ABSTRACT

Title of Dissertation: FRAMEWORK FOR SMALL FATIGUE CRACK PROPAGATION AND DETECTION JOINT MODELING USING GAUSSIAN PROCESS REGRESSION

Reuel Calvin Smith, PhD, 2017

Dissertation Directed By: Professor Mohammad Modarres, Department of Mechanical Engineering

Engineers have witnessed much advancement in the study of fatigue crack detection and propagation (CPD) modeling. More recently the use of certain damage precursors such as acoustic emission (AE) signals to assess the integrity of structures has been proposed for application to prognosis and health management of structures. However, due to uncertainties associated with small crack detection of damage precursors as well as crack

size measurement errors of the detection technology used, applications of prognosis and health management assessments have been limited.

This dissertation defines a new methodology for the assessment of CPD parameters and the minimization of uncertainties including detection and sizing errors associated with a series of known CPD models that use AE as the precursor to fatigue cracking. The first step of the procedure is defining the separate crack propagation and crack detection models that are to be used for the testing of a joint-CPD model. The two propagation models for this study are based on a Gaussian process regression model that correlates crack shaping factors (CSFs) to the propagation of the crack. One of these propagation models includes a particle filtering technique that includes several AE data. The testing of this joint-CPD model is facilitated by the Bayesian inference of the CPD likelihood where the posterior models are extracted and tested for correctness.

The CSFs, the CPD data, and the AE signal data used for testing of this methodology come from a series of fatigue tests done on dog-bone Al 7075-T6 specimens. The data is first corrected for measurement error that is present based on the initial crack measurements. Then the data is used to generate the prior CPD models that is needed for the Bayesian inference procedure. With the resulting posterior CPD models, a correlation procedure that estimates the CPD model parameters of validation specimens based on the relationship that exists between the CSFs and the CPD model parameters is performed as well as a model error correction procedure. The result of this correlation provides reasonable estimates for the remaining useful life of a given validation specimen.

FRAMEWORK FOR SMALL CRACK PROPAGATION AND DETECTION JOINT MODELING USING GAUSSIAN PROCESS REGRESSION

By

Reuel Smith

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy
2017

Advisory Committee:

Professor Mohammad Modarres, (Chair)
Professor Monifa Vaughn-Cooke
Professor Enrique Lopez Droguett
Professor Aris Christou
Professor Sung Lee (Dean's Representative)

Dedication

To my Lord and Savior,

my father, my mother,

my little sisters,

and family and friends

Acknowledgements

I would like to thank my advisor Professor Modarres for taking me on and encouraging me to continue and complete this research through his guidance and knowledge over the last few years. I would also like to thank Professor Enrique Droguett for his helpful assistance and guidance in implementing particle filtering techniques into my research. In addition, I would also like to thank the rest of the members of my dissertation committee for their support throughout the process.

A special thanks to all of my friends from the Center of Risk and Reliability who have helped me throughout the course of this research as well as my friends from the Center for Minorities in Science and Engineering who have been with me throughout my doctorate education. Finally I thank my family for their support and prayers during my studies and during my whole life.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AE–Acoustic Emission

CP–Crack Propagation

CPD–Crack Propagation and Detection

CSF–Crack Shaping Factors

GPR–Gaussian Process Regression

NDT–Non-Destructive Testing

POD–Probability of Detection

RBE–Recursive Bayesian Estimation

RUL–Remaining Useful Life

Chapter 1: Introduction

1.1 Background and Motivation of Research

Fatigue *crack propagation and detection* (CPD) research has seen significant progress in the field of prognosis and health management (PHM) over the past sixty years. From the crack propagation (CP) model by Paris and Erdogan (Paris & Erdogan, 1963), many other models such as those of Forman (Forman, Kearney, & Eagle, 1967) and Walker (Walker, 1970) have been developed to model CP as a function of material and/or test properties. Some more recent applications directly correlated CP as a function of non-destructive testing (NDT) fatigue markers. For example Keshtgar (Keshtgar, 2013) has established a correlation between initiating CP and certain acoustic emission (AE) signal indices (Keshtgar, 2013) while Naderi et. al. established a correlation with dissipated thermal energy (Naderi, Kahirdeh, & Khonsari, 2012). Similarly, many crack detection models have been proposed as cumulative density functions (CDFs) of the detected crack length. Early works in probability of detection (POD) research assigned a binomial probability distribution to represent POD (Rummel & Matzkanin, 1997), but because POD data highly depend on the type of NDT technique used (Georgiou, 2006), other probability density function (PDF) models such as lognormal and logistic have been proposed among others (Georgiou, 2006). Such variety in CPD modeling has resulted in a lot of options for PHM, the majority of which are based on empirical models, assumptions, and uncertain data and observations.

However, because empirical models are often assumed as a form of CP behavior, many PHM assessments include uncertainties, several of which may not be accounted for. As consequence, engineers may fail to characterize uncertainties and take the observed data

and the model output as the true representation of observed damage. There are three principle types of uncertainties to consider (Sankararaman, Ling, Shantz, & Mahadevan, 2009):

1. Data uncertainty
2. Physical variability uncertainty
3. Model uncertainty/error

Data uncertainty comes from the NDT methods used for crack detection purposes which are notorious for missing extremely small flaw sizes (Thornton & Tiffany, 1970) and for detecting different flaw sizes depending on the method (Georgiou, 2006). This uncertainty in NDT is part of the reason why initial crack sizes are generally unknown (Sankararaman, Ling, Shantz, & Mahadevan, 2009), which Sankararaman, et al. (Sankararaman, Ling, Shantz, & Mahadevan, 2009) consider as being a major reason for inaccuracies associated with the empirical crack growth models. Additional uncertainty comes from the variation in material properties which directly affect the shape and length of the crack such as grain size or inclusions. Variation is also inherent in supposedly static test properties such as loading conditions and test frequency. The final source of uncertainty comes from modeling error or uncertainty, which is the direct result of the selected CP model. When addressing the validity of models in general, British statistician George Box stated:

> "Is the model true?" If "truth" is to be the "whole truth" the answer must be "No." The only question of interest is "Is the model illuminating and useful?" (Box, 1979)

which in the context of this study can be taken to mean that all CP models are going to have modeling error, but some will have less error than others. For example, the AE

intensity dependent CP model possesses a significantly high modeling error precision of 85% (Keshtgar, 2013).

An accounting of all uncertainties within PHM assessments is necessary in order to improve upon existing empirical models. Sankararaman, et al. (Sankararaman, Ling, Shantz, & Mahadevan, 2009) proposed several methodologies for accounting these uncertainties, including applying measurement error correction upon collected data and representing the variation in material properties as distributions. Another key step is the selection of an appropriate CP model that best represents historical fatigue crack data. But Moore and Doherty (Moore & Doherty, 2005) cite that unless model input properties that have a direct bearing on the output are considered, predictions made by that model may still possess model error. Therefore, the appropriate CP model needs to include a firm correlation between observed CP data and test and material properties relevant to that propagation or *crack shaping factors* (CSFs). Such a model was developed by Mohanty (Mohanty, Chattopadhyay, & Peralta, 2011; Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007) that correlates CSFs to the detected crack length by way of a machine learning tool called multivariate Gaussian Process Regression (GPR) (Rasmussen, Evaluation of Gaussian Processes and other Methods for Non-Linear Regression, 1996). Mohanty's input data was limited to fatigue cycles, minimum load, and maximum load and load ratio (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007). The advantage that the GPR model has over most CP models is a stricter adherence to the characteristics of the source data depending on the kernel functions used to train the GPR model (Rasmussen, Evaluation of Gaussian Processes and other Methods for Non-Linear Regression, 1996). The drawback to this model, however, is that its effectiveness is

dependent on the presence of relevant CSFs and the variety therein (Smith & Modarres, 2017) so a generous number of properties and data is usually required.

For this research, the main motivation is to design a methodology that can determine the level of realism that is present in CPD models. For example, when performing a detection analysis on a singular aircraft or a fleet of aircrafts, one needs to address which CPD models best represent the actual rate of growth of a crack in the beam or frame. Naturally, the answer to this will vary from case to case due to differences in material, environmental conditions, or the time of operation. The means to which this determination of realism is implemented takes into account the fact that crack propagation and detection are otherwise related, and should therefore be examined as an integrated model. This relation will aid in determining the realism in CPD models. In addition to this, the methodology will take an adequate account of the uncertainties that CPD models possess and address them as the methodology is implemented.

1.2 Research Contributions

The following is a list of contributions from this research project.

1. Development and validation of a new methodology that will model more realistically an integrated CPD model
2. Introduction and assessment of a list of relevant CSFs that contribute to CPD
3. Establishment of a relationship between CSFs and CPD and the CPD model
4. Demonstrate the relationship between CSFs to the remaining-useful-life (RUL)

1.3 Methodology and Scope of Research

This research is dependent upon the acquisition of a large number of data from a series of fatigue life tests. The first set of data comes from fatigue tests from previous research (Keshtgar, 2013) while additional data comes from new fatigue tests using the same data acquisition methods with some differences (Smith, Modarres, & Droguett, 2017). Therefore, this research continues previous research in order to explore certain aspects that are related to the AE studies (Keshtgar, 2013). In particular, the AE related CP model that was developed by (Keshtgar, 2013) is an important part of this study. The data from that study is made up of a set of CP measurements as well as AE signal data and CSF data (loading conditions and material conditions) that have been identified and selected for this research. A broader probabilistic Bayesian estimation methodology is performed on the acquired data to develop an integrated likelihood model consisting of a CP model and a crack detection model. For example, the likelihood can consist of an exponential (log-linear) CP model (Molent, Barter, & Jones, 2008) and a lognormal crack detection (or POD) model (Georgiou, 2006). All CPD models are chosen from existing models and each CPD pair undergoes the Bayesian analysis. The output from each analysis yields the posterior CPD model parameters and models, the model error, and the remaining useful life (RUL) estimates for each fatigue test specimen. This proposed methodology is depicted in Figure 1-1, where the options for the CP models and the POD models have previously been proposed and tested in other literature (Smith & Modarres, 2017; Smith, Modarres, & Droguett, 2017) including this dissertation. The flowchart outlines the routine used for the processing of the input data (CSFs, CP data, AE signal data, and measurement

error) through the Bayesian inference of a unique CPD model, to produce the desired output.

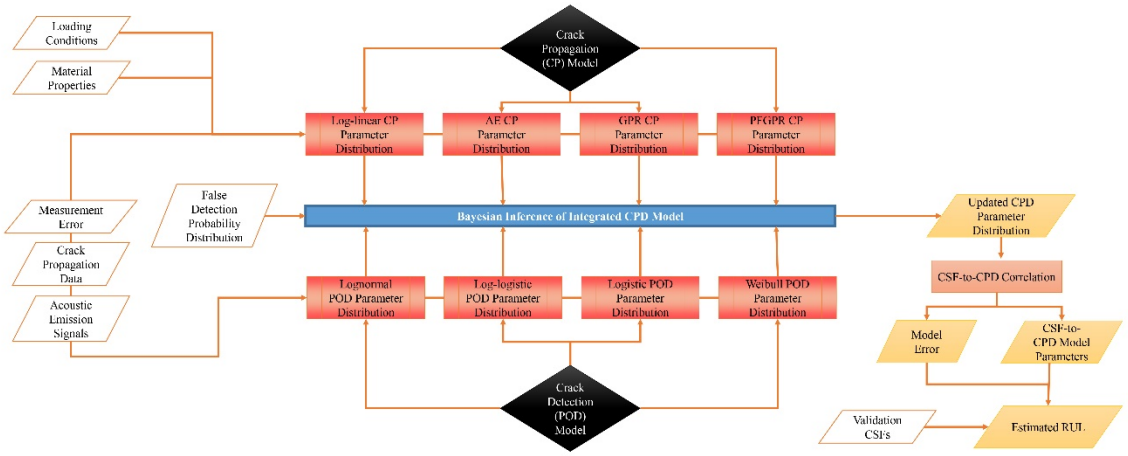

Figure 1-1: Outline of the methodology

1.4 Dissertation Overview

The remainder of this dissertation is divided into five principal chapters. Chapter 2 presents an overview of some of the existing models used for CPD. This literature review includes a comprehensive list of material and test properties that are known to contribute to CPD as well as some of the known uncertainties involved. In Chapter 3 the experimental procedure used for this research is outlined and the Bayesian procedure that is used to process the integrated models is introduced. The test parameters of all of the specimens used to collect the data from the two rounds of tests are defined as well as certain definitions with regard to different crack lengths and measurement errors. The chapter also introduces the Gaussian procedure and components that are necessary for this methodology. Chapter 4 explores further some aspects of the shaping factors including their correlative effect on the model parameters and the uncertainties. It is here that the recursive Bayesian CP

modeling procedure will be expanded as well as the final check for modeling error. The results from the total methodology are presented in Chapter 5 which include the posterior propagation and detection models, the shaping factor to model parameter correlation effects, the validation of this correlation, and finally the end-of-life analysis. Finally, Chapter 6 concludes the dissertation with final thoughts, contributions to the PHM community, and suggestions for future work.

Chapter 2: Background on Crack Modeling

2.1 Overview

To understand where CPD modeling is going, it is often necessary to understand where it has been. CP research has been a study within the PHM community for many years. It is based on the findings of several engineers who subsequently developed and designed CP models based on those findings. Several of these findings relate CP as a function of several shaping factors (CSFs) that are inherent in testing and material properties. We define the CSFs as correlated properties that directly affect the size, shape, and propagation of a crack. As CP research developed, the necessity to detect cracks soon became apparent and research around that advanced in conjunction to CP research. Many techniques in CPD modeling have been designed and while there are inherent strengths in them, there are also considerable weaknesses as well such as inherent modeling uncertainties (Sankararaman, Ling, Shantz, & Mahadevan, 2009). It is when these weaknesses are recognized that these existing models may be improved for the benefit of PHM research.

This chapter is dedicated to a brief overview of the history of CP research as well as how it led to and connects to crack detection research. Models and techniques that have been designed for both will be discussed including what their strengths and weaknesses are in CPD modeling. Additionally, this chapter will cover a list of CSFs that are known to contribute to the shape, propagation, and thus detection of a crack.

2.2 Crack Propagation

2.2.1 Overview and History

The foundation for modern fatigue CP research, more specifically metal fatigue CP research, began in the 1920s under Alan Griffith. He was the first engineer at the time to discover the presence of microscopic cracks as a result of common surface treatment and specimen preparation techniques such as sanding, polishing, and milling; and he was also the first to hypothesize the correlation between fatigue CP and increasing surface energy (Griffith, 1921). George Irwin later linked Griffith's findings to ductile materials as well as brittle metals and further linked CP to increasing strain energy (Irwin, Fracture Dynamics, 1948). The more groundbreaking of his contributions was finding the local stress values $\sigma_{x/y}$ near the tip of a crack (Irwin, Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate, 1957),

$$\sigma_{\frac{x}{y}} = \frac{K}{\sqrt{2\pi r}}\left(\cos\frac{\theta}{2}\right)\left(1 \mp \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) + \cdots \tag{2.1}$$

where K is the stress intensity factor and r and θ are the cylindrical coordinates radius and angle respectively in accordance to the position of a point with respect to the tip of the crack. The foundation of Equation (2.1) is where the equation for stress intensity factor K comes from.

$$K = f(g)\sigma\sqrt{\pi a} \tag{2.2}$$

Note that the term $f(g)$ is a material constant and a represents the crack length.

