred. Consider a two dimensional posterior, \( p(x, y|\mathcal{D}) = p(x|\mathcal{D}) \) which is a ridge, as illustrated in Figure 3.1. The methods described sample the posterior, by starting at some location
in the plane, and moving first in the $x$ direction, then in the $y$ direction. Each move is either accepted or rejected. It is clear from Figure 3.1 that the methods will take a substantial amount of time to move from the lower left to the upper right of the target density. In Section 5.6 it is mentioned that the chain took a very large number of steps to examine the posterior, and that this difficulty could have been avoided. The difficulty presented is that the chain takes many steps to sample from the complete support of the distribution; that is the chain does not mix well.

A reparameterisation of the problem allows a much quicker sampling strategy. Consider the reparameterisation shown in Figure 3.2. The transformation $z = Ax$ yields a basis for which the density is now a ridge running in the direction of $z_2$ only, with width along the $z_1$ direction. The samplers now have the ability to traverse the target in just one step of $z$. 

Figure 3.1: Posterior Density in $x$ and $y$. 
Of course, in order to carry out such a reparameterisation, one needs to have knowledge of the posterior distribution. As such, the transformation may not in general be specified a priori. It therefore requires an initial running of the chain with untransformed variables and based on this a transformation can be made. In the case of the models discussed later it was found that the transformation speeded up convergence by an order of magnitude.

3.6 Graphical Representation of a Model

A hierarchical model is one which has an ordered structure, such that a sequence of parameters of interest are conditionally independent of each other. For example, if one has a population of similar specimens, then one might expect that they have similar average crack growth rates. Let the average rate in specimen $i$ be $\lambda_i$, the observed rates at $j$ times be $x_{ij}$ and the overall average be $L$. Then $x_{ij}$ are assumed to be distributed with mean $\lambda_i$, and $\lambda_i$ distributed with mean $L$. Then $f(x_{ij}|L, \lambda_i) = f(x_{ij}|\lambda_i)$ that is that the $x_{ij}$ are conditionally independent of $L$ given the $\lambda_i$. The parameter $L$, refers to the distribution of the parameters $\lambda_i$, and is sometimes called a hyperparameter. Such models are used in many different circumstances to model
population effects [56].

In order to visualise the relationships between different parameters and hyperparameters in a model, a useful tool is the directed acyclic graph representation. Such graphs may be referred to as DAGs, and it is noted that the important properties of these graphs are that the arrows have direction, and that no cycles exist in the graph. Spiegelhalter has been an active proponent of directed acyclic graphs for some time [50], [26].

The DAG in Figure 4.8 demonstrates such a graphical structure. Following [50], \( V \) represents a node from the set of nodes, \( V \). A parent of \( v \) is any node which has an arrow pointing from it to \( v \). A descendant of \( v \) is any node for which a sequence of arrows exists, starting at \( v \) and finishing at the descendant.

A box around a node indicate that it is a constant (or fixed, known) quantity, whereas a circle around a node indicates that the quantity is a random variable. Solid arrows represent a probabilistic dependency, whereas a dotted arrow represents a deterministic relationship. The stacked boxes represent a sequence of plates, which may be thought of as a collection of exchangeable random variables. The graph is structured, so that all dependencies are visible. The model suggests that, conditional on knowing the parents of a node, the random variable is independent of all others in the graph, apart from its own descendants.

A DAG assists one with writing down the distribution of quantities of interest, since the independence structure is given by the graph. Define

\[
V_v = \{ w \in V : w \neq v \}
\]
and

$$v^+ = \{ w \in V : w \text{ is a parent of } v \}$$

Since by definition $v$ is conditionally independent of $V_v$ given $v^+$, in order to factorise the joint distribution of all the parameters one need only consider, for each node, $p(v|v^+)$. This is useful for Gibbs sampling. The computer package WinBUGS [51], allows one to specify certain types of models directly using DAGs.
Chapter 4

Growth Model

The primary cause of failure of metal structures is the growth of a single crack to such a length that the load bearing ability of the overall structure is compromised. For this reason, many models for the growth of cracks have been developed. Chapter 2 deals with some of the previous methods used. Attempts at modelling crack growth using the various deterministic models, such as Paris–Erdogan and Forman equations have had some success, when used to model the growth of long cracks. Indeed, very many variations have been proposed to take into account different features of the data. However, most attempts to explicitly model crack growth have concentrated on cracks which are microstructurally long. As seen, other methods exist for modelling damage without explicit reference to the underlying mechanism.

4.1 Stages of Crack Growth

Most models have been developed for the macro-crack phase of the crack growth process. However, the crack must first be formed, and must develop
to a length that is covered by the model.

As mentioned in Section 2.6, there are a number of stages to crack growth. Initially, a structure may contain no defects. This is called the dormant phase. During this stage the crack does not exist. The next stage is the formation of a crack. This is termed nucleation or initiation. This phase consists of the material gradually deforming, until such time as the molecules of the material have become sufficiently dislocated to the extent that a recognisable defect, or crack has formed.

Then there follows an initial growth phase which is known as microcrack propagation or short crack propagation. This is a haphazard growth, is not well explained by the traditional models described in Chapter 2, and is the area of research interest here.

As a crack grows, it enters the macrocrack phase, by which time the various models in the literature become appropriate. The final phase of interest is sometimes referred to as the failure phase, which is the component or specimen giving up its strength. Failure is generally very quick, since the rate of growth of long cracks is exponential. In comparison, microcracks can be present in a structure for a very large proportion of its lifetime. For this reason, they play an important part in determining total lifetime.

There are a number of qualitative differences exhibited by the microcracks in the specimens that were examined, as compared to what happens for macrocracks. The first of these is the rate of growth. Thus the primary aim for this section of research was to model the rate of growth of microcracks within the specimens, and make reliability predictions based upon these models. This aim fundamentally differs from the work of previous au-
Figure 4.1: Illustration of Different Phases of Crack Growth through the Material Microstructure.

Authors such as Beretta and Clerici [4], who have concentrated on predicting reliability based upon an overall damage statistic.

4.2 Characteristics of Microcracks

Figure 4.1 shows qualitatively the different stages in crack propagation. Within the granular structure of the material, cracks grow reasonably quickly, until they are slowed due to the obstacle of the grain boundary, after which time their rate of growth speeds up once again. To have a clear understand-
ing about what is being analysed, it is useful to clarify what a microcrack is.

**Definition 4.1** A microcrack is defined to be a crack within a specimen which has length of the order of magnitude of the microstructure of the material in which it resides.

The above definition, which comes from the introductory remarks in the work of Miller and de los Rios [32], is necessarily vague. Indeed, there are a number of ways of specifying what constitutes short fatigue crack growth in the literature ranging from the at first seemingly unhelpful [47],

Short crack propagation (SCP): fatigue crack growth which cannot be described by linear elastic fracture mechanics.

to the more useful detailing of typical lengths in terms of grain diameters (< 5 – 15 diameters) for different materials as in Blom [6]. Of course, this analysis is based on the behaviour of cracks within the materials in question, and whether they follow the usual rules for growth. Further, more recent classifications of different types of short crack growth are given in Brown [8] and Zhao et. al. [60].

However, it should be remembered that the primary reason for interest in these cracks is that the models which have been developed for (long) crack propagation do not work. As such, the definition by Smith et al. [47], is in fact more accurate than it first seems.

Brown [8], as well as many other authors cite the original work of Kitagawa-Takahashi [24] who produced a diagram which relates stress level, initial crack length, and type of growth. The theory suggests that cracks below a certain
length have a higher fatigue limit, and do not follow linear elastic fracture mechanics. Developments from the Kitagawa-Takahashi diagram have led to new theories for the classification of short, transitional and long crack growth.

Figure 4.2 shows a plot of rates of growth for short cracks. This is a plot of the average of $\frac{\Delta a}{\Delta N}$ versus $a(N)$ for a set of cracks present in a specimen subjected to cyclic stressing at 140kN. The slowing of the rate of growth is consistent with previous experience, and is the well known grain effect. In this data, the effect of the grain boundary on growth can clearly be seen, by three distinct minima in growth rate.

4.3 Experimental Details

In the laboratory, a prepared specimen was fitted to an apparatus, and then subjected to repeated stressing. The nature of the stressing was controlled, and in this case, the loading was sinusoidal with a maximum stress $S_{\text{max}}$ and minimum stress $S_{\text{min}}$. The specimen was inspected periodically, at which time a replicate tape was placed on the surface, and an image of any cracks present was recorded. At the end of the experiment, the images recorded were subjected to microscopic examination and measurements of cracks recorded.

Data was recorded for three different specimens, at three different stress levels. The first specimen did not fail in the laboratory; the time of failure for the other two specimens was recorded.

The data consist of the lengths of cracks at the final observation point, followed by the lengths of any cracks which subsequently coalesced to form these cracks, recorded at the earlier time points.
Figure 4.2: Plot of Rates of Growth for 140kN Data
One cycle is defined to be the time between the application of $S_{max}$ and the next application of $S_{max}$. It is known that the deterioration of the specimen depends upon the number of times it is stressed. It is possible, therefore, to simulate what might amount to years of use in a real application, in a shorter time in the laboratory, by just increasing the frequency of the stressing. In this case the stressing was carried out at 30Hz. Time, therefore is recorded as number of cycles and represented by $N$, which is defined to be the number of cycles of stress undergone by a specimen.