In time this initial research spurred exploration into other equations for CP which is often denoted as a rate function da/dN where N is the number of fatigue cycles at a given a and dN is the rate change in fatigue cycles (Bannantine, Comer, & Handrock, 1990). The

behavior of CP came to be known as the relation between CP rate and the stress intensity factor range ΔK. Typically this sigmoidal curve is divided into three regions because they represent different phases of the crack's life depicted in Figure 2-1.

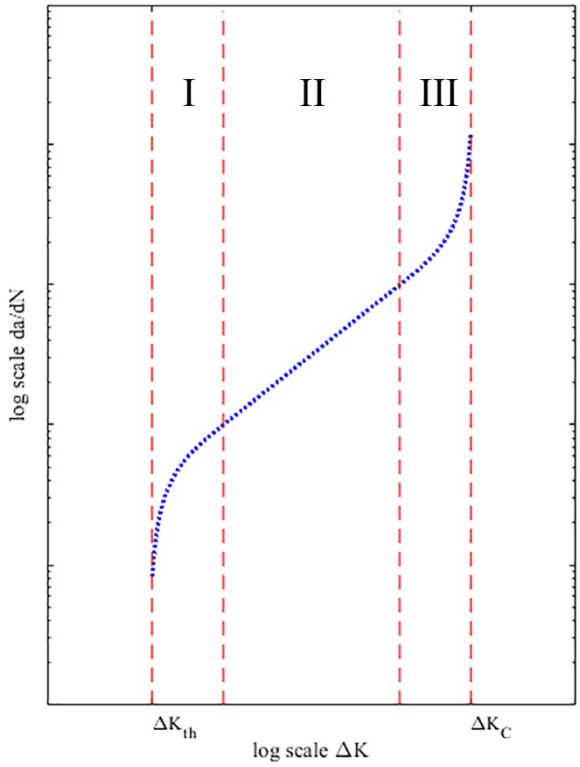

Figure 2-1: A crack propagation rate curve outlining the three principle regions of crack propagation

Region I is known as the threshold region where very minimalist crack growth is exhibited. That is the crack is just beginning to grow at this early phase in life. According to studies, cracks cannot propagate (or propagate extremely slowly) at or below the stress intensity factor range threshold ΔK_{th}. Region II is the region of the most stable crack growth in that the log-log relation between ΔK and da/dN is linear. This region is famously

characterized by the Paris Law which was developed in the early 60's (Paris & Erdogan, 1963).

$$\frac{da}{dN} = C(\Delta K)^m \tag{2.3}$$

Note that C and m are material constants and that the Paris Law is simply a power relation between ΔK and da/dN. Region III is the last region where the most rapid and most unstable crack growth occurs. This is usually as the sample is getting ready to fail or when the crack rapidly approaches its end of life. The term K_c stands for the fracture toughness of the material, at which point the crack is at the end of life.

Of particular importance in CP studies is the small crack growth Region I, because it is the region where the earliest stages of CP take place. Part of this study requires the definition of a "small crack" in regards to the research or material used for the research. However, according to Keshtgar, there is no consensus as to what crack length qualifies as a "small crack," (Keshtgar, 2013) so this value must be arbitrarily selected based on the researchers (Kujawski & Ellyin, 1992). Kujawski and Ellyin state that this measure is based on the diameter of material grains which is variable for different materials (Kujawski & Ellyin, 1992). Keshtgar for example, used $50\ \mu m$ as her upper threshold for small crack measurement (Keshtgar, 2013).

2.2.2 Models for Crack Propagation

The majority of CP research since the inception of Paris Law is focused on modeling the rate of CP. For example, Walker's equation is a variation of Paris Law,

$$\frac{da}{dN} = C\left[\frac{\Delta K}{(1-R)^{1-b}}\right]^m \tag{2.4}$$

where b is another material constant, except that it also considers the effect of load ratio R as well as the effect of stress intensity factor range ΔK on CP rate da/dN (Walker, 1970). And this is just one of many CP rate models that exist (Forman, Kearney, & Eagle, 1967). The majority of CP models may also adhere to a time series. When a CP model is characterized as an integrated probabilistic model, a lognormal PDF,

$$f\left(a|\vec{A},\vec{x}\right) = \frac{1}{a\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln a - \ln g\left(\vec{A},\vec{x}\right)}{\sigma}\right)^2\right] \tag{2.5}$$

can be used to account for it where σ is the lognormal standard deviation parameter, $\vec{A}$ is a vector of parameters that make up the CP model, and the vector $\vec{x}$ is a set of CSFs. However, CP models are not limited to models depicting CP rate. Many CP models operate as simple functions of CSFs, such as the four following models that were given as options in the methodology flowchart in Figure 1-1 (Smith, Modarres, & Droguett, 2017).

2.2.2.1 Mechanistic Approach

The *log-linear* or *exponential CP model* (Molent, Barter, & Jones, 2008; Rusk, 2011; Davidson & Lankford, 2013; Smith, Modarres, & Droguett, 2017) is one of these models which may be expressed as,

$$\ln[a(N)] = b + m\ln N \tag{2.6}$$

where, m and b are parameters representing the slope and intercept of this relationship, respectively. In this form, the initial crack length ($a_0 = a(N = 0)$) may also be defined as e^b. The slope m is actually a parameter that is dependent upon the geometry and the load of and on the crack (Jones, Peng, Huang, & Singh, 2015). Several studies support the position that CP curves can be expressed in exponential form (Molent, Barter, & Jones,

2008; Davidson & Lankford, 2013). Jones et. al for example supports this position by plotting simulated flight hours vs. crack length curves from a range of Australian military aircraft (Jones, Peng, Huang, & Singh, 2015).

2.2.2.2 Acoustic Emission

The *AE propagation model* is a recently proposed model based on a weighted measure of Keshtgar's (Keshtgar, 2013) AE signal the AE intensity $I(N)$, which is a function of two AE signal readings: cumulative counts $C(N)$ and cumulative signal amplitude $A(N)$.

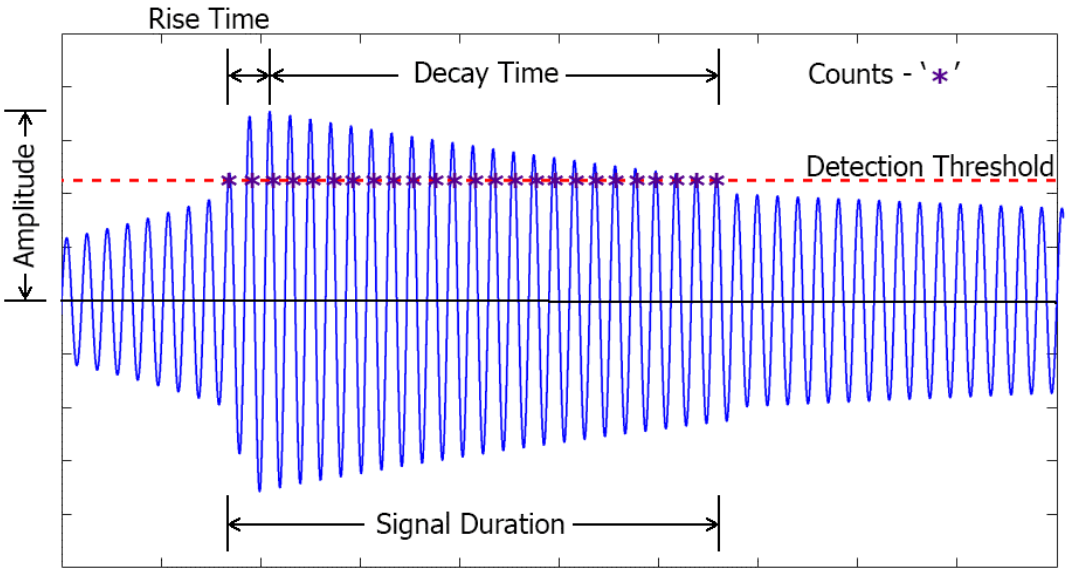

Figure 2-2: A standard depiction of an AE signal waveform (Kappatos & Dermatas, 2007)

The "counts" (or ring-down counts) (Kappatos & Dermatas, 2007) are defined as the number of times the AE signal amplitude rises above a predefined detection threshold amplitude A_0 within the signal waveform duration. The signal duration is the time between when the signal waveform first exceeds the detection threshold and when the waveform goes back under the threshold. The detection threshold amplitude also functions as a noise reduction mechanism as the testing facilities generates a lot of background noise that can

be taken as data. The AE signal represented in Figure 2-2, for example, shows that the signal passes the detection threshold 25 times within the defined signal duration. For this signal duration therefore, there are 25 cumulative AE counts. Both the time-based counts and amplitudes are normalized and cumulative as functions of fatigue cycle N (Keshtgar, 2013). The AE intensity $I(N)$ is defined (Keshtgar, 2013) as,

$$I(N) = \frac{C(N)A(N)}{A_0} \tag{2.7}$$

where,

$$a(N) = h[I(N)] \tag{2.8}$$

Depending on the behavior of the data, Equation (2.8) can take several forms including but not limited to the linear form,

$$a(N) = \alpha I(N) + \beta \tag{2.9}$$

the exponential form,

$$a(N) = \alpha \exp[\beta I(N)] \tag{2.10}$$

or the power form,

$$a(N) = \alpha I(N)^{\beta} \tag{2.11}$$

where α and β are unknown parameters. Initial study of this model shows that AE intensity can effectively detect and potentially measure crack length (Keshtgar, 2013). However, because it is an NDT technique, AE readings are still likely to miss very small crack lengths (Thornton & Tiffany, 1970). Additionally, as stated in the Introduction (Section 1.1) the model has been known to possess a high relative model error (Keshtgar, 2013).

2.2.2.3 Gaussian Process Regression

A more complex propagation model is the *GPR propagation model*, which is based on Mohanty's (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007) work of correlating a set of input variables (fatigue cycles, minimum load, maximum load and load ratio (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007)) to the output variable (fatigue crack length) through multivariate GPR (Mohanty, Chattopadhyay, & Peralta, 2011). Note that the previously mentioned CP models (the power and AE models) only use one CSF: fatigue cycles. This crack length model is a function of a set of CSFs such as: fatigue cycles, load ratio, minimum and maximum loads, frequency, mean grain diameter, and mean inclusion diameter.

$$a = g(\vec{x}) = g([CSF_1 \quad CSF_2 \quad \cdots \quad CSF_Q]) \tag{2.12}$$

where Q is the number of input CSFs being considered for the output crack length a. Note that for each observation data i, there is an output/input data pair of crack length measurement a_i and CSF vector $\vec{x}_i$. The complete set of observation data is typically divided into two groups when developing a multivariate GPR model: (1) the training data which establishes the initial model parameters, and (2) the validation data which validates the model. A multivariate GPR function may now be built primarily for application to the training data. When arranged as a group of M training data, the relations between the $M \times 1$ crack growth training data vector $\vec{a}_t$ and the $M \times Q$ CSF training data matrix $[X]_t$ can be defined as follows.

$$\vec{a}_t = g([X]_t) \tag{2.13}$$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}_t = g\left(\begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_M \end{bmatrix}_t\right) = g\left(\begin{bmatrix} CSF_{1,1} & CSF_{1,2} & \cdots & CSF_{1,Q} \\ CSF_{2,1} & CSF_{2,2} & \cdots & CSF_{2,Q} \\ \vdots & \vdots & \ddots & \vdots \\ CSF_{M,1} & CSF_{M,2} & \cdots & CSF_{M,Q} \end{bmatrix}_t\right) \quad (2.14)$$

The Gaussian crack length PDF, therefore, is depicted as,

$$f\left(a_v|[X]_t, \vec{a}_t, \vec{x}_v, \vec{A}\right) \quad (2.15)$$

for any validation crack length a_v and its validation CSF vector $\vec{x}_v$. The basis for the input/output relation comes from a simple Gaussian relationship $G(0, K)$, where K is the $M \times M$ covariance matrix or kernel matrix that correlates $\vec{a}_t$ and $[X]_t$. Kernel matrices are made up of kernel functions $k(\vec{x}_i, \vec{x}_j, \vec{A})$ which take two sets of CSF data $\vec{x}_i$ and $\vec{x}_j$ and the Gaussian crack length model parameters $\vec{A}$ to produce one element of the kernel matrix. In Gaussian modeling the objective is to develop a kernel function k based on the assumptions of the input and output relation being modeled (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007). There is an extensive list of kernel functions that can be used to fit a given model (Rasmussen, 1996; Rasmussen, Nickisch, & Williams, 2015), however, proper development requires trial and error as well as validation to see whether or not the kernel function is the best fit to the data (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007). The $M \times M$ training kernel matrix may be denoted as $[K]_t\left([X]_t, \vec{A}\right)$.

2.2.2.4 Particle Filtering

The final CP model based on particle filtering (PF), a recursive Bayesian estimation technique that is suited for non-linear processes. Recursive Bayes estimation is a probabilistic inference method in which a set of unobserved values are estimated based on

a set of observed values (Rabiei M. , 2011; Rabiei & Modarres, 2013). The *PF propagation model* can be characterized by the state-space model illustrated in Figure 2-3.