### 4.4 Data

The data consists of measurements of 435 final cracks for the three specimens. This breaks down as 190 final cracks measured at 4 times for the 140kN data, 185 final cracks measured at 8 times for the 160kN data and 60 final cracks measured at 7 times for the 200kN data. Some final cracks that were measured consisted of a set of coalesced smaller cracks. This was observed in the 200kN data in particular, with each final crack consisting of up to 23 smaller cracks.

The data is plotted in Figures 4.3, 4.4 and 4.5. These are plots of the total length of the cracks which were present at the final observation point, versus number of cycles.

The raw data exhibits a feature known as 'coalescence' which has to be dealt with. Coalescence is the process by which two cracks, which are spatially close, combine to form one single crack. This phenomenon is dependent upon how many cracks are in a specimen, and how they are spatially distributed [4], [60]. This can be an important mechanism for crack elongation,
Figure 4.3: Plot of 140kN Data.
Data at 160kN. Length versus N.

Figure 4.4: Plot of 160kN Data.
Figure 4.5: Plot of 200kN Data.
Table 4.1: Raw Data as for Cases where 2 Cracks Coalesce; Part of 140kN Data.

The primary focus of the original analysis is concerned with modelling growth of single cracks throughout time. Interactions, and thus coalescence, will not be considered directly in this chapter.

### 4.4.1 Treatment of Raw Data - Coalescence

Table 4.1 shows a sample of the raw data as recorded by the engineers. The ‘,’ indicates that cracks did not coalesce by the next observation, whereas the lack of a ‘,’ indicates that they did. In order to fit a model using data from each time, the data must consist of a sequence of observations, one for each observation time for each independent crack.

In other words, it is not possible to analyse data that consists of values for crack $j$ and crack $k$ at timepoint $m$ and then, following coalescence of $j$
and $k$, a single data point at $m + 1$. There are a number of possible options for dealing with this data:

- Ignore the data for cracks which were involved in coalescence, since these grow in a different way from other cracks;

- Construct a model for crack coalescence so that such data can be taken into account;

- Modify or transform the data, so that it may be considered as a set of data of sequential observations from single cracks.

An initial examination of the data suggested that coalescence was not all that common a phenomenon, and where it did occur, it often occurred quite early on in the specimen lifetime.

While it would be preferable to model coalescence, it initially appeared that the extra complexity would not be worthwhile. However, since failure is caused by the largest crack, which itself may have been involved in coalescence, it is important that real data is not just ignored. This meant that the only option left was to transform the data set.

It was decided to use the simplest of adaptations, that is, each crack extant at the final timepoint would be numbered from 1 to $J$. Then if a particular crack had been involved in coalescence at the length used at time $n$ for crack $j$ would be just the sum of the lengths of the coalescing cracks. This is consistent with the approach adopted by authors in the engineering field. An example of this is shown in table 4.2, based upon the data in table 4.1.

Implicit in this adaptation of the data is that the rate of growth of two
<table>
<thead>
<tr>
<th>Crack Name</th>
<th>Time</th>
<th>Crack 1</th>
<th>Crack 2</th>
<th>Length Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dcl6_1</td>
<td>1400000</td>
<td>12</td>
<td>39</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>2100000</td>
<td>54</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>3000000</td>
<td>60</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>4000000</td>
<td>60</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Ec3_2</td>
<td>140000</td>
<td>87, 36</td>
<td></td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>210000</td>
<td>87, 42</td>
<td></td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>300000</td>
<td>145</td>
<td></td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>400000</td>
<td>145</td>
<td></td>
<td>145</td>
</tr>
</tbody>
</table>

Table 4.2: Data as Interpreted for Growth Model.

coalescencing cracks is the same as that of a single crack with length the sum of the lengths of the two cracks. That is, given two cracks, \( C_k \) and \( C_l \), that subsequently coalesce to form \( C_m \), with lengths given by

\[
a_i \equiv a_i(N; \phi_i) \quad \forall i \in \{k, l, m\}.
\]

Then the assumption that the transformed data represents what truly occurs is the same as the assumption that

\[
a_m(N_1) = a_k(N_1) + a_l(N_1) \Rightarrow a_m(N_2) = a_k(N_2) + a_l(N_2),
\]

or equivalently

\[
\int_{N_1}^{N_2} \frac{da_m(N)}{dN} dN = \int_{N_1}^{N_2} \frac{da_k(N)}{dN} dN + \int_{N_1}^{N_2} \frac{da_l(N)}{dN} dN,
\]

which, given \( a_m(N_1) = a_k(N_1) + a_l(N_1) \) implies

\[
\frac{da_k(N)}{dN} + \frac{da_l(N)}{dN} = \frac{da_m(N)}{dN} \quad \forall N,
\]

(4.1)
which is clearly not the case unless \( \frac{da}{dN} \propto a \), which does not hold for this model.

However, even though this assumption is technically incorrect, it was made initially, in the hope that it will not be of substantial importance with regard to the overall reliability prediction.

An extra complicating factor was that the coalescence did not necessarily result in a crack which had length the sum of the constituent cracks. The reason for this was that the length that was measured for any particular crack was determined by a straight line distance. Thus, if two cracks were not parallel, the resultant crack would be shorter than the sum of the constituent parts. This is most apparent in the 200kN data, where many early cracks appear to shrink quite dramatically.

4.5 Model

To date, the majority of models have been developed for macrocrack data. In the graph of rates of growth, observe that the rate of growth for the microcracks tends to slow down, and then speed up again. This is inconsistent with what happens for macrocracks. It is widely recognised that this slowing down is caused by the microstructural properties of the material. The crack hits a grain boundary, and either stops altogether, or is greatly slowed, until such time as it overcomes the boundary.

This phenomenon is modelled, starting from the following assumptions:

1. When the crack gets out of the microcrack phase, the growth rate can be modelled by a macrocrack technique;
2. It is possible to model the deviation of microcrack growth from macrocrack models, by using a collection of random variables;

3. The underlying physical cause for the variation is the grain boundary, and the presence of this boundary should be modelled directly.

Previous work has been done on modelling the effect of the grain boundary. In general, the method is to model the variation in rate from that expected by the macrocrack model according to some function of crack specific parameters, \( \theta \). e.g. Miller [31], Plumtree [39].

One of the main features of the data requiring modelling is the fact that there are a large number of cracks in the specimen. It is assumed that it is possible to encapsulate the information about each crack in terms of parameters \( \theta_i \), and that the cracks are exchangeable, and a hierarchical model for the family of cracks in the specimen may be used.

### 4.5.1 Hierarchical Population Model

In the construction of a model for this data, it is assumed that the cracks are exchangeable, that is, that the length of one crack is conditionally independent on the lengths of the other cracks, given some set of parameters.

It is possible to focus on the growth of a single crack, and to explicitly model that. The model, is that the expected crack length is determined from a rate of growth given by:

\[
\frac{da(N)}{dN} = C(Q\Delta\sigma)^n a^{n/2} \left(1 - \phi \cdot \exp \left(-m \cdot \left(\frac{(a - D)}{D}\right)^2\right) \right).
\]

The observed length \( A(N) \) is normally distributed with mean \( a(N) \). Two types of variance were considered, namely constant variance \( \sigma^2 \) and multi-
The constant variance model was computationally easier, but after consultation with the engineers it was decided to use multiplicative variance.

**Definition 4.2** $D$ is the distance from the initial location to the first grain boundary experienced by the crack.

The model proposed, explicitly takes into account the grain boundary. In the limit, as the length of the crack becomes much larger than $D$, the model tends toward Paris-Erdogan.

**Definition 4.3** $\phi \in [0, 1]$ represents the strength of the first grain boundary. When $\phi = 0$ the first boundary has no effect on growth rate. When $\phi = 1$ the first grain boundary has the effect of stopping the crack altogether.

The effect of $\phi$ on rate of growth is shown for different values of $\phi$ in Figure 4.6. Large values of $\phi$ yield a slower rate for a given length. Thus in the plot, $\phi = 0.5$ is toward the top of the graph and $\phi = 0.9$ is toward the bottom. The plotted values are $\phi = 0.9, 0.8, 0.7, 0.6, 0.5$. The actual values on the $x$ and $y$ axes of the plots give an indication of scale only, and do not refer to any real data.

**Definition 4.4** $a \equiv a(N)$ is the expected length of a crack after $N$ cycles of stress.

**Definition 4.5** $A \equiv A(N)$ is the observed length of a crack after $N$ cycles of stress.
Figure 4.6: Effect of $\phi$ on Rate of Growth.

Figure 4.7: Effect of $m$ on Rate of Growth.
Definition 4.6 $m$ is a measure of the range of the effect of the boundary. If $m$ is large, then the boundary only has a slowing effect, when the tip of the crack is nearby. If $m$ is small, then the boundary has an effect even when the crack tip is far away from the boundary.

In a similar fashion to $\phi$, the effect of $m$ on rate of growth is shown in Figure 4.7. The other parameters are held fixed. Thus in the plot, $m = 10$ is toward the top of the graph and $m = 0.01$ is toward the bottom. The plotted values are $m = 10, 1, 0.1, 0.01$.

Thus, $\phi$ dictates the depth of the trough in growth rate, whereas $m$ how wide it is. It should be noted that since depth is fixed, widening the trough means lowering the rates on either side.

To summarise, the rate of growth is modelled by an adaptation of Paris-Erdogan, such that the observed length after $N$ cycles for crack $i$ within the specimen, $A_i(N)$, is distributed with mean value $a_i(N; \theta_i)$, where $\theta_i$ is a collection of crack specific parameters, $\theta_i \equiv (\phi_i, D_i, m_i)$.

4.5.2 Directed Graph

The relationship between the random variables and observables for the model may be represented by a directed graph as outlined in Section 3.6. This is done for this model in Figure 4.8.

Shown in the directed acyclic graph are the relationships between the various parameters of interest. The crack specific parameters yield, through integration of the differential equation, a deterministic link to the mean crack length $a_i$. The observed data is assumed to be normally distributed around the mean. This is shown as a probabilistic link directed from $a_i$ to $A_i$. The
The model proposes that \( \log(m_i) \) come from a normal with mean \( M \) and variance \( \sigma^2 \), and \( \logit(\phi_i) \) come from a normal with mean \( \Phi \) and variance \( \sigma^2_{\Phi} \). It is assumed that the \( \phi_i \) are exchangeable, which is consistent with a hierarchical population model. This is represented in the directed graph by hyperparameters \( \Phi, M, d, \sigma_{\phi}, \sigma_m \).

The distribution assumed for the hyperparameters is somewhat arbitrary and comes about from consideration of the allowable values for \( \phi_i \) and \( m_i \). Specifically, the \( \logit(\cdot) \) transforms \([0, 1]\) to \([-\infty, \infty]\), and \( \log(\cdot) \) ensures \( m_i \) positive.

Recall that the aim is to estimate the reliability for the specimen, \( R(N) \). Since this is defined as the probability that none of the cracks has reached
the threshold length, $A_{th}$, it may be written;

$$R(N) = P(\max_{\forall i} (A_i(N)) \leq A_{th}) \quad (4.2)$$

It is demonstrated later that the exchangeability within the model simplifies the evaluation of $R(N)$ greatly.

## 4.6 Analysis

The framework for the analysis is Bayesian. The set of all data may be denoted $A$, and the collection of parameters, $\Theta$. As referred to in Section 3.2 the method consists of the following:

1. Construct a model, obtaining a likelihood $p(A|\Theta)$;

2. Elicit a prior density $\pi(\Theta)$;

3. Deduce the posterior density $p(\Theta|A)$;

4. Estimate other quantities of practical interest, such as $R(N)$.

### 4.6.1 Likelihood

From the discussion of the model in Section 4.5 the likelihood may be written down. Let $I = \{1, 2, \ldots, k\}$ be the indexing set for all cracks, and $J = \{1, 2, \ldots, m\}$ be the indexing set for observation times. The data are then $A = \{A_{ij}, \forall i \in I, \forall j \in J\}$ where $A_{ij}$ is the length of crack $i$ at time $N_j$, and so the likelihood is

$$p(A|\Theta) = p(A_{11}, A_{12}, \ldots, A_{1m}, \ldots, A_{km}|\Theta) \quad (4.3)$$
Since the $A_{ij}$ are exchangeable (that is, conditionally independent given the $N_j$ and the parameters $\Theta$), the joint conditional distribution factorises as the products of the conditional distributions, then;

$$p(A_{11}, A_{12}, \ldots, A_{1m}, \ldots, A_{km}|\Theta) = p(A_{11}|\Theta) \cdot p(A_{12}|\Theta) \cdot \cdots \cdot p(A_{km}|\Theta), \quad (4.4)$$

and the $A_{ij}$ are normally distributed with mean $a_{ij}$ and (in the case of constant variance,) variance $\sigma^2$, so the likelihood is the product of Gaussians;

$$p(A|\Theta) = \prod_{i \in I, j \in J} \left( \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2\sigma^2}(A_{ij} - a_{ij})^2 \right) \right). \quad (4.5)$$

While the parameters, $\Theta$, do not appear explicitly in the above expression, they are present in the $a_{ij}$, which are deterministic functions of the crack parameters.

### 4.6.2 Priors for the Parameters

The prior distribution of the parameters has to be specified. This consists of specifying the distributional form of the priors, together with any of the necessary parameters. In the hierarchical model, the crack specific parameters already come from distributions with hyperparameters. It remains to specify and elicit priors for these hyperparameters.

**Priors on $\Phi$, $M$ and $d$.**

Vague priors are used for $\Phi$, $M$ and $d$.

In the absence of evidence to the contrary, it was felt reasonable that the knowledge of the location of the $\phi_i$ and $m_i$ was insufficient to be precise when specifying the distribution or parameters for $\Phi$ and $M$. Vague normal priors were used for these parameters, with zero mean. The prior for $d$ was
lognormal with mean 4.8 and variance 1. This prior was based upon the average grain length for the material under investigation.

4.6.3 Evaluation of the Posterior distribution

The posterior distribution is given by Bayes theorem in the way described in Chapter 3:

\[ p(\Theta | A) \propto p(A | \Theta) p(\Theta). \]

The constant of proportionality is

\[ \frac{1}{\int_{V_0} p(A | \Theta) p(\Theta) d\Theta}. \]

The relationship between the \( A \) and \( \Theta \) is complicated due to the nature of the dependency of \( a_{ij} \) on the \( \theta_i \). It is still possible to sample from the posterior distribution using the methods described in Chapter 3.

4.6.4 Evaluation of Other Quantities

The primary quantity of interest is an estimate of \( R(N) \) as defined in Equation 4.2. In particular it is the posterior estimate,

\[ R(N | A_{ij}, \forall i \in I, \forall j \in J \text{ s.t. } N_j < N) \]

that is of interest, as a function of \( N \).

Using the form of \( R(N) \) in Equation 4.2 this reduces to

\[
R(N | A) = \int_{V_0} P(\max_{v \in I} (A_v(N)) \leq A_{th}|A) \, d\Theta
= \int_{V_0} P(\max_{v \in I} (A_v(N)) \leq A_{th}|A, \Theta) f(\Theta|A) d\Theta
= \int_{V_0} P(A_1(N) \leq A_{th}, \ldots, A_k(N) \leq A_{th}|A, \Theta) f(\Theta|A) d\Theta
\]

71
\[ \int_{V_{0}} f(\Theta|\mathcal{A})d\Theta \prod_{i=1}^{k} (P(A_i(N) \leq A_{th}|\mathcal{A}, \Theta)) \]

\[ \int_{V_{0}} f(\Theta|\mathcal{A})d\Theta \prod_{i=1}^{k} (P(A_i(N) \leq A_{th}|\Theta)) \]

\[ E_{\Theta|\mathcal{A}} \left[ \prod_{i=1}^{k} (P(A_i(N) \leq A_{th}|\Theta)) \right]. \quad (4.6) \]

using the result in Equation 4.4. Since

\[ P(A_i(N) \leq A_{th}|\Theta) = P(A_i < A_{th}|a_i, \sigma^2) \]

is Gaussian, these are easy to evaluate. The expected value in Equation 4.6 can be estimated using samples from the posterior. If there are \( S \) samples from the posterior, \( \{\Theta_1, \Theta_2, \ldots, \Theta_S\} \) then the approximation

\[ E_{\Theta|\mathcal{A}} \left[ \prod_{i=1}^{k} (P(A_i(N) \leq A_{th}|\Theta)) \right] \approx \frac{1}{S} \sum_{s=1}^{S} \left( \prod_{i=1}^{k} (P(A_i(N) \leq A_{th}|\Theta_s)) \right), \]

is used for \( R(N|\Theta) \). This is a kernel density estimate as outlined in Section 3.2.3.

### 4.7 Practicalities

Although the principles of the Bayesian method as outlined in Section 4.6 are straightforward, there are many practical difficulties in the calculation of the details.

Of particular interest are:

- solving for \( a_i(\theta_i) \);
- sampling from the posterior via MCMC;
- issues of convergence of the MCMC sampler;
• correlation between variables and its effect on convergence;

• programming issues.

4.7.1 Evaluation of $a_i$

Since $a_i$ is given as a function of the parameters only in the form of a differential equation, it is necessary to integrate to evaluate $a_i$ for any set of parameters $\theta_i$. This cannot be done analytically, and so numerical techniques are used. The method employed is adaptive stepsize fourth order Runge–Kutta. The algorithm comes directly from the work of Press et. al. [40].

The method was chosen to give a reasonable amount of accuracy, and a reasonable amount of speed. This algorithm, together with the sampling algorithms discussed were programmed in the C programming language. A number of difficulties arose during the course of the programming and were solved by reference to numerical analysis techniques.

4.7.2 Sampling from Posterior

The methods outlined in Section 3.3 are used to sample from the posterior distribution. In particular Metropolis Hastings (Section 3.3.6) and Gibbs sampling (Section 3.3.7) are used.

The posterior distribution may be written $p(\Theta|A)$, where

$$\Theta = (\Phi, \sigma_\Phi^2, M, \sigma_\sigma^2, d, \sigma_d^2; \theta_I) \quad \text{and} \quad \theta_I = \{\phi_i, m_i, D_i\}_{i=1}^I.$$  

The distribution of the hyperparameters is independent of the data, condi-
tional on the crack specific parameters, which yields the following result.

\[
p(\Theta | A) = p(\Phi, \sigma_\phi^2, M, \sigma_m^2, d, \sigma^2; \theta_1 | A) = p(\Phi, \sigma_\phi^2, M, \sigma_m^2, d, \sigma^2 | \theta_1, A) p(\theta_1 | A) = p(\Phi, \sigma_\phi^2, M, \sigma_m^2, d, \sigma^2 | \theta_1) p(\theta_1 | A),
\]

and extending the \( \theta_i \) notation in the obvious fashion,

\[
p(\Phi, \sigma_\phi^2, M, \sigma_m^2, d, \sigma^2 | \theta_1) = p(\Phi, \sigma_\phi^2 | \phi_1) p(M, \sigma_m^2 | m_1) p(d | D_1). \tag{4.7}
\]

The sampling methods as described are used to sample from the posterior distribution of \( \theta_i \) in the following fashion. The actual calculations for the constant and multiplicative variance are provided in the appendix.