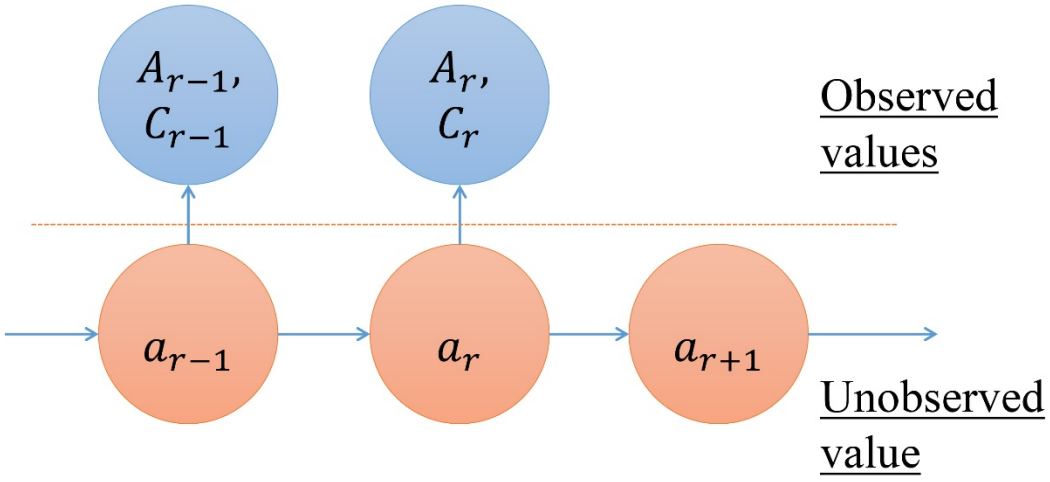

Figure 2-3: The dynamic state-space representation of the particle filtering propagation model where r is a time-step

Since crack length a is largely unknown, it is the unobserved value in this model. Whereas the AE variables cumulative count C and cumulative amplitude A are known, so these are the observed values in the model. The mathematical form for this probabilistic propagation model therefore is,

$$\pi(a_r|C_{1:r}, A_{1:r}) \tag{2.16}$$

which estimates the distribution of crack length a at time-step r based on all C and A values up to that time-step. Equation (2.16) is obtained by way of Bayes Rule,

$$\pi(a_r|C_{1:r}, A_{1:r}) = \frac{\pi(C_r, A_r|a_r) \times \pi(a_r|C_{1:r-1}, A_{1:r-1})}{\pi(C_r, A_r|C_{1:r-1}, A_{1:r-1})} \tag{2.17}$$

where the prior distribution is given as,

$$\pi(a_r|C_{1:r-1}, A_{1:r-1}) = \int \pi(a_r|a_{r-1}) \times \pi(a_{r-1}|C_{1:r-1}, A_{1:r-1}) da_{r-1} \tag{2.18}$$

and the normalizing factor is,

$$\pi(C_r, A_r|C_{1:r-1}, A_{1:r-1}) = \int \pi(C_r, A_r|a_r) \times \pi(a_r|C_{1:r-1}, A_{1:r-1}) da_r \tag{2.19}$$

Because exact calculation of Equations (2.17)–(2.19) is either not possible or very complex, the propagation model is approximated by generating a set of N_p crack length samples (or particles) and their associated weights w_r^i for each time step r. This approximation is depicted as,

$$\pi(a_r|C_{1:r}, A_{1:r}) \approx \sum_{i=1}^{N_p} w_r^i \delta\left(a_{0:r} - a_{0:r}^i\right) \tag{2.20}$$

where δ is the Dirac delta function and the weights are normalized such that sum from $i = 1 \ldots N_p$ is 1. The weights for each of these particles are chosen by sequential importance sampling (SIS) a Monte Carlo method (Doucet, De Freitas, & Gordon, 2001; Arulampalam, Maskell, Gordon, & Clapp, 2002). The method is based on the idea that while direct particle extraction from the target distribution Equation (2.20) is difficult, extraction *is* possible from a proposed distribution $q(a_r|C_{1:r}, A_{1:r})$ called the importance distribution. From this idea the weights (called importance weights) may be defined as the ratio between the target distribution and importance distribution.

$$w_r^i \propto \frac{\pi(a_r|C_{1:r}, A_{1:r})}{q(a_r|C_{1:r}, A_{1:r})} \tag{2.21}$$

The importance weight Equation (2.21) may be further simplified by factorizing the importance distribution,

$$q(a_r|C_{1:r}, A_{1:r}) = q(a_r|a_{1:r-1}, C_{1:k}, A_{1:r}) \times q(a_r|C_{1:r-1}, A_{1:r-1}) \quad (2.22)$$

and the Bayes form of the PF propagation model Equation 2.17,

$$\pi(a_r|C_{1:r}, A_{1:r}) = \frac{\pi(C_r, A_r|a_r) \times \pi(a_r|a_{r-1})}{\pi(C_r, A_r|C_{1:r-1}, A_{1:r-1})} \times \pi(a_r|C_{1:r-1}, A_{1:r-1}) \quad (2.23)$$

or,

$$\pi(a_r|C_{1:r}, A_{1:r}) \propto \pi(C_r, A_r|a_r) \times \pi(a_r|a_{r-1}) \times \pi(a_r|C_{1:r-1}, A_{1:r-1}) \quad (2.24)$$

Therefore, Equation (2.21) is redefined as,

$$w_k^i \propto \frac{\pi(a_r|C_{1:r-1}, A_{1:r-1})}{q(a_r|C_{1:r-1}, A_{1:r-1})} \times \frac{\pi(C_r, A_r|a_r) \times \pi(a_r|a_{r-1})}{q(a_r|a_{1:r-1}, C_{1:r}, A_{1:r})} \quad (2.25)$$

and simplified to,

$$w_r^i \propto w_{r-1}^i \times \frac{\pi(C_r, A_r|a_r) \times \pi(a_r|a_{r-1})}{q(a_r|a_{1:r-1}, C_{1:r}, A_{1:r})} \quad (2.26)$$

Further simplification of Equation (2.21) results from assuming that the importance distribution $q(a_r|a_{1:r-1}, C_{1:r}, A_{1:r})$ is equal to the prior distribution $\pi(a_r|a_{r-1})$ such that,

$$w_r^i \propto w_{r-1}^i \times \pi(C_r, A_r|a_r) \quad (2.27)$$

It has to be noted though that the traditional SIS approach eventually iterates to a point when all but one importance weight is negligible. This is called the "degeneracy problem" (Doucet, Godsill, & Andrieu, 2000) and is generally rectified by generating new particles to replace those with negligible weight.

The resulting propagation is then treated as output for the GPR propagation model with the CSFs as input again, so in that sense this is really a *PF/GPR Propagation Model*. This iteration of the GPR propagation model is more effective than the three first models. CP

is approximated as a Bayesian function of AE readings and then that same propagation is correlated to the CSFs by way of the GPR model. The PF/GPR propagation model takes the most effective aspects of existing models (AE data and CSFs) and relates CP with them all (Smith, Modarres, & Droguett, 2017).

2.3 Crack Detection

2.3.1 Overview and History

As CP study was developing in the late ‘60s, at the same time that the space exploration program was steadily developing, it was becoming more and more pertinent to study the probability of detecting cracks (Georgiou, 2006). The majority of the groundwork in this field was conducted by the aerospace industry, most notably the National Aeronautics and Space Administration (NASA), as the need to ensure structural integrity of critical components became more and more important (Petrin Jr., Annis Jr, & Vukelich, 1993). This was because NASA was beginning development of the space shuttle program where flaw detection and fracture control protocol was of utmost importance in order to prevent catastrophic structural failure due to crack initiated fractures (National Aeronautics and Space Administration, 1971). The early studies of flaw detection probability placed emphasis on finding the smallest detectable crack or flaw length by means of existing NDT methods like radiographic, ultrasonic, X-Ray, and magnetic particles. However, it was becoming more apparent that very small flaws are extremely difficult to detect using NDT methods. As a consequence of this, it was made bad policy to assume that potentially catastrophic flaws would be able to be found or detected at all times (Thornton & Tiffany, 1970). This all led to NASA seeking information about the largest flaw that could possibly

be missed as it is easier to detect a large flaw by way of NDT methods (notably well before critical flaw size) (Georgiou, 2006; Thornton & Tiffany, 1970).

These early flaw detection probability concepts were called POD by NASA in 1973 (Rummel & Matzkanin, 1997). Since this inception, aerospace researchers have worked to obtain POD data for the various NDT methods although initially the distribution for the POD of all flaw types and sizes for all NDT methods was assumed to be binomial (Georgiou, 2006; Rummel & Matzkanin, 1997). By the mid-80s however, POD soon became known as a function of flaw size a or $POD(a)$, thus additional POD distribution types were considered including the lognormal distribution and the log-logistic distribution (Georgiou, 2006). Furtherance of CPD modeling has gone hand-in-hand as a result.

2.3.2 Models for Crack Detection

The crack detection models (or POD models for crack) are all in form of cumulative density functions of crack length a. Finding the model parameters is done by identifying the crack lengths between a_{lth} and a_{hth},

$$a_{lth} < a < a_{hth} \tag{2.28}$$

where $a_{lth}$is the smallest crack length that can be detected and a_{hth} is largest crack length that can be missed using the an NDT technique (Georgiou, 2006). These crack lengths can then fit to a signal response POD function for NDT variable Z (Georgiou, 2006),

$$POD(a) = 1 - F\left[\frac{\ln \hat{Z}_{th} - \ln[Z(a)]}{\sigma_Z}\right] \tag{2.29}$$

where σ_Z represents the standard deviation associated with the error between log forms of model NDT variable $\hat{Z}$ and the true NDT variable Z (Georgiou, 2006; Keshtgar, 2013) written as,

$$\ln \hat{Z} = \ln Z + NOR(0, \sigma_Z) \tag{2.30}$$

The $\hat{Z}_{th}$ term in Equation (2.29) is the NDT threshold where above this value cracks are detected and below this value the crack goes undetected. For example, in the case of an AE based analysis, AE Intensity I may be used as the NDT variable to estimate the POD based on Equations (2.29) and (2.30).

2.3.2.1. Lognormal

As stated, the *lognormal POD model* was one of the first standards developed after the binomial model became impractical (Georgiou, 2006). The form of this model may be represented as,

$$POD(a|\zeta_0, \zeta_1, a_{lth}) = \int_{a_{lth}}^{a} \frac{1}{(x - a_{lth})\sqrt{2\pi\zeta_1^2}} \exp\left\{-\frac{1}{2}\left[\frac{\ln(x - a_{lth}) - \zeta_0}{\zeta_1}\right]^2\right\} dx \tag{2.31}$$

where the parameters ζ_0 and ζ_1 represent the log mean and standard deviation elements. The random variable a has to be additively adjusted as $(a - a_{lth})$ for all POD models because $0 \leq POD(a) \leq 1$ for crack lengths greater or equal to the lower crack length threshold a_{lth}.

2.3.2.2. Log-logistic

The *log-logistic POD model* (sometimes dubbed the "log-odds" model) was also among the first standards developed after the binomial POD model (Georgiou, 2006) and is defined as,

$$POD(a|\beta_0, \beta_1, a_{lth}) = \frac{\exp[\beta_0 + \beta_1 \ln(a - a_{lth})]}{1 + \exp[\beta_0 + \beta_1 \ln(a - a_{l\,th})]} \tag{2.32}$$

where β_0 and β_1 are model parameters. The log-logistic and the lognormal detection models are among the most commonly used POD models (Georgiou, 2006). The log-logistic model is especially prevalent primarily because of its simple mathematical form, and because of its ease of use when dealing with censored data (Georgiou, 2006; Singh, Warsono, & Bartolucci, 1997). However, just because they are common does not make them the norm for all fits.

2.3.2.3 Logistic

The *"logistic" POD model* for example was first proposed by Yuan et al. as a model to represent in-service inspection data (Yuan, Mao, & Pandey, 2009). This model takes a different form than the standard logistic CDF function as shown by the equation,

$$POD(a|\eta_0, \eta_1, a_{lth}) = 1 - \frac{1 + \exp(-\eta_0\eta_1)}{1 + \exp[\eta_0(a - \eta_1 - a_{lth})]} \tag{2.33}$$

where η_0 and η_1 are the model parameters or more specifically parameters that control the detection quality of crack lengths (Yuan, Mao, & Pandey, 2009). As with all POD models, the nature of these model parameters will depend on the NDT procedure being used for detection (Georgiou, 2006).

2.3.2.4. Weibull

The Weibull POD model is another out of the norm detection model represented by,

$$POD(a|\alpha_0, \alpha_1, a_{lth}) = 1 - \exp\left[-\left(\frac{a - a_{lth}}{\alpha_0}\right)^{\alpha_1}\right] \quad (2.34)$$

where α_0 and α_1 are the scale and shape Weibull parameter respectively. It was Bencala and Seinfeld who brought to attention that the Weibull distribution (as well as the Gamma distribution) is capable of replicating lognormal distribution data to some degree (Bencala & Seinfeld, 1976). Since both the lognormal and log-logistic distributions have been adopted as principle POD models, other researchers such as Sekine, Mao, and Rountree (Rountree, 1990; Sekine & Mao, 1990) have adopted the Weibull distribution as a viable POD model. Disregarding the similarity between the lognormal and Weibull distributions however (Bencala & Seinfeld, 1976), Rountree cites Schleher's research (Schleher, 1976) which states that the lognormal POD tends to overestimate some detections making it often the worst-case scenario POD model to use (Rountree, 1990). Hence the Weibull distribution was seen as an option for representing flaw detection in this case (Rountree, 1990).

2.3.3 False Detection

Unlike the governing models for detection which consist of two or more parameters, the probability of false detection P_{FD} can be represented by a single parameter. "False detection probability" is a term that is synonymous with "false positive probability (Wacholder, Chanock, Garcia-Closas, El Ghormli, & Rothman, 2004)" which in this case means the probability that a detected crack or flaw isn't really there or that it was a false-alarm detection. Some very basic probability analysis is applied to obtain P_{FD}. First two

hypotheses are defined as H_0 and H_1. H_0, the null hypothesis is when a crack is detected ($D = 1$) whereas H_1 the alternate hypothesis, is if the crack measurement is actually bigger than the low crack threshold a_{lth}. Second, the truth and falsity of these hypotheses are the basis for the probability of false detection as well as the probabilities of true detection, false non-detection, and true non-detection; all of which are presented as follows:

$$True\ Detection = \Pr(D = 1|a > a_{lth}) \Pr(a > a_{lth}) \tag{2.35}$$

$$False\ Detection = \Pr(D = 1|a < a_{lth}) \Pr(a < a_{lth}) \tag{2.36}$$

$$True\ Non - Detection = \Pr(D = 0|a < a_{lth}) \Pr(a < a_{lth}) \tag{2.37}$$

$$False\ Non - Detection = \Pr(D = 0|a > a_{lth}) \Pr(a > a_{lth}) \tag{2.38}$$

Note that the sum of Equations (2.35) and (2.36) represent the total probability that $D = 1$ or $\Pr(D = 1)$; likewise, the sum of Equations (2.37) and (2.38) represent the total probability that $D = 0$ or $\Pr(D = 0)$. The probability of false detection P_{FD} is depicted simply as,

$$P_{FD} = \Pr(D = 1|a < a_{lth}) \Pr(a < a_{lth}) \tag{2.39}$$

2.4 Crack Shaping Factors

The term *"crack shaping factor"* (CSF) is a terminology that comes from the term performance shaping factor (PSF), a term typically used in human reliability assessment (HRA) when gaging the performance of a worker. Here "crack shaping factor" refers to a unique property that gages the performance or propagation and detectability of a fatigue crack (Smith, Modarres, & Droguett, 2017). Although the term is fairly new, the study of

determining CSFs goes back to the time inception of fatigue CP research. The following is a comprehensive list of several CSFs:

- **Fatigue Cycles**–Time is among the most prevalent of units in modeling the propagation of a crack. Therefore, any time-dependent variable (fatigue cycles, minutes, flight hours, days, etc.…) may be used to model CP. All CP models are governed by a time-dependent variable whether it is propagation form $a(N)$ or propagation rate form da/dN.
- **Load Ratio Effects**–As Forman (Forman, Kearney, & Eagle, 1967) and Walker's (Walker, 1970) research details, the applied load ratio R has a very significant effect on CP rate and thus the crack length or shape. Since the load ratio is a ratio of the minimum and maximum loads, (or a ratio of the minimum and maximum stress intensity factors) they are considered to be CSFs (Paris & Erdogan, 1963; Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007). Generally, the higher the load ratio is, the faster CP takes place; however, this is dependent upon material CSFs.
- **Load Frequency**–Load frequency is a major CSF that affects CP; however, its effect has been known to vary depending on the material. In general decreasing load frequency causes an increase in CP rate and a decrease in fatigue life (Bannantine, Comer, & Handrock, 1990). However, in some instances (such as in polymers) increasing load frequency can result in the CP rate being increased (Hertzberg, Manson, & Skibo, 1980), decreased (Skibo, 1977), or unchanged (Hertzberg, Manson, & Skibo, 1975).