### 4.7.3 Algorithm

**Preliminaries**

\[
\theta_i = (\phi_i, m_i, D_i)
\]

\[
p(\theta_i | A) = p(\theta_i | A_{i1}, A_{i2}, \ldots, A_{ij})
\]

\[
p(\theta_i | A_{i1}, A_{i2}, \ldots, A_{ij}) \propto p(A_{i1}, A_{i2}, \ldots, A_{ij} | \theta_i) \pi(\theta_i)
\]

\[
p(A_{i1}, A_{i2}, \ldots, A_{ij} | \theta_i) = \prod_{j=1}^{J} p(A_{ij} | \theta_i)
\]

\[
a_{ij} = a_i(\theta_i; N_j)
\]

\[
p(A_{ij} | \theta_i) = \frac{1}{\sqrt{2\pi \sigma}} \exp \left( -\frac{1}{2\sigma^2} (A_{ij} - a_{ij})^2 \right)
\]

\[
\alpha = \min \left\{ 1, \frac{\pi(\theta_{i\text{new}})}{\pi(\theta_{i\text{old}})} \prod_{j=1}^{J} \left[ \frac{p(A_{ij} | \theta_{i\text{new}})}{p(A_{ij} | \theta_{i\text{old}})} \right] \right\}
\]

The proposal density \( q(\logit(\phi_{\text{old}}), \logit(\phi_{\text{new}})) \sim N(\logit(\phi_{\text{old}}), \sigma_{\phi\phi}^2) \) is chosen for the random walk chain, and similarly for \( \log(m) \) and \( D \).
Algorithm

1. Start chain at $i(0)$ with some initial value; Set $s=0$;

2. Propose $\phi_{i_{\text{new}}}$ from $q(\cdot)$;

3. Evaluate $a_{ij} \equiv a(\phi_i(s), m_i(s-1), D_i(s-1); N_j)$ for each $N_j$ by integration;

4. Accept $\phi_{i_{\text{new}}}(s)$ with probability $\alpha$ then $\phi_i(s) = \phi_{i_{\text{new}}}$ else $\phi_i(s) = \phi_i(s-1)$;

5. Repeat step 2 through 4 for $m_i$ and $D_i$;

6. $s = s+1$;

The integration in the algorithm is evaluated three times per sweep, and this is critical to the speed of the sampler. Evaluation of $\alpha$ is not always necessary and is greatly simplified as demonstrated in the appendix in Section A.5.

Assessing Convergence

Convergence of the Markov chains is assessed by examination of the traces from the sampler. These are included for completeness for a number of runs in the results section.

The chains exhibited very high autocorrelations, which was a feature of the posterior distribution, and for this reason very many samples were taken. The autocorrelation within the chain was reduced by thinning the samples, and only using one in every 10. Consequently, the number of samples needed
for the estimates in the results section was quite large indeed. After a num-
ber of exceedingly long run-times, it was decided to examine the sampling
strategy more closely.

Transforming the Variables

The high autocorrelations were caused by the shape of the posterior distri-
bution, which was a diagonal ridge (approximately cigar shaped) in \((\phi, m)\).

In order to sample more efficiently, the \(m\) and \(\phi\) variables were trans-
formed. An alternative basis \(z = (z_1, z_2)\) as outlined in Section 3.5, would
allow the routine move more efficiently through the posterior.

The transformation takes the form of \(z = A \cdot (\phi, m)\). Sampling from
\(z\) allowed much better mixing for the chains, required less thinning of the
results and so yields faster run times.

4.7.4 Programming

The sampling algorithm was coded in the C programming language. The
programs were compiled using gnu c++ compiler on a machine with dual
Pentium Pro 200MHz processors running FreeBSD 3.2, and subsequently
FreeBSD 3.3 (stable). Programs were tested and debugged using gdb de-
bugger. A set of 5000 iterations of the algorithm for a single crack took 4
minutes 23 seconds.
4.8 Results

The results consist of samples from the posterior distribution, together with kernel density estimates for the reliability $R(N)$.

Figures 4.9, 4.10, 4.11 show typical traces for the MCMC samples from the posterior distributions of $\phi$, $m$, and $D$. These traces are from a single run of the chain for single cracks. It was necessary to draw samples from the posterior distribution for each of the crack specific parameters, that is for each crack in the specimen in order to calculate the kernel density estimates which are provided in Figures 4.12 and 4.13. These kernel density estimates are approximations to the posterior distribution for $R(N)$ for each of the sets of data, that is $R(N|A)$, as outlined in Section 3.2.3.
Figure 4.10: A Typical Trace of Samples for $m$.

Figure 4.11: A Typical Trace of Samples for $D$. 
Figure 4.12: Kernel Density Estimate for $R(N)$ at 140kN. $A_{th} = 1000$
Figure 4.13: Kernel Density Estimate for R(N) at 160kN. $A_{th} = 2000$
4.9 Conclusions

The model provides an adaptation to Paris-Erdogan in order to take into account the variation of the growth rates caused by the first grain boundary.

The reliability predictions were consistent with what was predicted by the model for the 140kN and 160kN data. The specimen at 140kN did not fail. The specimen at 160kN failed at 486900 cycles.

The results for the 200kN weren't so promising. Failure actually occurred between 50000 and 60000 cycles. The 200kN data exhibited coalescence to a greater degree than the other data, and on closer examination it was clear that this phenomenon was the cause of failure at such an early time.

In order to generalise the model further the question of coalescence had to be reexamined.
Chapter 5

Coalescence

As mentioned previously, the primary objective is to model growth of microcracks. To this end, the growth model as detailed was developed. However, an important issue of concern as mentioned in Section 4.4.1 was the treatment of the data when two cracks coalesced.

5.1 Motivation

For a number of reasons, at the outset of the research, the influence of coalescence on the data was ignored. It was initially felt that it was not necessary to include coalescence explicitly in the model. Of course, the model for growth does allow coalesced data to be used, after transformation.

However, it was realised, subsequent to the analysis of the 200kN data set in particular, that coalescence could be a significant factor. The main factor of concern was that the growth only model did not allow for the possibility of coalescence. Thus inferences from the model may be incorrect for the situation where coalescence is a major factor.
In addition to the concerns above, it had been recognised by this author in other research [29] that failure mechanisms are sometimes due to general structural breakdown, caused by large families of cracks throughout a specimen.

5.2 Fatigue in PMMA

During the course of research on the propagation of microcracks in steel, another interesting project presented itself. Brendan McCormack [28] had carried out experimental tests on hip replacements. PMMA is an abbreviation of polymethylmethacrylate, which is a polymer that is used to secure a metal implant to the bone. Damage was caused by cracks appearing within the PMMA, and on the boundaries between the PMMA and the bone and metal. In particular it was noted that damage accumulation and failure could be caused by the initiation of a large number of small cracks throughout the structure, in such a fashion that the cement disintegrated. In such a situation the implant would loosen, causing failure, and would have to be replaced.

This method of failure is different from the case of a dominant crack growing beyond some threshold length, since damage is due to the family of cracks within the specimen rather than one single crack.

The analysis for this situation consisted of modelling the rate of formation of new cracks as a Poisson process, together with some modelling of the underlying growth [29].

While this analysis was for a different material, it still raised important questions about populations of cracks, which in turn raises the question of the possibility of interactions, i.e. coalescence. The next focus, then, in the case
of the steel data was to model the nature of coalescence within a specimen, with a view to improving the reliability predictions. Spatial knowledge would allow direct consideration of \( P(\text{Coalescence | Location} ) \). Previous research revealed models for coalescence, in a variety of different materials, such as for fiber reinforced composites (Yushanov et al. 1995) [59]. Of particular interest was research into coalescence in stainless steel, albeit for different loading at high temperatures [17], which describes some of the ways in which cracks coalesce.

Previous work had been done in examining coalescence for the current data by Taylor et. al. (1995) [53]. This was a simple model which investigated some possible mechanisms underlying crack coalescence. Essentially two cracks were deemed to coalesce when they entered into a zone of influence between them. In the main, this paper recognised also that crack coalescence could be an important factor.

5.3 Data

Spatial data would allow a fuller analysis of the problem, and would allow a spatial model to be applied. However, detailed spatial data was not recorded. When a crack was involved in coalescence, all that was known was that a number of cracks existed before time \( N_j \) and at time \( N_{j+1} \) a single crack remained. The lengths of the constituent cracks, together with the length of the final crack were also recorded.

The form of the data was \( \mathcal{A} = \{ A_{ij} : j \in [1, J], i \in [1, I] \} \) where there were a total of \( I \) cracks recorded at \( J \) timepoints. If the data regarding lengths is ignored for the time being then a pattern of coalescence may be observed.
Figure 5.1: Data, Transformation and Mechanism
Figure 5.1 shows the various interpretations of the data. This data is for crack Dc16.1 from data at 140kN, as described in Table 4.1. The upper quadrant is a plot of observed raw data versus number of cycles. The observed data are marked with an ‘x’. The second plot is of the data used for the growth model. This reflects the transformation of the data as described in Section 4.2.

The third plot demonstrates one way in which to think of coalescence. One considers that the first of the two cracks ‘disappears’ with the second crack taking on a corresponding elongation and subsequent growth. This is shown on the plot with a solid line for the second crack, and a dotted line for the first crack, which disappears at the coalescence time. The question of identifiability complicates issues for this case.

The fourth plot represents the way in which we considered the data after many attempts at modelling the situation. All that is really known is the following:

• Crack one and crack two were measured at the first observation time. This is marked as ‘x’ on the plot. All that can be deduced is that the two cracks coalesced at some time between the first observation time and the second observation time.

• The lengths are unknown at coalescence, but will depend upon the lengths at the first observation time, and the time of coalescence. This is detailed as a circle on the plot.

• The evolution from the time of coalescence to the second observation time is unknown, since no data was recorded during this interval. This,
Table 5.1: Data as Provided by Engineers for a Case where 2 Cracks Coalesce.

<table>
<thead>
<tr>
<th>Crack Name</th>
<th>Time</th>
<th>Crack 1</th>
<th>Crack 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dc16.. 1</td>
<td>1400000</td>
<td>12</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2100000</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3000000</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4000000</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

together with the evolution of each crack up to coalescence are marked with dotted lines on the plot.

• The remainder of the data are marked with a solid line.

By considering the data as a set of cracks which grow during time, and which die at coalescence, it is possible to consider the growth being independent of coalescence. Coalescence is the birth and death effect, whereas the growth effect results only in elongation as previously examined. Graphically, this idea may be represented in a DAG (see Figure 5.2.)

Tables 5.1 and 5.2 show the treatment of the crack data. Instead of considering the cracks as one single crack (growth model) or a combination of cracks, one of which grows, they are considered as cracks which die at coalescence, with a new crack forming. That is, in this case, there are three cracks to consider; two initial cracks and a third crack which is born at the time of death of the other two.
Figure 5.2: DAG to Represent Joint Activity.
### 5.3.1 Preliminary Analysis

Following the thought processes above, and having constructed the model as shown in Figure 5.2, it remained to model the phenomenon of coalescence independently of the growth effect.

A few simplifying assumptions were made, regarding the phenomenon of coalescence, which may be summarised as follows:

- Coalescences were independent of each other;
- The probability of coalescence for a crack was independent of other cracks, given its location and global parameters;
- Coalescences were distributed throughout time, independent of each other.

Examination of the model described by Figure 5.2 shows that what was initially required was to model a rate of coalescence for the cracks. The number of coalescences since the previous interval is given in Table 5.3, for the 200kN set of data.

<table>
<thead>
<tr>
<th>Crack Name</th>
<th>Time</th>
<th>Growth</th>
<th>Coalescence Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crack 1</td>
</tr>
<tr>
<td>Dc16. 1</td>
<td>1400000</td>
<td>51</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2100000</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3000000</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4000000</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2: Data in the Form in which it was Used for the Growth and Coalescence Model Respectively.
<table>
<thead>
<tr>
<th>Time</th>
<th>Total Extant Cracks</th>
<th>Coalescences ($D_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20000</td>
<td>217</td>
<td></td>
</tr>
<tr>
<td>30000</td>
<td>198</td>
<td>19</td>
</tr>
<tr>
<td>35000</td>
<td>174</td>
<td>24</td>
</tr>
<tr>
<td>40000</td>
<td>146</td>
<td>28</td>
</tr>
<tr>
<td>45000</td>
<td>114</td>
<td>32</td>
</tr>
<tr>
<td>50000</td>
<td>79</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 5.3: Number of Coalescences since Last Time Point for the 200kN Data.

### 5.4 Model

A very basic model was used for the coalescence. It was decided to model the rate of coalescence directly. The actual spatial mechanism itself can only be addressed at a later date in the light of spatial data. The rate of coalescence is modelled using a binomial model, with non-constant $p$ as a function of time. The directed graph of this model is shown in Figure 5.3.

The observed number of coalescences at the $i$th stage is $D_i$, and is taken to be binomial. The number of cracks available for coalescence is $K_i$, and $p_i$ is the proportion coalescing. The time of the observation is $T_i$. The proportion, $p_i$ is assumed to be such that $\text{logit}(p_i)$ depends linearly on time, that is

$$\text{logit}(p_i) = \alpha + \beta T_i.$$  (5.1)

A model for coalescence was programmed in WinBugs [51]. The directed acyclic graph associated with the model is shown in Figure 5.3. Convergence
was assessed by visual examination of the chains, and the normal estimates provided from the monitored variables.

5.5 Analysis

The WinBUGS package is a convenient method for carrying out a Bayesian analysis for models containing many variables that can be expressed as a DAG. The model is easily specified in a graphical form by drawing the DAG and specifying distributions and relationships between nodes. While the package cannot be used for the more complicated circumstances encountered for the growth model, due to the computational complexity introduced by the integration of the differential equation to obtain mean length, the coalescence model is very tractable.
5.5.1 Likelihood

The likelihood is just the binomial, with parameters $K_i, p_i$. Let $C = D_2, ..., D_N$ be the data, $C_N = D_2, ..., D_{N-1}$ be the data up to time point $N$, and $\vartheta_i = (K_i, p_i, T_i)$ be the parameters, with $\Theta = (\alpha, \beta)$ being the hyperparameters. Let $\vartheta$ represent the collection of all $\vartheta_i$.

Then, the likelihood is given as

$$p(C|\vartheta, \Theta) = p(C|\vartheta)$$

(5.2)
since $C$ independent of $\Theta$ conditional on $\vartheta$. And, this is just

$$p(C|\vartheta) = \prod_{i=2}^{N} \binom{K_{i-1}}{D_i} p_i^{D_i} (1 - p_i)^{(K_{i-1} - D_i)}$$

(5.3)

Certainly one is interested in predicting $D_N$ given $C_N$ and the $\vartheta_1, \cdots, \vartheta_{N-1}$.

5.5.2 Priors

The prior distributions as initially specified were very flat in the hyperparameters. Closer examination of the situation revealed an experience which is well known, that of specifying priors for parameters, where the true parameter of interest is in fact a transformed variable. For this model, it was desired to specify a prior distribution for the hyperparameters, $\alpha, \beta$, which were related to the logit of the proportion of interest, $p$. Graphically, this observation is shown in Figure 5.4. A central prior for $\logit(p)$ translates as a flat prior for $p$. Very slight difference in location has a noticeable affect on the distribution of $p$. In fact, a moderately high precision on $\logit(p)$ turns out to yield a not too precise prior for $p$. 

92
Figure 5.4: Reason for care in prior specification when transforming variable.
In fact, a certain amount was known about $p$. It was known that $p$
would not be likely to exceed 0.5, and would probably be smaller than that.
With the limited amount of observation, this knowledge should certainly be
incorporated in any specification of the prior. The prior for $\alpha$ was specified
as normal with mean $-0.5$ and variance $0.6^2$ and for $\beta$ as normal with mean
$\frac{2.5}{T_{av}}$ and precision $\frac{T_{av}}{0.25}$, where $T_{av}$ was the average time of observation.

5.5.3 Evaluation of Posterior

After specification of the model in WinBUGS, the analysis is carried out
using the normal MCMC techniques as described in Chapter 3 to sample
from the full conditionals.

As outlined in Section 3.3.7 Gibbs sampling takes from the full conditional
distribution of each of the variables, and the joint posterior is the target
distribution.

For example, letting $p_N^*$ denote \{\(\alpha, \beta, D_1, \ldots, D_N, K_1, \ldots, K_N, p_1, \ldots, p_{N-1}\}\),
that is the parameters, except $p_N$,

$$f(p_N|p_N^*) \propto \left(\frac{K_{N-1}}{D_N}\right) p_N^{D_N} (1 - p_N)^{(K_{N-1} - D_N)} \cdot \pi(p_N^*),$$

and similarly for the other variables.

5.6 Results

The kernel density estimates for the marginal distribution of the various
parameters may be of interest, but as far as model checking is concerned, the
procedure adopted was to censor the final datapoint and make predictions
based on the rest of the data.
That is;

\[
p(D_N|C_N) = \int p(D_N|\theta, C_N)f(\theta|C_N)d\theta
\]

\[
p(D_N|\theta, C_N) = p(D_N|\theta)
\]

\[
\int p(D_N|\theta)f(\theta|C_N)d\theta = E_{\theta|C_N}[p(D_N|\theta)]
\]

(5.4)

And so we use the posterior predictive distribution in order to estimate the value of \( D_N \). Quite favourable results are achieved, as outlined in the Figure 5.5.

The figure shows the posterior kernel density estimate for \( D_6 \) conditional on \( D_1, \cdots, D_5 \). The actual observed value for \( D_6 \) was 35. The mean of the kernel density was 33.4 and the median was 34. This analysis was for the 200kN data. Similar results were obtained for the 140kN and 160kN sets and are included in Figure 5.6 and Figure 5.7.

The actual observed number of coalescences for the 140kN Data was 7. The mean and median from the kernel density estimate were 5.4 and 4 re-
The actual observed number of coalescences for the 160kN Data was 28. The mean and median from the kernel density estimate are 33.2 and 34 respectively. In all these cases the observed value falls comfortably within the 25th and 75th percentiles of the kernel density estimate.

For completeness, the trace for $D_3$ for the 140kN chain is included in Figure 5.8. The chains were similar for the other parameters of interest.

Thus, a straightforward model is employed in the modelling of rate, and
it appears to fit the data which we have. Concern should be expressed about a number of the details; this is done in what follows. However, even this simple model allows a very important insight.