- **Material Grain Size**–Neindorf and Halon et. al. have cited that the overall CP rate is known to decrease as the grain size increases and vice versa (Niendorf, Rubitschek, Maier, Canadinc, & Karaman, 2010; Hanlon, Kwon, & Suresh, 2003).
- **Inclusions**–Inclusions are defined as type of non-metallic deformity or aberration in metallic materials that can be harmful to fatigue properties. For example, MacKenzie cites that the concentration and size of material inclusions can reduce the ductility of steels (MacKenzie, 2008). As inclusions are based on a corrosive effect which occurs either by handling or preparation of the material, they are known to adversely affect CP (Randelius, 2008; Ekengren, 2008).
- **Temperature**–The temperature of the environment has a strong effect on CP particularly in the Region II area. The studies of Nelson et. al. for example show that the CP rate da/dN has a log-linear relationship with the inverse of temperature (Nelson, Williams, & Tetelman, 1971). That is for rising temperature the CP rate also increases thus reducing fatigue life.
- **Pressure**–Williams cites his own research which correlates pressure, as well as temperature, to CP (Williams, 1973). Along with increasing temperature, increasing pressure contributes to a deterioration in fatigue life and da/dN can be simplified to an exponential function of temperature and pressure (Williams, 1973).
- **Dissipated Energy/Entropy**–Although it is a product of CP, dissipated energy or entropy is still correlated to it, and therefore may be used as a means to tracking CP. Early research by Bao et. al. concluded that entropy may be treated as a cumulative value (like the cumulative AE count and AE amplitude) in relating to CP behavior (Bao, Peng, Cong, & Wang, 2010).

2.5 Uncertainties

As stated in Chapter 1, CPD modeling has always been prone to several uncertainties for various reasons. Identifying the three principle uncertainties associated with CPD modeling (Sankararaman, Ling, Shantz, & Mahadevan, 2009) is a necessary step in any PHM research. CPD models are primarily affected by the detected data gathered from NDT methods which is a prominent data uncertainty. This is confirmed in Georgiou's assessment of flaw detection data obtained by the Nondestructive Testing Information Analysis Center (NTIAC) in which several types of NDT methods are used to detect crack lengths from the same flaw specimens (Rummel & Matzkanin, 1997; Georgiou, 2006). In that assessment, the different NDT method results from the same operator given the same flaw specimen is documented. The result of that assessment shows that the POD curves are different for each NDT method used (Rummel & Matzkanin, 1997; Georgiou, 2006). Despite the differences Georgiou insists that one should not claim that on POD curve or set of data is better than another (Georgiou, 2006). It *is* important however, to have a sufficient collection of data since a lack thereof can contribute to data uncertainty in modeling (Sankararaman, Ling, Shantz, & Mahadevan, 2009). A good example comes from this research, where some specimens only possess five to seven AE data while others provide up to twenty-thousand AE data for the purpose of modeling. It is also important to see that the data detected is correctly measured. While different NDT methods may detect and measure different crack sizes, there is a definite size of the detected crack which needs to be confirmed in order to address this data uncertainty. This is also true for the measured CSFs and other material properties which require accurate instrumentation in order to minimize data uncertainty.

Physical variability uncertainty applies to the material properties that contribute to CP. CSFs such as material grain sizes and inclusions are neither constant throughout a single sample, nor throughout a set of samples. Because of this it is strongly advised that these CSFs be represented as probabilistic distributions (McDonald, Zaman, & Mahadevan, 2009; Sankararaman, Ling, Shantz, & Mahadevan, 2009). This also applies to all the environmental CSFs (temperature, pressure, etc.) that are entirely subject to fluctuation during fatigue testing. In some cases, this is also an applicable step in reducing physical variability uncertainty in CP models. The log-linear CP model for example the initial crack length a_0 at $N = 0$ is e^b. This initial crack length may be represented as a distribution due to the small size and the uncertainty associated with it (Sankararaman, Ling, Shantz, & Mahadevan, 2009).

Model uncertainty comes down to the selection of the CPD model. As stated by Georgiou however, one model cannot and should not be discounted in favor of anther in spite of the uncertainty attached (Georgiou, 2006). That said the model uncertainty of each model should be documented for use in the field.

2.6 Summary

The CPD models presented in this chapter have been developed for the same purposes: To estimate the life expectancy of certain materials and structures that are prone to fatigue cracking, to estimate the POD of a crack or flaw under a given NDT methodology, and to identify weaknesses and uncertainties associated with each CPD models. Isolation of these weaknesses and uncertainties is essential in producing more fact-based modeling methodology that are used in the PHM field. In the process, proper accounting of CSF variability through representation as probabilistic distributions robustness to the models.

Chapter 3: Development of an Integrated Probabilistic Crack Propagation and Detection Modeling Methodology Based on Gaussian Process Regression

3.1 Overview

The CPD models and their association with various CSFs forms the basis for this research. CP models are already directly related to CSFs, and by extension crack detection models are also related to CSFs. Thus a joint-modeling methodology can be designed by treating the separate CPD models as one model. This is done thorough a likelihood function that is made up of a given CPD model set and by way of Bayesian parameter estimation of the parameters. By way of this Bayesian parameter estimation approach, the CP model parameter vector $\vec{A}$, the crack detection parameter vector $\vec{B}$, and the false crack detection probability parameter P_{FD}, can be estimated for a given CPD model set. The research described in this dissertation tested this Bayesian analysis on several CPD model sets including the four CP assessment models from Chapter 2 Section 2.2.2 and the four crack detection models from Chapter 2 Section 2.3.2 (Smith, Modarres, & Droguett, 2017).

This chapter introduces the integrated probabilistic CPD modeling methodology and discusses its development with regard to the sub-models selection and the proposed Bayesian inference methodology used to extract the parameters. The majority of the chapter is devoted to the experimental methods used to obtain and process the data used for the Bayesian analysis.

3.2 The CPD Likelihood Function

The integrated probabilistic CPD modeling methodology starts with development and definition of the likelihood function. A simple statement of the likelihood function is presented as,

$$l\begin{pmatrix} D=0,1;\vec{x}_{i=1},\dots,\vec{x}_{n_D}, \\ \vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}|\vec{A},\vec{B},P_{FD} \end{pmatrix} = \prod_{i=1}^{n_D}\frac{d}{da}P\left(D=1;\vec{x}_i|\vec{A},\vec{B},P_{FD}\right)\times \prod_{j=1}^{m_{ND}}\left[1-P\left(D=1;\vec{x}_j|\vec{A},\vec{B},P_{FD}\right)\right] \tag{3.1}$$

which stands for the likelihood of a set of n_D detection data points $(\vec{x}_{i=1},\dots,\vec{x}_{n_D};a_{i=1},\dots,a_{n_D})$ and m_{ND} non-detection (missed) data points $(\vec{x}_{j=1},\dots,\vec{x}_{m_{ND}};a_{j=1}=0,\dots,a_{m_{ND}}=0)$, where detection state D is 1 for positive detection and 0 for non-detection. The CDF $P(D=1)$ is the probability of detecting a crack of length a given a CSF vector $\vec{x}$. The full expression for $P(D=1)$ integrates both a CP model and a POD model. This is expressed as,

$$P\left(D=1;\vec{x}|\vec{A},\vec{B},P_{FD}\right) = \int_{a_{lth}}^{\infty} POD\left(D=1|\vec{B},a>a_{lth}\right)f\left(a|\vec{A},\vec{x}\right)da + \int_{0}^{a_{lth}} P_{FD}(D=1|a<a_{lth})f\left(a|\vec{A},\vec{x}\right)da \tag{3.2}$$

Equation (3.2) is inclusive of the detection probability of cracks longer than the threshold a_{lth} as well as false detections of cracks below that threshold. However, the CP model Equation (2.5) is a represented as a lognormal distribution, so the integration from 0 to a_{lth} would have no analytical solution. Therefore Equation (3.2) is restated as Equation (3.3),

$$P\left(D=1;\vec{x}|\vec{A},\vec{B},P_{FD}\right) = (1-P_{FD}) \times \int_{a_{lth}}^{\infty} POD\left(D=1|\vec{B},a>a_{lth}\right) f\left(a|\vec{A},\vec{x}\right) da \quad (3.3)$$

The left hand side of the likelihood Equation (3.1) has to be stated as the product of a PDF function by definition of a standard likelihood composed of detections and non-detections[1], therefore, the derivative of the CDF Equation (3.3) is computed with respect to the crack length a resulting in,

$$\frac{d}{da}P\left(D=1;\vec{x},a\middle|\vec{A},\vec{B},P_{FD}\right) = (1-P_{FD}) \times POD\left(D=1|\vec{B},a>a_{lth}\right) f\left(a|\vec{A},\vec{x}\right) \quad (3.4)$$

Finally, the likelihood function Equation (3.1) can be redefined as the *CPD likelihood function*:

$$l\begin{pmatrix} D=0,1;\vec{x}_{i=1},\dots,\vec{x}_{n_D}, \\ \vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}|\vec{A},\vec{B},P_{FD} \end{pmatrix} = \prod_{i=1}^{n_D}\left[\begin{matrix}(1-P_{FD})\times \\ POD\left(D=1|\vec{B},a_i>a_{lth}\right) f\left(a_i|\vec{A},\vec{x}_i\right)\end{matrix}\right] \times \prod_{j=1}^{m_{ND}}\left\{\begin{matrix}1-(1-P_{FD})\times \\ \int_{a_{lth}}^{\infty}\left[\begin{matrix}POD\left(D=1|\vec{B},a>a_{lth}\right)\times \\ f\left(a|\vec{A},\vec{x}_j\right)da\end{matrix}\right]\end{matrix}\right\} \quad (3.5)$$

Bayesian inference for the posterior CPD model parameters is written according to Bayes' Theorem (Bayes, 1763) as follows,

$$\pi\left(\vec{A},\vec{B},P_{FD}\middle|\begin{matrix}D=0,1;a_{i=1},\dots,a_{n_D}, \\ \vec{x}_{i=1},\dots,\vec{x}_n,\vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}\end{matrix}\right) = \frac{l\begin{pmatrix}D=0,1;a_{i=1},\dots,a_{n_D},\vec{x}_{i=1},\dots,\vec{x}_{n_D}, \\ \vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}|\vec{A},\vec{B},P_{FD}\end{pmatrix}\times\pi\left(\vec{A},\vec{B},P_{FD}\right)}{\iiint l\begin{pmatrix}D=0,1;a_{i=1},\dots,a_{n_D},\vec{x}_{i=1},\dots,\vec{x}_{n_D}, \\ \vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}|\vec{A},\vec{B},P_{FD}\end{pmatrix}\times\pi\left(\vec{A},\vec{B},P_{FD}\right)d\vec{A}d\vec{B}dP_{FD}} \quad (3.6)$$

[1] A full likelihood function can be defined as $l=\prod_{i=1}^{n_D} p\left(D=1;\vec{x}_i|\vec{A},\vec{B},P_{FD}\right)\prod_{j=1}^{m_{ND}}\left[1-P\left(D=1;\vec{x}_j|\vec{A},\vec{B},P_{FD}\right)\right]$ where $P\left(D=1;\vec{x}|\vec{A},\vec{B},P_{FD}\right)$ is the CDF and $p\left(D=1;\vec{x}|\vec{A},\vec{B},P_{FD}\right)$ is the PDF

$$\pi\left(\vec{A},\vec{B},P_{FD}\middle|\begin{matrix}D=0,1;a_{i=1},\dots,a_{n_D},\\ \vec{x}_{i=1},\dots,\vec{x}_n,\vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}\end{matrix}\right)\propto$$
$$l\left(\begin{matrix}D=0,1;a_{i=1},\dots,a_{n_D},\vec{x}_{i=1},\dots,\vec{x}_{n_D},\\ \vec{x}_{j=1},\dots,\vec{x}_{m_{ND}}|\vec{A},\vec{B},P_{FD}\end{matrix}\right)\times\pi\left(\vec{A},\vec{B},P_{FD}\right) \quad (3.7)$$

where $\pi(\vec{A},\vec{B},P_{FD}|D=0,1;a_{i=1},\dots,a_n,\vec{x}_{i=1},\dots,\vec{x}_n,\vec{x}_{j=1},\dots,\vec{x}_{m_{ND}})$ is the posterior PDF for the model parameters $\vec{A},\vec{B}$, and P_{FD}, and $\pi(\vec{A},\vec{B},P_{FD})$ is the joint prior PDF for the model parameters.

3.3 Model Development

The Bayesian inference on the CPD Likelihood was tested on combinations of the propagation models listed in Section 2.2.2 and detection models listed in Section 2.3.2. Ultimately, this research will only be exploring the effectiveness of the GPR CP model (Smith & Modarres, 2017) in conjunction with a recursive Bayesian representation of CP based on AE readings and PF (Smith, Modarres, & Droguett, 2017). This particular CP model expands upon previous research in the GPR methodology stated by Mohanty et. al. (Mohanty S. , Chattopadhyay, Peralta, Das, & Willhauck, 2007; Mohanty, Chattopadhyay, & Peralta, 2011), the AE methodology proposed by Keshtgar (Keshtgar, 2013), and PF (Doucet, Godsill, & Andrieu, 2000) methodologies proposed by several researchers (Rabiei & Modarres, 2013; Rabiei, Droguett, & Modarres, 2016).

All detection models listed in Chapter 2 Section 2.3.2 are used in conjunction with the GPR CP model based on the overall ambiguity that is inherent in POD research (Mage, 1981; Ott, 1995; Singh, Warsono, & Bartolucci, 1997). For example, Singh et. al. cites that when modeling their data using both the lognormal POD model and the log-logistic POD model, the two curves look extremely similar to each other (Singh, Warsono, & Bartolucci, 1997).

However, a researcher may take a different set of data and conclude that a gamma distribution may be a better representation than a lognormal distribution (Berger, Melice, & Demuth, 1982; Jakeman & Taylor, 1985). Based on this ambiguity, each of the four listed detection models will be used for this research and compared based on the behavior of the data.

As for the probability of false detection P_{FD}, this will be represented as a beta distribution because it is bounded between 0 and 1 at all times which is representative of a probability.

3.4 Experiment Procedure

The following is a detailed description of the equipment, materials, and procedures used to obtain the data needed to execute the Bayesian inference of the CPD Likelihood model. The majority of the data collected came from a series of fatigue tests which consisted of:

- Time-based and end-state crack length measurements (and CP)
- AE readings
- Loading conditions for the tests
- Material properties and dimensions of the specimens

The loading conditions and material properties were used as the defined CSFs for the GPR and PF/GPR based CP models. From these CSFs, as well as the CP data, a kernel function by which the two data are correlated is defined.