5.7 Conclusions

The model for coalescence is a start. It makes reasonable sense to suggest that in a population of \( K_i \) cracks the number of coalescences can be modelled by a binomial. The implicit assumptions that go along with this include the fact that a coalescence is an event which consists of exactly two cracks merging (or dying) and one crack being born. Where three cracks are observed at \( T_n \) and only one is extant at \( T_{n+1} \), this is described adequately as the combination of two coalescences under this model. That is that two of the initial cracks die, a new crack is born and then subsequently the new crack dies with the third crack and a crack is born from these two. Of course the intermediate crack is not observed.

The information on which the data is based is just the number of coa-
lescences in a particular time interval. The $p_i$ are a function of $T_i$, and this needs further examination. The fact that $p_n$ differs from $p_{n+1}$ raises questions as to how the true underlying rate should be modelled across the time interval, and indeed how real the relationship is. Also, the times of coalescence determine how the cracks grow up to and after coalescence, and these are not observed, but nor are they modelled.

Further investigation of the nature of coalescence was undertaken in the examination of a joint model for coalescence and growth.
Chapter 6

Combining Coalescence and Growth

Having carried on an initial analysis considering growth only, it was clear that coalescence had a critical impact on time to failure. This was also the experience of Taylor et. al. [53] when they analysed the data previously. A difficulty faced by Taylor et. al., was that there is no spatial data available.

6.1 Background

The literature contains a little about coalescence of short cracks in steel. Gao et. al. [17] describe the mechanism of coalescence for high temperature bending fatigue. Reference was also found to the modelling of coalescence phenomena in metal matrix structures, plastics and fiber reinforced composites [59]. It is well recognised that damage accumulation due to a population of cracks is an important mechanism of failure in many different materials [29], [31]. Short cracks are less well understood, and less well researched than
their longer counterparts. They are more difficult to examine (since they are so small, and there are more of them), and they grow in a more haphazard fashion. In practice longer cracks tend to be involved in growth only failure, which is caused by a single crack. In contrast, mechanisms for dealing with damage accumulation due to short cracks have focused on quantities such as microcrack density and total length of cracks per unit area.

The microcrack problem involves a large family of cracks that interact with one another. It is observed that these grow into each other to form a dominant long crack which causes failure. For this reason the modelling of the coalescence phenomenon is at least as important as the model for growth.

Initial attempts included examining the data with the idea that each crack had a certain probability of growth, and a probability of being involved in coalescence. This is similar to the jump models used by [48] in analysing growth of cracks.

Some headway was made in the context of focusing on coalescence, and the directed acyclic graph in Figure 5.2 certainly formed a start of the solution to the problem.

6.2 Data

Some of the conceptual difficulties experienced when attempting to model this phenomenon and analyse the data were in part due to jumping to conclusions when presented with data. In the literature the total length of cracks is commonly analysed. This is how the data was analysed for the growth model. When turning the focus to the coalescence problem the initial raw data that was received from the engineers was reordered, after some
consideration, as described in Table 5.2. Then the data $A$ are a collection of lengths $\{A_{ij}\}$ for each crack $i$, of a total of $I$ cracks, observed at some or all of the times $N_1, \ldots, N_J$. Not all cracks existed at all times, that is, it was possible that the cracks had been involved in coalescence, in which case, the later timepoints would not be relevant, or that the crack had not yet been formed, in which case the earlier timepoints might not be relevant. For the subset of the data included in Table 5.2 there are three cracks, and four timepoints. The data consist of:

$$A = \{A_{11}, A_{21}, A_{32}, A_{33}, A_{34}\}$$ (6.1)

since the first and second crack coalesce between the first and second timepoint. When the 200kN data is so arranged an alternative plot of the data, as compared with Figure 4.5 is obtained. This is plotted in Figure 6.1.

Thus, the number of coalescences, the time interval during which the coalescences took place, the lengths of cracks at the observation time before and after coalescence, together with the details of which crack coalesced with which, form the set of data. Clearly, one may be interested in the context of the modelling problem in the actual time of coalescence and the lengths at coalescence, and while these are not observed, it is still possible to make probability statements about them.

6.3 Model

The growth of cracks is modelled using the growth model. Based upon the coalescence model already discussed, and following the methods used in the analysis of crack formation in PMMA [29], the rate of coalescence is modelled
Figure 6.1: Plot of 200kN Data, Coalescence Taken into Account
using a Poisson process. The modelling of which cracks are involved in coalescence is dependent upon the spatial distribution of cracks, and is really a question of modelling $p(\text{Coalescence}|\text{Location})$. In this context, spatial data was unavailable, and the probability of being involved in coalescence was assumed to be $\propto 1$ for all cracks, that is, assumed to be discrete uniform.

The model used is represented by a directed graph in Figure 6.2 for an example of three cracks involved in a coalescence. The important feature of the DAG is the dependence of the new crack on the two previous cracks, or in particular, the parameters for the previous cracks. There are no cyclic dependencies in the graph.

In the DAG, the model for growth is the same as the model in Chapter 4. The new parameters are $(B_i, X_i, L_i)$ which are the times of birth, death and
length at birth respectively for each crack. The hyperparameters are $\lambda, O$
which represent the rate of a Poisson, and an order for births and deaths,
which also determines which cracks coalesce with which. The times of birth
and death are realisations from the Poisson process. Of course these will not
be observed, but they are an important part of the model.

The dependence of the initial length of crack 3 on the parameters and
initial lengths of cracks 1 and 2 and their parameters is modelled as the sum
of the computed values from the differential equation governing growth at
time of birth with random noise.

The model represented by this DAG is valid for the total data. This would
obviously be much larger than that for just three cracks. The links depend
on which cracks coalesce, and in the full DAG, there are still no cycles.

6.4 Analysis

The DAG allows one to clearly read the dependencies from the model, and
makes it easier to write down the likelihood. The samples from the posterior
distribution are drawn using the numerical techniques already discussed.

6.4.1 Likelihood

As in the earlier situation, note that:

$$f(A_{ij}|a_{ij}, \sigma^2) \sim N(a_{ij}, a_{ij}^2 \sigma^2),$$  (6.2)

Also,

$$a_{ij} = a(\theta_i, L_i; N_j) \quad \forall N_j : B_i \leq N_j \leq X_i,$$  (6.3)
that is, \( a_{ij} \) is a deterministic function of the various parameters obtained by numerically solving the differential equation model. The ordering of coalescence determines that crack \( i \) is the \( k^{th} \) crack to be born, and also which crack dies when crack \( i \) is born. The crack only exists at times between birth and death, and we define

\[
f(B_i|O, \lambda) = T_k \text{ from Poisson Process with rate } \lambda; \quad (6.4)
\]

and

\[
f(X_j|O, \lambda, B_i) = \delta_{Bi} \text{ where crack } i \text{ is born at } X_j. \quad (6.5)
\]

The \( k^{th} \) realisation time from a homogeneous Poisson process is gamma distributed with scale \( \lambda \) and shape parameter \( k \), since it is the sum of exponentially distributed random variables.

Let \( c_i \) be the number of coalescences in interval \( i \), where interval 1 is [20,30], interval 2 is [30,40], interval 3 is [40,50] and interval 4 is [50,60] (where intervals are given in thousands of cycles). Then \( c_i \) are realisations of a Poisson distribution with rate \( \lambda \). That is to say the likelihood for the \( c_i \) is:

\[
p(c_i|\lambda) = \frac{\lambda^{c_i} \cdot e^{-\lambda}}{c_i!} \quad (6.6)
\]

\( O \) is an ordering of the times of birth and death for the different cracks and determines which cracks coalesce. In the context of spatial data, \( O \) would obviously depend upon some parameters representing the spatial nature of the specimen, such as local microcrack density, but this data is not available here.
6.4.2 Priors

If the prior for the rate, $\lambda$ is a gamma distribution, that is:

$$\pi(\lambda) = \frac{b^{-a}}{\Gamma(a)} \lambda^{a-1} e^{-\frac{\lambda}{b}}$$

then the posterior distribution will also be gamma. That is, the gamma is a conjugate prior for this likelihood, which makes the derivation of the posterior distribution a straightforward exercise. A vague prior, $G(2, 20)$ is taken for the rate, $\lambda$.

6.4.3 Evaluation of Posterior

$$p(\lambda | \mathcal{A}) = p(\lambda | c_1, \ldots, c_4)$$

$$\propto \left( \prod_{i=1}^{4} p(c_i | \lambda) \right) \pi(\lambda)$$

$$= \left( \prod_{i=1}^{4} \frac{\lambda^{c_i} e^{-\lambda}}{c_i!} \right) \frac{b^{-a}}{\Gamma(a)} \lambda^{a-1} e^{-\frac{\lambda}{b}}$$

$$\propto \lambda^{(a-1+\sum_{i=1}^{4} c_i)} \exp \left( -\frac{(4b+1)\lambda}{b} \right)$$

Thus the posterior distribution is $G(a_{\text{new}}, b_{\text{new}})$ where $a_{\text{new}} = (a + \sum_{i=1}^{4} c_i)$ and $b_{\text{new}} = \frac{b}{4b+1}$, thus in this case the posterior distribution for $\lambda$ is $G(166, \frac{20}{81})$.

The posterior distribution for the crack specific parameters may be obtained in the same fashion as for the growth model, and are sampled using MCMC. The additional pair of parameters which are sampled for each crack comprise $B_i$ and $L_i$. These are sampled in a similar fashion to the other crack specific parameters since $a(\theta_i, B_i, L_i; N)$ is a deterministic function of these values. Thus $p(B_i | \mathcal{A}, \Theta) = p(a(B_i) | \mathcal{A}, \Theta)$ which is sampled in the same fashion as the other parameters for the growth model.
6.5 Results

The posterior distribution for the rate, $\lambda$ is plotted in Figure 6.3. The trace for $B_i$ and a histogram of samples for $L_i$ are given in Figure 6.4 and Figure 6.5. The value of the sum of lengths of cracks before coalescence is 1014, which is far larger than any of the samples from the posterior. The trace for $B_i$ is consistent with a uniform[35,40].
Figure 6.4: Samples from the Posterior Distribution of $B_1$ for a Single Crack.
Figure 6.5: Histogram of Samples from Posterior of $L_i$ for Crack d1ea.
6.6 Conclusions

The combination of the growth and coalescence models yields some useful results. The posterior distributions of the $L_i$ are interesting, in that they demonstrate that the total crack length after coalescence is often substantially less than the sum of the lengths of the constituent cracks before coalescence. The posterior distribution of the $B_i$ does not differ substantially from the prior. This is likely to be due to the fact that $B_i$ is confounded with $L_i$, since the new crack grows from time of coalescence, a shorter original length will be sampled for an earlier time within the interval and vice-versa. The posterior for the parameters obtained for the growth only model, assumed that there was no coalescence. As such the data was not pure growth, but was rather transformed data. The fact that the joint model separates out the growth effect from the coalescence effect allows a more correct consideration of the data. Cracks which have been involved in coalescence are observed at fewer times, however, so there is less data with which to make accurate estimates of the crack specific parameters.

For computational reasons, and since spatial data are not available, the reliability has not been calculated for this case. However, the reliability of the specimen under this model could, in principle, be estimated, since $R(N)$ is a calculation based upon the crack specific parameters and depends on the order of future coalescences, and which crack coalesces with which. For example, in the case of 10 cracks, let $C_i$ denote the $i^{th}$ crack. One possible future ordering would be $C_{10}$ coalesces with $C_1$ forming $C_{11}$, then $C_2$ with $C_3$ forming $C_{12}$ then $C_{12}$ with $C_4$, and so on until there is only one crack left. This is only one possible sequence of coalescence. Let $o$ be a particular
coalescence event, and let $E$ be the set of all possible future coalescences. Then $E$ is finite, albeit very large indeed. For example, for a set of 10 cracks, $\#E = 2.6 \cdot 10^9$. Then;

$$p(R(N)|A) = \sum_{\forall \varepsilon \in E} p(R(N)|A, O = o)p(O = o|A)$$

$$= \int_{V_0} \sum_{\varepsilon} p(R(N)|A, O = o, \Theta)p(O = o|A)p(\Theta|A)d\Theta$$

$$= \int_{V_0} \sum_{\varepsilon} p(R(N)|O = o, \Theta)p(O = o|A)p(\Theta|A)d\Theta.$$  

And recall that;

$$p(R(N)|O, \Theta) = \prod_{\forall i} p(A_i(N) < A_{th}|O, \Theta), \quad (6.8)$$

where the dependence on $O$ determines which crack coalesces with which. The $p(O|A)$ depends upon the spatial distribution of the cracks. Any information on the location of the cracks is useful in determining this; for example if the data were observed in a number of different regions, it is likely that $p(O|A)$ is close to zero when $O = o$ consists of coalescences of cracks in distinct regions.

A natural extension to the model is the case where the microcrack density (which is a recognised damage function in the literature) is recorded for a number of different regions. Following a similar methodology to [29] it is natural to extend the hierarchical model, by having a rate $\lambda_i$ specific to each region, with an overall rate $\Lambda$.

Thus the proposed joint model constitutes an attempt at modelling the two distinct effects; coalescence and growth. It is a flexible model which may be adapted to the case where spatial data is available.
Chapter 7

Summary & Conclusions

7.1 Summary of Growth Model

Given data on lengths of cracks at a set of times, a differential equation with random parameters is fitted to the data. A posterior estimate for specimen reliability, given experimental data is obtained.

The model is useful in that it provides a very natural way of summarising material properties that have been determined by experiment. The posterior distributions for the hyperparameters may be used for the general modelling of fatigue.

Coalescence turned out to have an important effect in the data which were analysed. For this reason it was decided to investigate methods of modelling this phenomenon.
7.2 Summary of Coalescence Model

The only data on coalescence consisted of the number of cracks that coalesce, and the lengths of such cracks at observation times before and after coalescence. It would be interesting to have some spatial data, and would allow a more thorough and profound analysis. This was not available, however. Hence, a model for the rate of crack coalescence was developed. An inspiration for modelling the phenomenon in this fashion was a statistical analysis of damage accumulation in orthopaedic hip replacements [29].

7.3 Summary of Joint Model

Combining the models for growth and coalescence took some insight. Different methods of reordering the data were tried. Finally a (somewhat natural in hindsight) reordering of the data allowed consideration of growth and coalescence in a reasonable fashion. Probability statements can be made regarding length at time of coalescence and the actual time of coalescence itself. This model is sufficiently flexible to take into account spatial data, should such become available.

7.4 Conclusions

The growth model and coalescence models developed allow direct consideration of the mechanism of growth for short cracks, which takes into account the microstructure of the material. The collection of the data on such small cracks is a lengthy process. In the context of this data, the reliability pre-
dictions are very reasonable. There are other results of interest, such as the empirical distribution of the lengths of long cracks caused by the coalescence of shorter cracks.

7.5 Further Work

Further work could be undertaken in a number of different areas, and indeed this dissertation has led me to ask as many questions as I have answered.

7.5.1 Spatial Aspect

Spatial data would have undoubtedly led to a different treatment of the research question. That said, the lack of spatial data has focused the research to concentrate on deducing as much as possible from the available information. It is acknowledged that the collection of detailed spatial data constitutes a difficult and tedious task. In principle, however, some spatial information could in the future be recorded at the same time as the lengths. Ideally, such information may take the form of detailed locations of the cracks; an alternative might consist of dividing the specimen into regions, noting in which region each crack resides and recording the microcrack density for that region. This would be useful in determining different rates for different regions, and for estimating reliability based on future coalescence. It would also be a much more reasonable expectation of engineering practitioners than the recording of detailed locations. It is possible that an analysis subsequent to the collection of such data may demonstrate that the crack parameters have also a dependence on local region, which may be incorporated into the...
7.5.2 Methodological Issues

There are a number of important questions in the context of any Bayesian analysis which need answering. In particular, the question of sampling from the posterior is still open for further research and is topical in the literature. When using MCMC techniques one has to consider convergence and convergence diagnostics; the feasibility of exact sampling for a particular situation may merit consideration. The nature of the shape of the posterior distribution and how it affects convergence is one which should be examined in detail, for any real application of the techniques.

7.5.3 Factors Influencing Coalescence

The incidence of coalescence varies between the different sets of data. The prevalence of this phenomenon at different stress levels and under different operating conditions could have relevance for real applications, and continuing research in this field is needed.

7.5.4 General Process of Failure

This research is only a part of the material properties work that needs to be done. It is only one model for crack growth, in very controlled circumstances. In the general case of fatigue failure, numerous other factors must be considered.

Other factors which influence failure are loading, environmental factors, and manufacturing details. It is possible of course that very good models...
be developed for any one of these aspects, but the overall uncertainty in the process will depend on all these issues. For example, it is possible that a structure be very well designed, and that the materials used are specified to a high degree of accuracy, and the operating conditions are well known. However, if the structure is to be assembled by a contractor with an unknown degree of reliability, then the high precision of the materials and models will be undermined by the low precision of the contractor.

This point is an important one for anyone involved in reliability prediction.

### 7.5.5 Alternative Growth Models

There are a large number of different models for long crack growth referred to in the literature. These could be modified in a similar fashion in order to model short cracks. The Paris–Erdogan was chosen because of its simplicity, and the ideas herein could easily be extended to more complicated cases.

### 7.6 Closing Remarks

It has already been acknowledged that this research has been a collaborative effort between the Departments of Statistics and Mechanical and Manufacturing Engineering at Trinity College, University of Dublin. The research of damage accumulation in PMMA was carried out in collaboration with the Department of Mechanical Engineering in National University of Ireland, University College Dublin. I feel that it is important to recognise here that there is still much to be gained from future collaboration. The insights that
can be gained in such a context, and the sharing of ideas and knowledge across disciplines are invaluable when modelling real applications.
Appendix A

Calculations

A.1 Bayes Theorem

For completeness, Bayes theorem is included;

**Theorem 3 (Bayes) for events A, B and probability measure P(·),**

\[ P(A|B)P(B) = P(B|A)P(A). \]

This extends to the case of random variables X and Y, which yields

\[ f(X|Y) \propto f(Y|X)f(X). \] (A.1)

*This is the result most often used throughout Bayesian analysis.*

A.2 Inverse Gamma

If X comes from a gamma distribution with parameters \( a \) and \( b \) then X has the following density;

\[ f(X) = \frac{1}{b^a \Gamma(a)} \cdot X^{(a-1)} \exp\left(-\frac{X}{b}\right). \] (A.2)
Consider a random variable $Y$ which is $\frac{1}{X}$. Then the Jacobian is $X^{-2} = Y^2$, so the distribution for $Y$ is

$$f(Y) = Y^2 \cdot \frac{1}{b^2 \Gamma(a)} \cdot \left[\frac{1}{Y}\right]^{(a-1)} \exp\left(-\frac{1}{bY}\right)$$

$$= \frac{1}{b^2 \Gamma(a)} \cdot Y^{-(a+1)} \exp\left(-\frac{1}{bY}\right),$$

(A.3)

and this distribution is termed Inverse Gamma.