3.4.1 Experimental Setup and Fatigue Testing

The data for this research was gathered from twenty-one fatigue life tests on Al 7075-T6 dog-bone specimens (Keshtgar, 2013; Sauerbrunn, 2016). All of the fatigue tests were

conducted on a uniaxial 22 kilo-Newton Material Testing System (MTS) 810 load frame and divided into three specimen groups.

The first group consists of six tests where the specimens fit the geometry presented in Figure 3-1a (Keshtgar, 2013). The second and third groups of specimens consists of two and thirteen specimens respectively as shown in Figure 3-1b (Sauerbrunn, 2016).

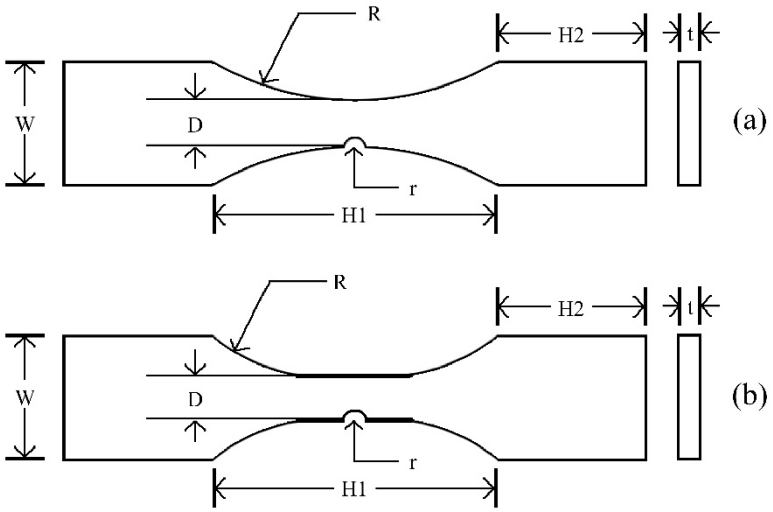

Figure 3-1: Dogbone specimen geometry

The dimensions of the three groups are listed in Table 3-1 (Keshtgar, 2013; Sauerbrunn, 2016).

Table 3-1. The dimensions for the three dog-bone specimen groups in millimeters (Keshtgar, 2013; Sauerbrunn, 2016)

	Group 1 (mm)	**Group 2 (mm)**	**Group 3 (mm)**
W	31.44	45	45
D	10	10	18
$H1$	80	124.88	175.74
$H2$	30	35	45
T	3.175	3.175	3.175
R	80	80	144
r	0.5	0.5	1

These “dog-bone” geometries were selected based on the ASTM-E466-2015 standard (ASTM E466-15, 2015). Each specimen was tested at different values of load frequencies, load ratios, and load ranges listed in Table 3-2 through Table 3-4[2].

Table 3-2: Loading conditions for Group 1 specimens (Keshtgar, 2013)

Specimen Designation	Loading Frequency (Hz)	Load Ratio	Min Force (kN)	Max Force (kN)
DB3	3	0.1	0.8	8
DB4	3	0.1	0.8	8
DB5	2	0.5	6.5	13
DB6	3	0.1	0.8	8
DB7	2	0.5	6.5	13
DB15	2	0.3	3	10

Table 3-3: Loading Conditions for Group 2 specimens (Sauerbrunn, 2016)

Specimen Designation	Loading Frequency (Hz)	Load Ratio	Min Force (kN)	Max Force (kN)
1B3	5	0.1	0.8	8
1A2	5	0.1	0.75	7.5

[2] Specimen Group is denoted by: Group 1 “*”, Group 2 “+”, and Group 3 “#”.

Table 3-4: Loading conditions for Group 3 specimens (Sauerbrunn, 2016)

Specimen Designation	Loading Frequency (Hz)	Load Ratio	Min Force (kN)	Max Force (kN)
5A2	5	0.1	1.2	12
5A3	5	0.1	1.1	11
5A4	5	0.1	1.1	11
5A6	5	0.1	1.05	10.5
5A8	5	0.1	1.05	10.5
5A9	5	0.1	1.05	10.5
5A10	5	0.1	1.1	11
5A20	5	0.1	1.5	15
5A21	5	0.1	1	10
5A22	5	0.1	1.3	13
5A23	5	0.1	0.9	9
5A24	5	0.1	1.2	12
5A25	5	0.1	0.9	9
5A26	5	0.1	1	10

Each fatigue test ran under these conditions until a crack of large enough size was formed. This typically went well past the first detection and sometimes all the way up until the specimen was destroyed or failed.

Through the course of each test, a series of time-lapse photographs are taken at the notch of the sample in order to monitor the cracks as they grow. It is by these photographs that the CP data is obtained. This is accomplished through a complex optical measurement system that allowed a magnification of $100\times$ to effectively spot and measure small cracks measuring as small as 5 μm (Keshtgar, 2013). The system was first designed by Keshtgar (Keshtgar, 2013) and then used again in an updated setting by Sauerbrunn (Sauerbrunn, 2016).

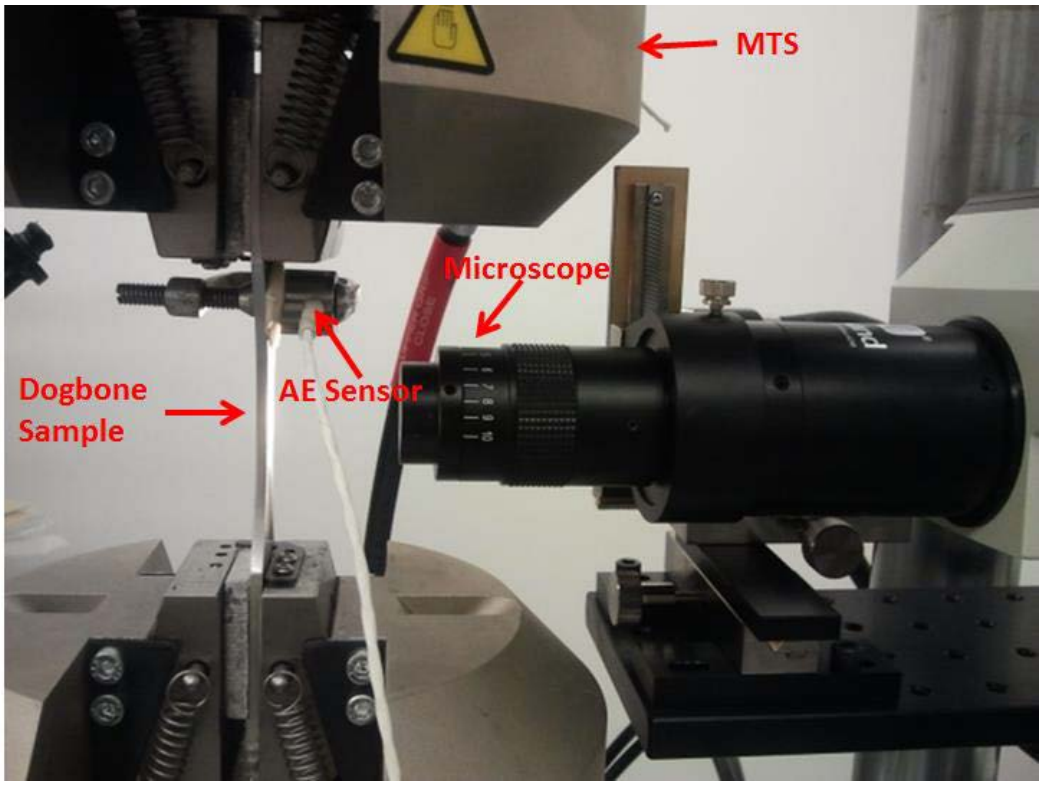

Figure 3-2: The optical microscopy and measurement system constructed to capture fatigue cracks and measurements. (Photo courtesy of Keshtgar (Keshtgar, 2013))

The system, pictured in Figure 3-2, is made up of the following components:

- A high magnification microscope
- A video camera
- A dual arm fiber optic illuminator
- A high resolution monitor
- Image processing software, and
- A micrometer scale

Hundreds of high frequency photograph data was taken in this fashion and marked at its given test time such that they can be correlated to the number of fatigue cycles into the test. Photographic or visual detection was noted only when there is physical evidence of a part

of the final crack path. That is, if a previous picture can be compared to a picture of the final crack and show that it has part of that final crack, then it is seen as a positive detection. The detected cracks are measured along the path using Java-based imaging software called ImageJ (NIH, 2015).

3.4.2 Acoustic Emission Specifications

Likewise, AE detection acts as a signal response correlating factor to the length of the detected crack hence the necessity to apply an AE sensor to each specimen (pictured in Figure 3-2). An advanced AE system (Physical Acoustic Corporation., 2007) was used to obtain the two AE signals of interest cumulative count $C(N)$ and cumulative amplitude $A(N)$, both of which are defined in Chapter 2 Section 2.2.2.2. Keshtgar and Sauerbrunn both used different AE instrumentation settings to obtain their AE signal data which is summarized in Table 3-5 (Keshtgar, 2013; Sauerbrunn, 2016).

Table 3-5: Acoustic emission software settings from Keshtgar and Sauerbrunn's tests (Keshtgar, 2013; Sauerbrunn, 2016)

Parameter	**Values**	
	Group 1 Specimens	**Group 3 Specimens**
Preamplifier	40 dB	40 dB
Detection Threshold A_0	35 dB	45 dB
Sampling Rate	5 MSPS[3]	1 MSPS
Peak Definition Time (PDT)	N/A	300 µs
Hit Definition Time (HDT)	N/A	600 µs
Hit Lockout Time (HLT)	N/A	1,000 µs
Pre-trigger length	100 µs	256 µs
Hit length	614 µs	2,048 µs
High-pass analogue filter	200 KHz	1 kHz
Low-pass analogue filter	3 MHz	3 MHz

[3] MSPS means "million samples per second"

Keshtgar's settings (Keshtgar, 2013) were based on standards set by the Physical Acoustic Corporation, while Sauerbrunn's settings (Sauerbrunn, 2016) were based on standards set by pencil lead break tests (ASTM, 2015).

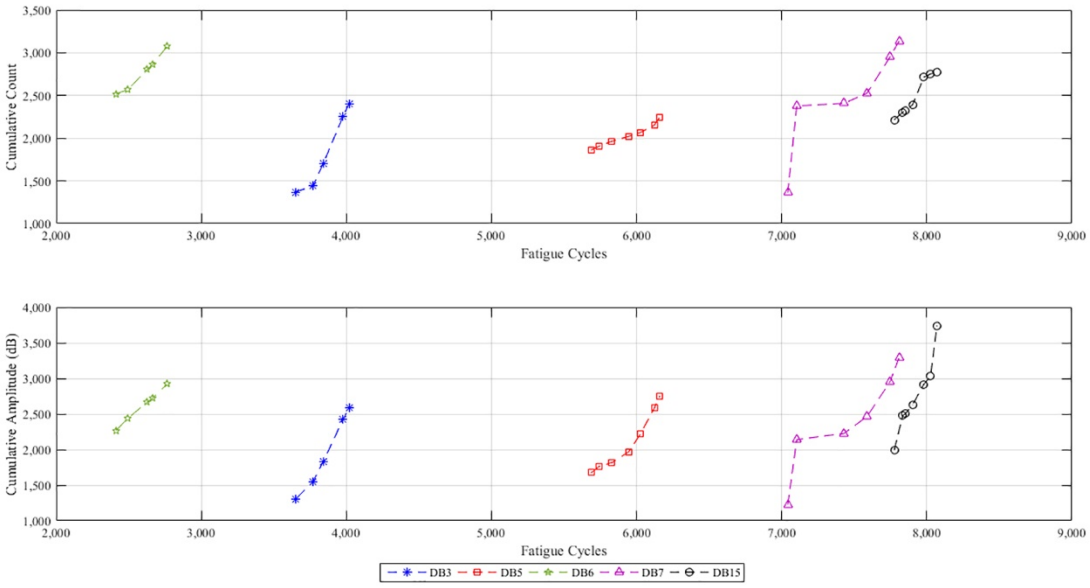

Figure 3-3: The normalized cumulative count and cumulative amplitude data from fatigue tests of specimen Set 1

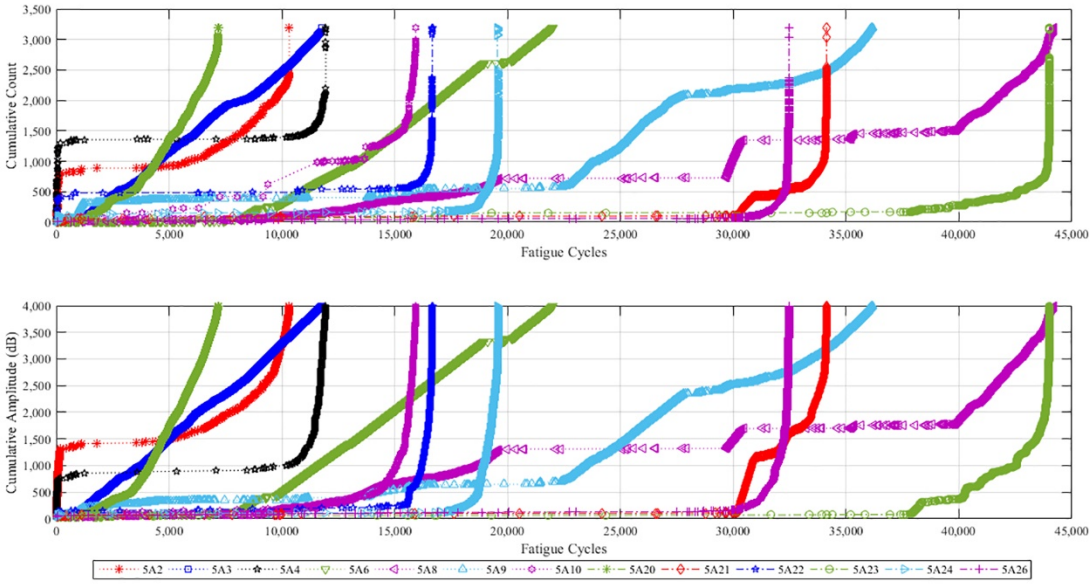

Figure 3-4: The normalized cumulative count and cumulative amplitude data from fatigue tests of specimen Set 3

As seen in Figure 3-3, the AE data was only available from five of the six tests from the first group of data with between five and seven AE index data points per test. However, as seen in Figure 3-4 the third group of specimen tests had up to twenty-thousand AE index data points per test.

3.4.3 Acquisition of Crack Length Data and Uncertainties

The following section goes into the detail of the processes involved in measuring and preconditioning the fatigue crack data and the AE data. This preconditioning is done by way of a simple measurement error correction based on three length variables used in this research (Smith & Modarres, 2017): (1) the experimental-based crack length (2) the "true" crack length, and (3) the model-based crack length. These crack length representations and their correlation to the AE data are further discussed in the following subsections.

3.4.3.1 Experimental-Based Crack Length

Experimental-based crack lengths a_e are the sized and detected lengths online as the experiment progresses. The lengths are denoted as $a_{e,m}$ in this research which stands for "experimental measurement."