### A.3 Constant Variance Growth Model

An often used prior for the variance, $\sigma^2$ also denoted $S$ is inverse gamma, that is to say;

$$f(S) = \frac{1}{b^2 \Gamma(a)} S^{-(a+1)} \exp\left(-\frac{1}{bS}\right).$$

(A.4)

Now the posterior for $S$ may be calculated as follows, where the data, $\mathcal{D}$ consists of $n$ observations;

$$f(S|\mathcal{D}) \propto f(\mathcal{D}|S)\pi(S),$$

and noting that

$$f(\mathcal{D}|S) = \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi}S} \exp\left(-\frac{1}{2S}(A_i - a_i)^2\right)\right].$$

The posterior will also be inverse gamma, as follows;

$$f(S|\mathcal{D}) \propto \left[\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}S} \exp\left(-\frac{1}{2S}(A_i - a_i)^2\right)\right] \cdot \frac{1}{b^2 \Gamma(a)} S^{-(a+1)} \exp\left(-\frac{1}{bS}\right)$$

$$\propto S^{-(a+1)} \cdot \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}S} \exp\left(-\frac{1}{bS}\right) \prod_{i=1}^{n} \exp\left(-\frac{1}{2S}(A_i - a_i)^2\right)$$

$$\propto S^{-(a+1)} \cdot S^{-\frac{a}{2}} \exp\left(-\frac{1}{bS}\right) \exp\left(-\frac{1}{2S} \sum_{i=1}^{n}(A_i - a_i)^2\right)$$

$$= S^{-(a+1+\frac{a}{2})} \exp\left(-\frac{1}{S} \left[\frac{1}{b} + \frac{1}{2} \sum_{i=1}^{n}(A_i - a_i)^2\right]\right)$$

$$\propto S^{-\left(a_{\text{new}} + 1\right)} \exp\left(-\frac{1}{b_{\text{new}}S}\right),$$

(A.5)
where
\[ a_{\text{new}} = a + \frac{n}{2} \quad \text{and} \quad b_{\text{new}} = \frac{b}{1 + b \frac{1}{2} \sum_{i=1}^{n} (A_i - a_i)^2}. \] (A.6)

### A.4 Multiplicative Variance Growth Model

If the variance for the model is multiplicative, then the calculations change slightly. The same prior applies, but the posterior for \( S \) may be calculated as follows, where the number of timepoints observed is \( n \);

\[ f(S|D) \propto f(D|S) \pi(S), \]

and noting that in this case

\[ f(D|S) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi S a_i^2}} \exp \left( -\frac{1}{2 S a_i^2} (A_i - a_i)^2 \right). \] (A.7)

The posterior will also be inverse gamma, since

\[
\begin{align*}
    f(S|D) &\propto \left[ \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi S a_i^2}} \exp \left( -\frac{1}{2 S a_i^2} (A_i - a_i)^2 \right) \right] \cdot \frac{1}{b^a \Gamma(a)} S^{-(a+1)} \exp \left( -\frac{1}{b S} \right) \\
    &\propto S^{-(a+1)} \cdot S^{-\frac{n}{2}} \exp \left( -\frac{1}{b S} \right) \exp \left( -\frac{1}{2 S} \sum_{i=1}^{n} \left( \frac{A_i - a_i}{a_i} \right)^2 \right) \\
    &\propto S^{-(a+1+\frac{n}{2})} \exp \left( -\frac{1}{S} \left[ \frac{1}{b} + \frac{1}{2} \sum_{i=1}^{n} \left( \frac{A_i - a_i}{a_i} \right)^2 \right] \right) \\
    &\propto S^{-(a_{\text{new}} + 1)} \exp \left( -\frac{1}{b_{\text{new}} S} \right) \quad \text{(A.8)}
\end{align*}
\]

where
\[ a_{\text{new}} = a + \frac{n}{2} \quad \text{and} \quad b_{\text{new}} = \frac{b}{1 + b \frac{1}{2} \sum_{i=1}^{n} \left( \frac{A_i - a_i}{a_i} \right)^2}. \] (A.9)
A.5 Acceptance Probabilities

In the context of Metropolis Hastings, in general, an acceptance probability needs to be calculated for the proposed state. Let the target distribution, the posterior, be denoted $f(\cdot | A)$ and the prior distribution denoted $\pi$. Let the parameter or parameters of interest be denoted $\vartheta$, with chain samples of $\vartheta$ denoted $x$. A proposed new state will be denoted $x_{\text{new}}$, and the current state denoted $x_{\text{old}}$. The proposal density is just $q(x_{\text{new}}, x_{\text{old}})$. The data, as before, may be denoted $A$, which consist of lengths for $I$ cracks at times $N_1, \ldots, N_I$.

It is known that the acceptance probability is:

$$
\alpha = \min \left\{ 1, \frac{f(x_{\text{new}} | A)q(x_{\text{new}}, x_{\text{old}})}{f(x_{\text{old}} | A)q(x_{\text{old}}, x_{\text{new}})} \right\}. \quad (A.10)
$$

Using Bayes, the form of the posterior is obtainable from the likelihood and prior as:

$$
f(x | A) \propto f(x | A) \cdot \pi(x),
$$

and the constant of proportionality is

$$
f(A) = \int f(A | x) \pi(x) dx,
$$

which is clearly independent of $x$. Since this is the case, the ratio of the posteriors is the ratio of the products of the priors and likelihoods.

A.5.1 Growth Model for $\phi_i$ with Multiplicative Variance

The $\phi$ parameter in the growth model has a hyperparameter $\Phi$ such that the $\phi_i$ are exchangeable with the following distribution for the parameters:

$$
f(\text{logit}(\phi_i) | \Phi, \sigma_\Phi^2) = \frac{1}{\sqrt{2\pi}\sigma_\Phi} \exp \left( -\frac{1}{2\sigma_\Phi^2} (\Phi - \text{logit}(\phi_i))^2 \right). \quad (A.11)
$$
The proposal density for the logit(\(\phi_i\)) is Gaussian, with mean \(x_{old}\) and sampler variance tuned for acceptance rate. Thus it is symmetric in \((x_{old}, x_{new})\), that is to say that in this case \(q(x_{new}, x_{old}) = q(x_{old}, x_{new})\). Thus the acceptance probability reduces to the ratio of products as above:

\[
\alpha = \min \left\{ 1, \prod_{j=1}^{J} \frac{f(A_{ij}|x_{new})}{f(A_{ij}|x_{old})} \right\}. \tag{A.12}
\]

And the prior comes from the above. That is, the product becomes:

\[
\prod_{j=1}^{J} \left[ \frac{(2\pi \sigma^{2} a_{ij_{new}}^{2})^{-1/2} \exp \left( -\frac{1}{2\sigma^{2}} \left( \frac{A_{ij} - a_{ij_{new}}}{a_{ij_{new}}} \right)^{2} \right)}{(2\pi \sigma^{2} a_{ij_{old}}^{2})^{-1/2} \exp \left( -\frac{1}{2\sigma^{2}} \left( \frac{A_{ij} - a_{ij_{old}}}{a_{ij_{old}}} \right)^{2} \right)} \right] \cdot \frac{1}{\sqrt{2\pi \sigma_{\Phi}}} \exp \left( -\frac{1}{2\sigma_{\Phi}^{2}} (\Phi - \logit(x_{new}))^{2} \right) \cdot \frac{1}{\sqrt{2\pi \sigma_{\Phi}}} \exp \left( -\frac{1}{2\sigma_{\Phi}^{2}} (\Phi - \logit(x_{old}))^{2} \right),
\]

\[
\tag{A.13}
\]

which can be simplified for computational purposes and by cancellation leaving the following:

\[
\left( \prod_{j=1}^{J} \frac{a_{ij_{new}}}{a_{ij_{old}}} \right) \cdot \exp \left( -\frac{1}{2\sigma^{2}} \left[ \sum_{j=1}^{J} \left( \frac{A_{ij} - a_{ij_{new}}}{a_{ij_{new}}} \right)^{2} - \sum_{j=1}^{J} \left( \frac{A_{ij} - a_{ij_{old}}}{a_{ij_{old}}} \right)^{2} \right] \right) \cdot \exp \left( -\frac{1}{2\sigma_{\Phi}^{2}} (2\Phi - \logit(x_{new}) - \logit(x_{old})) (\logit(x_{old}) - \logit(x_{new})) \right).
\]

\[
\tag{A.14}
\]

The acceptance probability is the minimum of this and 1. As an aside, note that when carrying out such calculations, if \(pe^{\alpha}e^{\beta} > 1\) then \(\ln(p) + \alpha + \beta > 0\). In this way, one avoids floating point errors which may occur when evaluating the acceptance probability, as they did for this author.

**A.5.2 Growth Model for \(\phi_i\) Constant Variance**

The calculations for the case of multiplicative variance differ due to the difference in the likelihood and hence posterior. The results regarding the sym-
metricity of the proposal density still hold. Equation A.14 is then:

\[
\exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^{J} (A_{ij} - a_{ijnew})^2 - \sum_{j=1}^{J} (A_{ij} - a_{ijold})^2 \right] \right) \\
\cdot \exp \left( \frac{-1}{2\sigma^2_{\Phi}} \left( 2\Phi - \logit(x_{new}) - \logit(x_{old}) \right) \left( \logit(x_{old}) - \logit(x_{new}) \right) \right)
\]

(A.15)
Bibliography


124


127