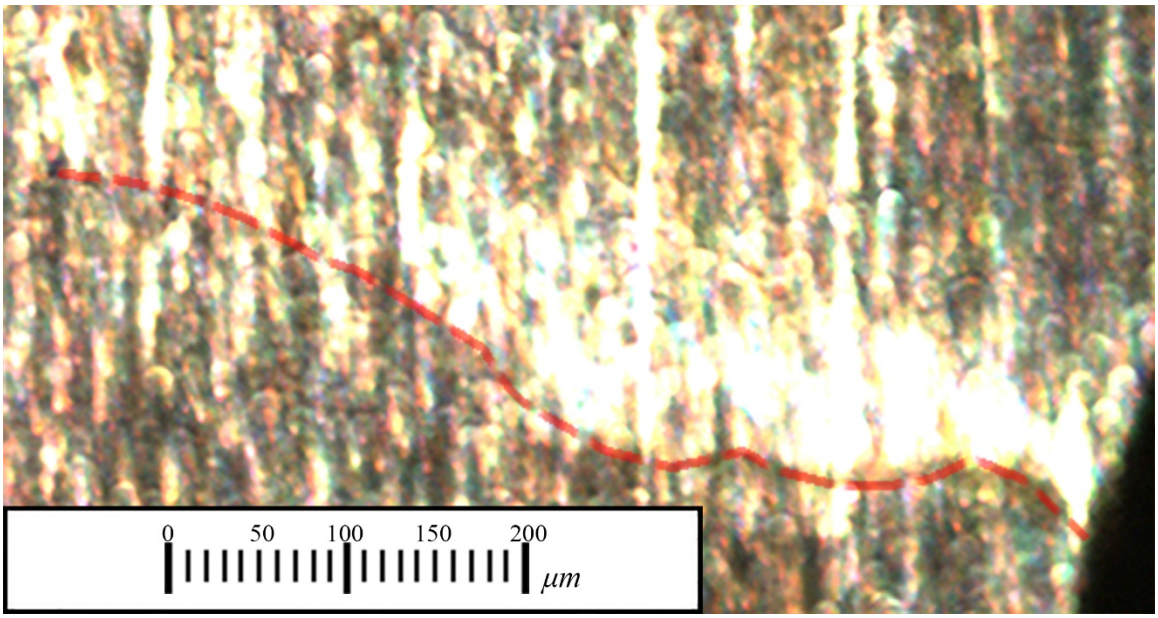

Figure 3-5: An example of an experimental-based crack measurement taken from the ImageJ program (NIH, 2015)

As illustrated in Figure 3-5, this crack length is subject to detection probability and measurement error because the best online images captured are not very clear because:

1. Due to inherent vibration of the specimen, in-test images at best can be taken at a magnification of $100\times$
2. Images are subject to motion blur
3. Crack initiation detection probability and errors in measurement tools

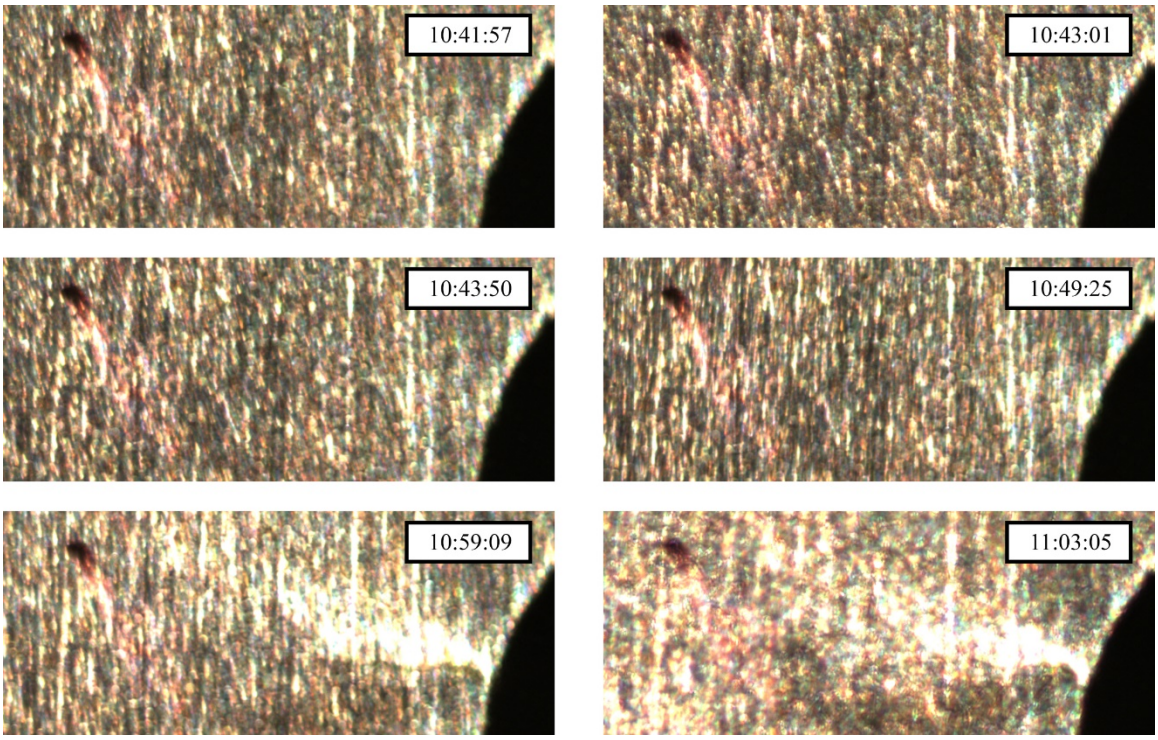

Figure 3-6: A time-lapse display representing fatigue crack propagation in Specimen DB7

In some specimens where the crack path was not apparent for each individual photo a time-lapse animation needed to be developed by overlaying several images in order to detect and isolate the crack path. An example of this is presented in Figure 3-6.

3.4.3.2 True Crack Length

The "true" crack lengths a are sized at the highest reasonable magnification after the completion of the fatigue tests. This is done to eliminate the motion blur and the specimens are polished and etched around the propagation region. The magnification scale of $200\times$ was selected as a result of a (stochastic) percent error analysis listed in Table 3-6. While there is still a small measurement error remaining, this can be tolerated and taken as the true measurement.

Table 3-6: Percent error of sample crack length measurement between different magnification scales

$\mathbf{Scale_i}$	$\mathbf{Scale_j}$	**Percent Error between $Scale_i$ and $Scale_j$**
100 ×	200 ×	46.79%
200 ×	400 ×	2.28%
400 ×	1,000 ×	2.07%

The post-test specimens were cut and polished at the CP area to prepare for the magnification procedure. When under the microscope several high-definition photographs were taken at different areas of the specimen around the fatigue crack. Afterwards the photos of each specimen were spliced together in Adobe Photoshop to create a single 200 × magnification image of the fatigue CP area.

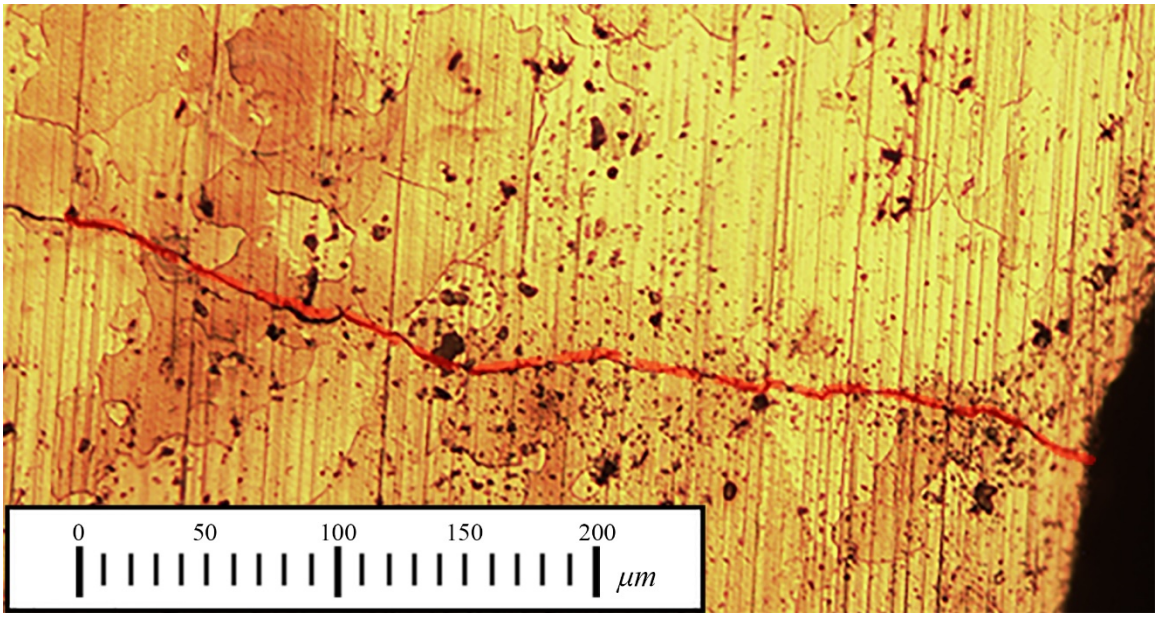

Figure 3-7: An example of a true crack measurement taken from the ImageJ program (NIH, 2015)

It is from photos such as in Figure 3-7 where the true crack lengths are acquired. Because this photo is taken after the test however, an overlay of some of the experiment-based crack measurements have to be made in order to obtain an approximation of the true crack length at a given in-test fatigue cycle. Forty such measurements were obtained.

3.4.3.3 Model-Based Crack Length

Model-based crack lengths a_m are predictively sized using the crack length data fit into a given propagation model (see Section 2.1). For example, crack lengths obtained from the GPR model are denoted as $a_{m,GPR}$ (See Section 2.1.3 and the CSFs defined in Section 3.1).

3.4.3.4 Acoustic Emission Data Conditioning

The camera used in testing to take time-lapse photos of the CP in progress is out of phase with the AE recording equipment, so the AE signals need to be approximated to the fatigue cycles of detection. The AE software settings for the two groups of specimens were different (see Table 3-5), so the procedure for synchronization was different as well. For Keshtgar's tests (the first group) (Keshtgar, 2013) the data is limited to five to seven data points of AE indices. As a result, both indices are modeled using an empirical form of the power law such that $C(N) = \alpha_C N^{\beta_C}$ and $A(N) = \alpha_A N^{\beta_A}$ where $[\alpha_C, \beta_C]$ and $[\alpha_A, \beta_A]$ are the model parameter pairs for the cumulative count function and the amplitude function respectively. However, because of the large number of data points from Sauerbrunn's tests (the third group) (Sauerbrunn, 2016) linear interpolation was used to synchronize the AE data to the respective crack lengths and fatigue cycles.

3.4.3.5 Measurement Error and Analysis

As only forty true crack length measurements were taken, the rest of the measurements are adjusted by way of measurement error correction. The measurement errors of true crack length with respect to the *experimental-based* and *model-based* crack lengths, a_e and a_m, are defined as:

$$E_{a,e} = \frac{a_e}{a} \tag{3.8}$$

$$E_{a,m} = \frac{a_m}{a} \tag{3.9}$$

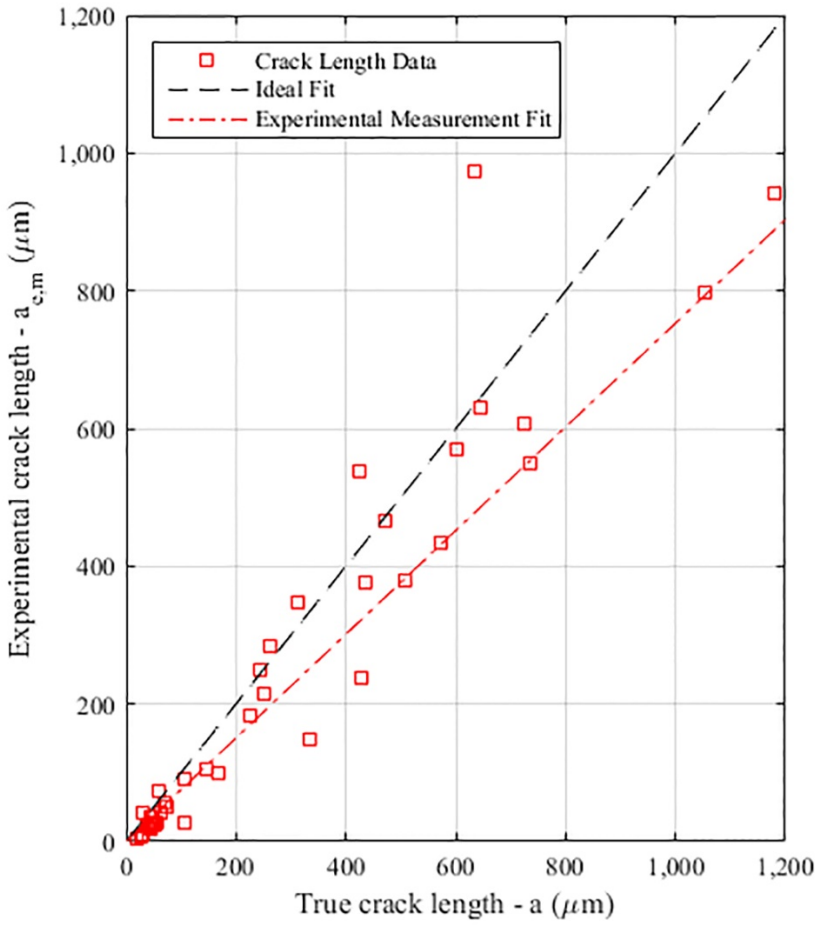

Figure 3-8: Crack length data that compares the true length to the experimental length

The forty true measurements compared to their experiment-based measurements are depicted in Figure 3-8. A mean experiment-based measurement error of 0.75 is obtained (Smith & Modarres, 2017) and is used as the correction ratio for the remaining experiment-based measurements. Alternately, the relation between a_e and a may be defined as the probabilistic model,

$$a = \frac{a_e}{LOGN(-0.36,0.38)} \tag{3.10}$$

The mean measurement error is used for preconditioning the data because the probabilistic measurement error would have been more expensive computationally as more crack length data per detection cycle would be generated. However, probabilistic measurement error with respect to both the model and experimental lengths is applied in the Bayesian estimation and validation phases as will be seen in Chapters 4 and 5.

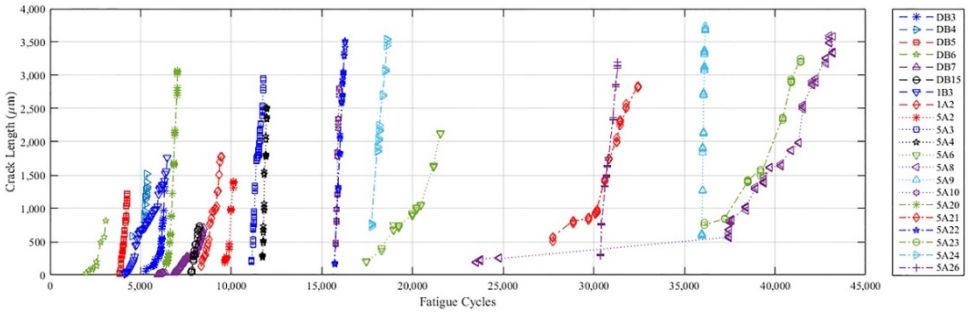

Figure 3-9: The crack propagation against fatigue cycles for all Al 7075-T6 specimens

The CP data in Figure 3-9 is the result of this mean measurement error correction procedure and is used for the remainder of the methodology demonstration.

3.4.4 Crack Shaping Factors

The CSFs for this research are selected from among the list presented in Chapter 2 Section 2.4. The following CSFs are extracted from the testing specimens for this purpose:

1) Fatigue cycles
2) Minimum force
3) Maximum force
4) Load ratio
5) Test frequency

6) Mean grain diameter

7) Mean inclusion diameter

The first CSF, fatigue cycle, is extracted from the time-stamp of the photographs and the loading frequency. The minimum and maximum force as well as the load ratio and test frequencies for each fatigue test are given from Table 3-2 through Table 3-4. The last two CSFs, mean grain diameter and mean inclusion diameter, are represented as distributions because of the variability that is present in these material properties.

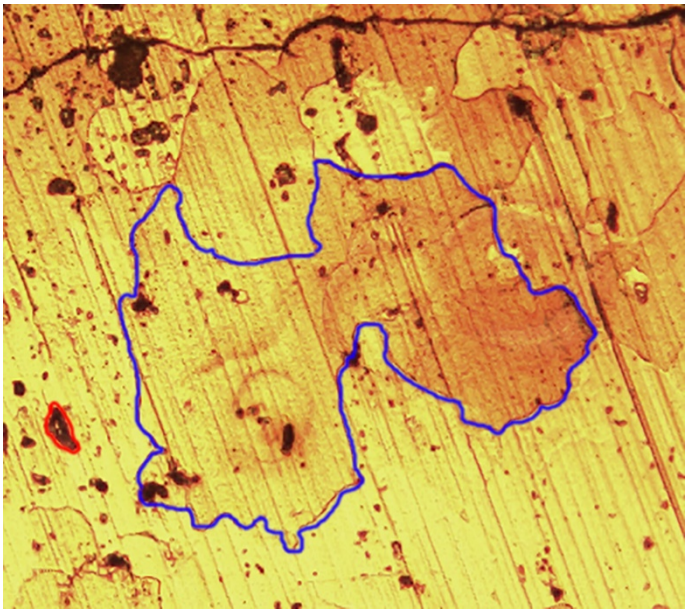

Figure 3-10: Example of a material grain (outlined in blue) and a material inclusion (outlined in red)

As exemplified in Figure 3-10, the documentation of mean grain diameter and mean inclusion diameter is done in the post-test phase as with the true crack length measurements. The procedure requires roughly twenty mean grain diameters and mean inclusion diameters to be measured and recorded for the data. A mean diameter for each

grain and inclusion is obtained by four multiaxial measurements followed by an averaging of those measurements (see Figure 3-11).

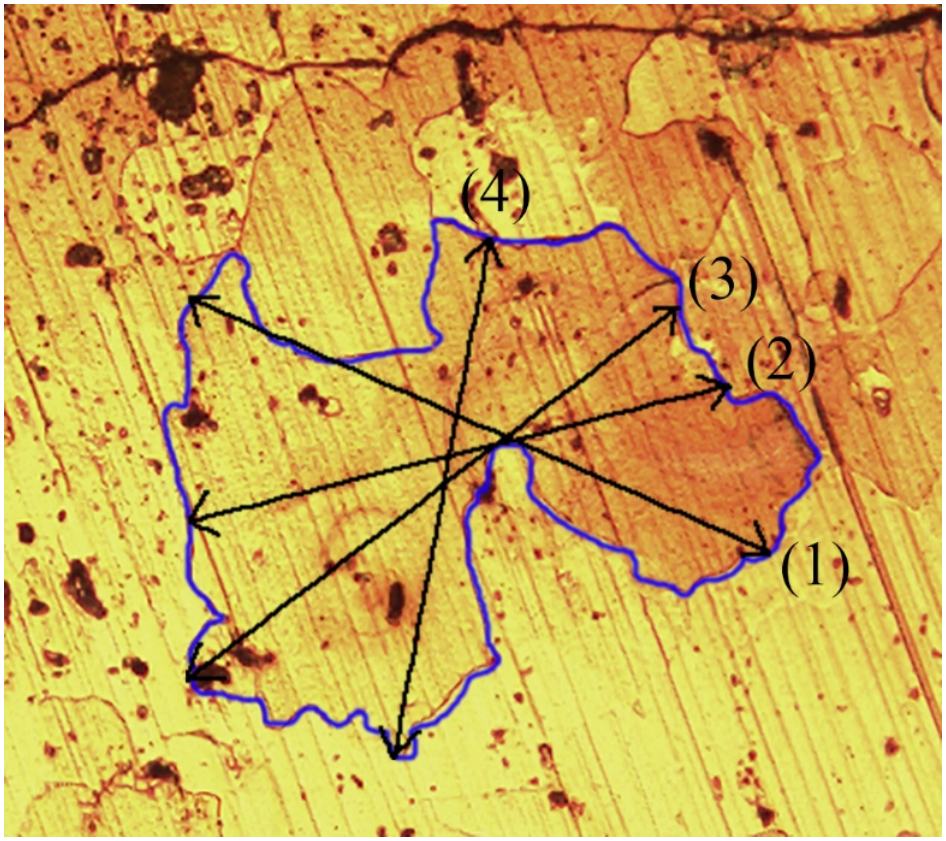

Figure 3-11: Mean diameter measuring procedure for grains and inclusions

The Weibull (WBL) distribution is assumed as the best fit to represent the random variability of these CSFs whose parameters are listed in Table 3-7 through Table 3-9 for each specimen. Both lognormal and Weibull distributions were considered for modeling these CSFs, but the Weibull distribution was selected based on a goodness-of-fit analysis.

Table 3-7: Material Properties for Group 1 specimens (Keshtgar, 2013)

Sample Designation	Mean Grain Diameter WBL Distribution		Mean Inclusion Diameter WBL Distribution	
	α_g	β_g	α_i	β_i
DB3	100	1.3	10	3.5
DB4	103	1.7	14	1.9
DB5	98	1.5	11	1.2
DB6	71	1.3	8	2.1
DB7	70	1.6	10	1.7
DB15	95	1.3	6	1.8

Table 3-8: Material Properties for Group 2 specimens (Sauerbrunn, 2016)

Sample Designation	Mean Grain Diameter WBL Distribution		Mean Grain Diameter WBL Distribution	
	α_g	α_g	α_i	β_i
1B3	111	1.4	12	2.7
1A2	101	1.2	11	1.6

Table 3-9: Material Properties for Group 3 specimens (Sauerbrunn, 2016)

Sample Designation	Mean Grain Diameter WBL Distribution		Mean Grain Diameter WBL Distribution	
	α_g	α_g	α_i	β_i
5A2	133	1.2	8	2.0
5A3	95	0.95	11	2.2
5A4	147	1.3	10	2.1
5A6	185	1.2	18	1.5
5A8	158	1.3	18	1.5
5A9	149	1.1	12	2.3
5A10	175	1.4	12	1.8
5A20	139	1.5	10	2.4
5A21	122	1.4	15	1.8
5A22	130	2.2	11	3.0
5A23	140	1.9	16	3.3
5A24	116	1.4	10	3.8
5A25	120	2.4	11	4.0
5A26	146	2.5	15	2.5

Thus, a total of nine CSFs ($Q = 9$) are used in this research to correlate to crack growth propagation: (CSF #1) the variable fatigue cycles N, (CSFs #2-5) the four loading

conditions, and (CSFs #5-9) the two material properties each represented as Weibull distributions. It must be noted that the loading conditions are also subject to variation within the test and could, therefore, be represented in form of a distribution function as well. However, for this research the variation is assumed to be minimal so they are treated as constants. It would be suggested that for future work, a similar study should be conducted where the loading conditions are also treated as random variables represented by probabilistic distributions.

3.4.5 GPR and PF/GPR Crack Propagation Gaussian Kernel Function

As stated in Chapter 2, a multivariate GPR function is characterized by its input and output data and the kernel function designed to correlate the two data. In this research there are two input/output relations that need to be characterized by a kernel. The first is the *CSF/CP* relation which is used for the *GPR* and *PF/GPR propagation models* (Section 2.2.2.3 and 2.2.2.4 respectively).

Development of these kernel functions has to be based on the behavior of the data involved. One such kernel function, a "*standard kernel function*" is very well-known in GPR literature (Rasmussen, 1996; Chen, Morris, & Martin, 2007).

$$k\left(\vec{x}_i, \vec{x}_j, \vec{A}\right) = g_1 + \sum_{q=1}^{Q} g_{q+1}\vec{x}_{i,q}\vec{x}_{j,q} + g_{2+2Q}\exp\left[-\sum_{q=1}^{Q} g_{q+Q+1}\left(\vec{x}_{i,q} - \vec{x}_{j,q}\right)^2\right] + g_{3+2Q}\delta_{i,j} \quad (3.40)$$

Its components include:

- A *constant bias or offset* term g_1
- A *linear* component $\sum_{q=1}^{Q} g_{q+1}\vec{x}_{i,q}\vec{x}_{j,q}$

- A *squared exponential* function that accounts for the correlation between output and input $g_{2+2Q} \exp\left[-\sum_{q=1}^{Q} g_{q+Q+1}\left(\vec{x}_{i,q} - \vec{x}_{j,q}\right)^2\right]$
- A *noise or random effect error* function $g_{3+2Q}\delta_{i,j}$ where $\delta_{i,j}$ is a Dirac delta function that equals 1 when $i = j$ and 0 elsewhere (Chen, Morris, & Martin, 2007)

This is often selected as an initial kernel because of its tendency to model the linearity, the curvatures, and randomness of most sets of data (Chen, Morris, & Martin, 2007). As such, the standard kernel function is the foundation of the two kernels for this research.

The kernel correlating CSFs to CP is developed based on the following conditions:

1. The CP function has to be a positive monotonically increasing function at all times. That is $a(\vec{x})$ must be a strictly increasing function.
2. The upper and lower confidence bounds of the crack length outside of the region where data exists (before first crack data point detected and after last data point) must also be strictly positive and increasing.
3. The mean CP model must be a good fit to its training and validation data.

Setting the output data as $\ln(a)$ rather than a maintains half of Condition 1. The other half of Condition 1 as well as Condition 2 are satisfied under the standard kernel function. However, Condition 3 is satisfied with fitness, and this is tested by the normalized mean square error (NMSE),

$$NMSE = \frac{1}{M\sigma_t^2}\sum_{i=1}^{M}\left(a_{GPR,i} - a_i\right)^2 \tag{3.41}$$

between the true crack lengths a_i and the GPR model crack length estimates $a_{m,GPR}$. The term σ_t^2 is the variance of the true crack lengths. New kernel functions are typically designed by adding or subtracting certain kernel function components to or from the overall function. The neural network function $g_3 \sin^{-1} \frac{g_2 \sum_{q=1}^{Q} x_{i,q} x_{j,q}}{\sqrt{\left(1+g_2 \sum_{q=1}^{Q} x_{i,q} x_{j,q}\right)^2}}$, for example (Rasmussen, 1996) was considered as a potential addition to the standard kernel function. This is the list of kernel functions selected for review:

- Kernel 1: the standard kernel function
- Kernel 2: a kernel consisting of a constant, neural network (Rasmussen, 1996), and noise component
- Kernel 3: the standard kernel function with the neural network component (Rasmussen, Nickisch, & Williams, 2015)

These kernel functions were tested on data from the first set of specimens (Geometry 1) with the results presented in Table 3-10.

Table 3-10: Kernel analysis showing increasing function possibilities and normalized mean square error

	Increasing Function	**NMSE**
Kernel 1	Yes	1.78×10^{-2}
Kernel 2	No	5.17×10^{-4}
Kernel 3	Yes	5.05×10^{-4}

The addition of the neural network component produces a kernel with a smaller NMSE than the other kernel options. Furthermore this kernel maintains adherence to the three stated conditions. Therefore the following kernel function,

$$k(\vec{x}_i, \vec{x}_j, \vec{A}) = g_1 + \sum_{q=1}^{9} g_{q+1}\vec{x}_{i,q}\vec{x}_{j,q} + g_{20}\exp\left[-\sum_{q=1}^{9} g_{q+10}(\vec{x}_{i,q} - \vec{x}_{j,q})^2\right] + g_{22}\sin^{-1}\frac{g_{21}\sum_{q=1}^{9} x_{i,q}x_{j,q}}{\sqrt{(1 + g_{21}\sum_{q=1}^{9} x_{i,q}x_{j,q})^2}} + g_{23}\delta_{i,j} \quad (3.42)$$

is used for the GPR (and PF/GPR) CP analysis of this research. Therefore, the GPR (and PF/GPR) propagation model parameter set $\vec{A}$ is a twenty-three-parameter vector $[g_1, g_2, \ldots, g_{23}]$.

3.4.6 Estimation of Bayesian Prior Model and Parameters

The test data that has been defined so far is used to obtain the prior distribution for $\vec{\Theta}$ which is the collection of all the model parameter sets of interest: The CP parameter set $\vec{A}$, the POD parameter set $\vec{B}$, and the false detection probability P_{FD}. At the beginning of the research only specimen Sets 1 and 2 were available, so the prior parameters described in this section only applies to them. The posterior parameter distributions that result from following the Bayesian analysis methodology cited in section 3.3 of this chapter are used as prior parameter distributions for analysis of specimen Set 3. This will be covered more in Chapter 6 of this dissertation.

The first parameter set $\vec{A}$ is obtained for each specimen based on a GPR propagation model (Section 2.2.2.3) using the CP data (seen in Figure 3-9), the CSF data (seen in Table 3-2, Table 3-3, Table 3-7, and Table 3-8), and the GPR kernel function (seen in Equation (3.39)). However, because of the probabilistic measurement error relation stated in Equation (3.10), the CP data as well as the prior CP parameters $\vec{A}$ are also probabilistic. Thus for each variant of the CP data, the prior CP parameters $\vec{A}$ are obtained by way of a

MATLAB (MathWorks, 2016) software package called GPML (Rasmussen, Nickisch, & Williams, 2015).

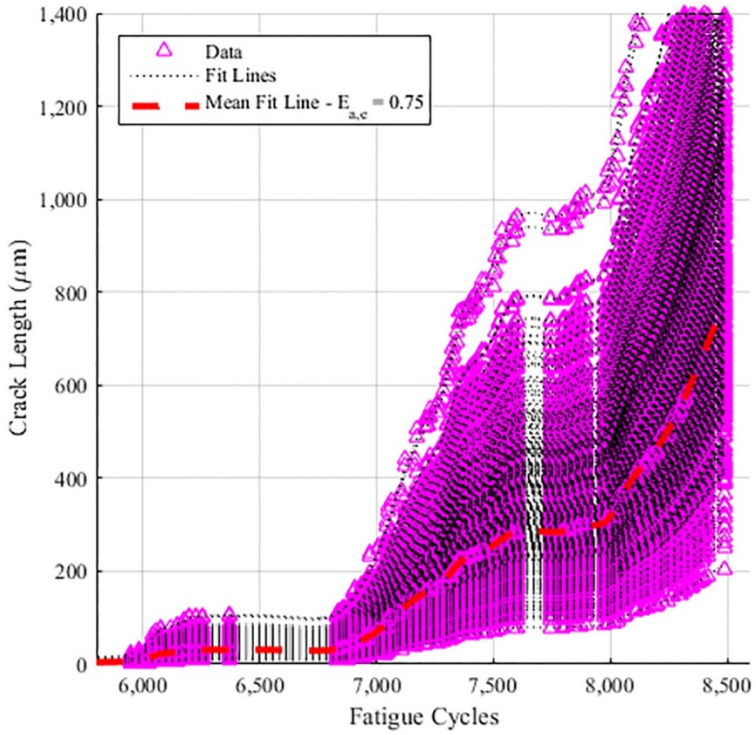

Figure 3-12: CP data of dog-bone specimen DB7 based on the distribution of the experimental measurement error

This then produces a specimen-specific distribution for the CP prior parameter set $\vec{A}$. As stated in Section 3.4.3.5 of this chapter, the mean fit line (Figure 3-12) is used for the complete Bayesian analysis of the CPD likelihood function Equation (3.5). The parameters for the mean for each specimen is presented in Table 3-11.

Table 3-11: The prior mean parameter sets $\vec{A}$ for all specimens of Sets 1 and 2

Hyper parameter	DB3	DB4	DB5	DB6	DB7	DB15	1B3	1A2
g_1	0.0485	1.05	0.999	0.0179	0.161	1.00	0.882	1.00
g_2	2.52×10^6	1.94×10^{12}	115	382	980	164	641	502
g_3	1.43×10^4	5.71×10^9	418	5.27×10^4	1.11×10^8	203	1.61×10^3	1.96×10^9
g_4	2.13×10^6	1.4×10^{12}	1.09×10^3	5.30×10^5	1.22×10^7	590	1.45×10^5	2.55×10^4
g_5	1.69	0.999	1.00	6.38	2.75	1.00	1.00	1.00
g_6	60.1	0.644	1.00	368	13.2	1.00	4.74	0.963
g_7	4.10×10^3	835	5.09	4.83×10^3	2.30×10^3	2.62	87.9	1.43×10^3
g_8	25.3	0.861	1.00	68.4	10.6	1.00	1.27	0.997
g_9	196	4.85	1.08	475	64.4	1.01	36.2	1.03
g_{10}	72.9	0.831	1.00	211	11.1	1.00	1.98	0.995
g_{11}	1.00	2.54×10^4	33.8	1.00	0.0731	57.3	1.00	1.00
$g_{12} \cdots g_{19}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
g_{20}	0.0239	7.84×10^{-7}	0.449	0.0740	0.00463	0.421	1.19×10^{-7}	3.36×10^{-6}
g_{21}	39.9	13.0	0.164	0.113	63.4	0.995	22.4	3.23
g_{22}	5.90	1.97	0.586	1.61	3.70	0.0152	1.81	1.81
g_{23}	0.0239	0.0139	0.0298	0.0740	0.0312	0.0713	1.19×10^{-7}	3.36×10^{-6}

Note however that the applied measurement error is still subject to some random error which may be detected as noise. This noise, when applied to the CP data curves, creates an additional uncertainty in this example of the proposed methodology. In future use of this methodology, this random noise should be addressed and corrected prior to the step that establishes the CP prior parameter sets $\vec{A}$.

The next parameter set $\vec{B}$ is obtained based on the signal-response POD outlined in Equation (2.29) and (2.30). The true crack length data and their AE intensities (the signal response index) when fit to this signal response POD and the POD models defined in Chapter 2 Sections 2.3.2.1-2.3.2.4 (and the POD model decision block in Figure 1-1), produce the prior POD parameter sets $\vec{B}$.

A crack detection threshold of,

$$5\ \mu m < a < 48\ \mu m \tag{3.43}$$

was chosen for this extraction procedure and as the definition of the small CP region. The basis for this selection comes from both literature and test results. The data obtained for this research places crack detection to as small as about $5\ \mu m$. While literature states that the upper bound of detection typically falls between 48 and $50\ \mu m$ for most aluminums (Keshtgar, 2013). Thus, this crack detection threshold is used as the definition for the small CP region for this research. Only ten measurements within the threshold Equation (3.43) were identified. Combining Equation (2.11) and Equation (2.30) gives the following signal response POD equation,

$$POD(a) = 1 - F\left[\frac{\ln \hat{I}_{th} + \frac{1}{\beta}(\ln \alpha - \ln a)}{\sigma_Z}\right] \tag{3.44}$$

where AE intensity is the index in question and $\hat{I}_{th}$ is the AE intensity threshold where the cumulative amplitude $A(N)$ is set to the amplitude threshold A_0. Calculation of the threshold $\hat{I}_{th}$ is based on the power relationship between AE intensity and the true crack length as shown in Figure 3-13 and Table 3-12.

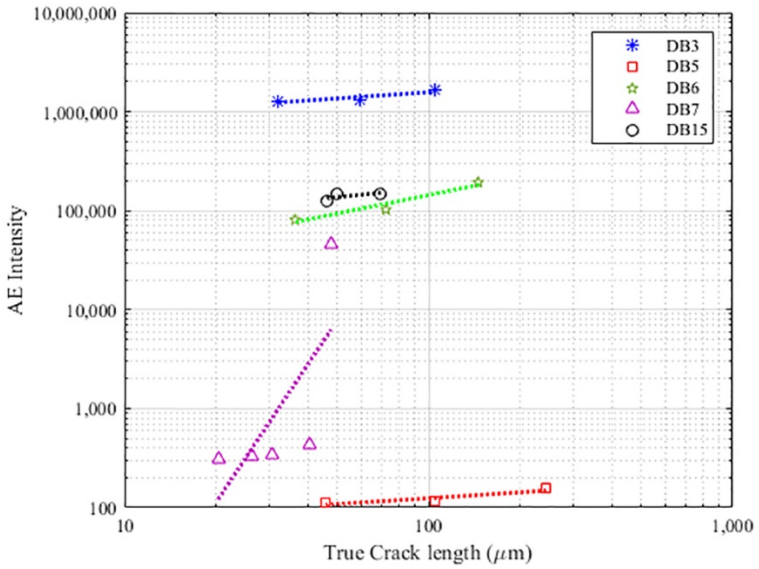

Figure 3-13: True crack length versus acoustic emission intensity

Table 3-12: Signal response parameters for five specimens from Set 1

	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{I_0}$
DB3	3.12×10^{-28}	0.0752	251
DB5	$3.07E \times 10^{-9}$	0.255	1345
DB6	4.88×10^{-7}	0.111	499
DB7	7.20	-0.110	532
DB15	1.32×10^{-15}	0.0942	1280

While each specimen has its own intensity threshold value, the total set of signal response PODs are used to fit to each POD model option defined in Chapter 2 Sections 2.3.2.1-2.3.2.4. The fits for these POD models and thus the prior sets $\vec{B}$ are presented in Figure 3-14 and Table 3-13.

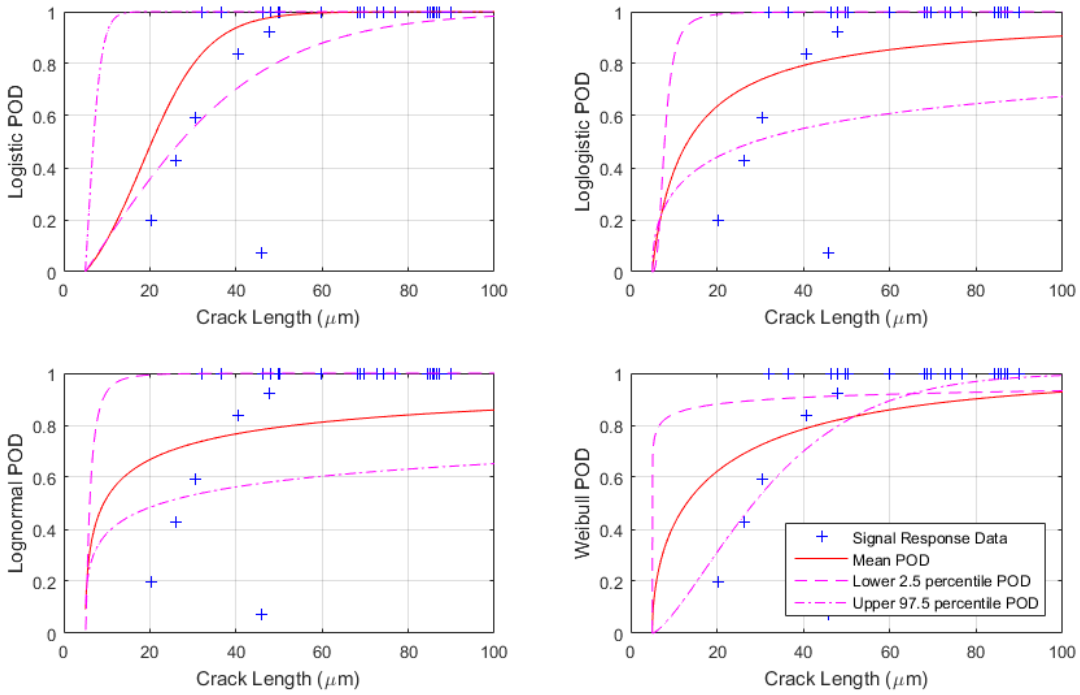

Figure 3-14: The prior logistic, log-logistic, lognormal, and Weibull POD curves

Table 3-13: The prior POD parameter distributions for each POD model type

	Hyper parameter	Mean	SD	2.50%	Median	97.50%
Lognormal	ζ_0	1.46	0.863	0.0395	1.47	2.88
	ζ_1^2	8.29	5.30	1.15	7.35	18.4
Log-logistic	β_0	2.11	0.640	1.05	2.11	3.16
	β_1	0.0735	0.460	−1.01	0.177	0.648
Logistic	η_0	0.133	0.169	0.0502	0.0612	0.674
	η_1	13.22	7.65	0.666	13.2	0.674
Weibull	α_0	15.76	9.08	0.0395	15.8	30.6
	α_1	0.539	0.335	0.127	0.451	1.38

The coefficients of determination R^2 for the four POD prior models are,

- $R^2_{logistic} = 0.39$
- $R^2_{loglogistic} = 0.25$
- $R^2_{lognormal} = 0.088$

- $R^2_{Weibull} = 0.29$

The prior parameter for the false crack detection probability P_{FD}, has been defined as a beta distribution $BETA(0.25,4.17)$ by way of the equations set in Chapter 2 Section 2.3.3. The three sets of priors form the basis for the complete prior distribution of the target parameter set $\vec{\Theta}$ for specimen Sets 1 and 2.

3.5 Summary

This chapter defined the Bayesian procedure and development of the new CPD likelihood model that is essential in this methodology. The primary focus of this test of this methodology will be the Bayesian analysis of the joint-CPD models of the GPR and PF/GPR CP models with the four POD models from Chapter 2 Section 2.3.2. Using an intricate experimental procedure, the needed data is acquired which includes detected crack length data, AE signals, and the CSFs selected for this research. The data in its preconditioned forms of measurement error corrected CP data and prior parameter distributions for the joint GPR propagation and detection models is to be used in the Bayesian analysis stage of the methodology.

Chapter 4: Effect of Crack Shaping Factors on Model Uncertainty

4.1 Overview

A major part of this research relies upon the CSFs that contribute to the CPD model; as such there is a need to understand the nature of the effect that each CSF has on the CPD model parameters. This understanding comes in the form of a CSF-to-CPD correlation technique that also makes use of the input-output relation found in GPR analysis. The scope of the GPR correlation is extended beyond a single correlation and tested through a validation procedure in which the CPD can be predicted with additional model error correction.

This chapter will cover the procedures that are used to correlate the CSFs to the CPD model parameters as well as estimate the model error and perform validation. It concludes with a more in-depth explanation will be provided on the recursive Bayes estimation procedure that is specific to this research where the known variables are the AE signals cumulative count and cumulative amplitude, and the unknown variable is the true crack length.

4.2 Crack Shaping Factor to Model Parameter Correlation

As stated in Chapter 3 Section 3.4.5, there are two input/output relations that need to be characterized by a kernel for this research. The second input/output relation is the *CSF/CPD model parameter* relation which is used primarily for validation of the posterior CP model parameters. This relation is based on the idea that CP of a material can be predicted by a limited set of detected cracks and a unique set of CSFs. The CSFs for the training input are chosen from among the specimens specific CSFs listed in the loading condition tables (Table 3-2, Table 3-3, and Table 3-4) and the material property tables

(Table 3-8, Table 3-9, and Table 3-10) in Chapter 3 Section 3.4.1 and 3.4.4. Thus the CSFs are correlated to each output posterior CPD parameter within the output vector of a training set; and the validation set is used to test this correlation.

4.2.1 CSF-to-CPD Correlation Gaussian Kernel Function

The conditions guiding the kernel correlating CSFs to the CPD parameters h are as follows:

1. The training set of CPD parameters must fall in the same range as the validation set of propagation parameters (positive, negative, or positive/negative)
2. The CPD parameter mean must be a good fit to its training and validation data

As with the GPR propagation kernel defined in Chapter 3 Section 3.4.5, Condition 1 is kept by setting the output data for CPD parameters h to $\ln(h)$ or $\ln(-h)$, which keeps the validation estimates in the positive or negative range respectively. Keeping the output data h as is frees the validation estimates to a positive/negative range. Condition 2 for fitness is also satisfied by finding the smallest NMSE between the estimated and actual posterior CPD parameters. The test is done on variants of Kernels 1 and 3 that removed the constant bias term g_1. Each kernel option produces twenty-five NMSE estimates where the means for Kernels 1 and 3 are 0.59 and 0.58 respectively. Both averages are very small, however, looking at the NMSE for each h parameter in $[\vec{A}, \vec{B}]$ shows that the NMSE for Kernel 3 is less than that of Kernel 1 for fifteen of the twenty-five parameters (60% of the CPD parameters). So the CSF/CPD correlation kernel for this research is defined as,

$$k(\vec{x}_i, \vec{x}_j, \vec{V}) = \sum_{q=1}^{8} v_q \vec{x}_{i,q} \vec{x}_{j,q} + v_{17} \exp\left[-\sum_{q=1}^{8} v_{q+9}(\vec{x}_{i,q} - \vec{x}_{j,q})^2\right] + v_{19} \sin^{-1} \frac{v_{18} \sum_{q=1}^{8} x_{i,q} x_{j,q}}{\sqrt{\left(1 + v_{18} \sum_{q=1}^{8} x_{i,q} x_{j,q}\right)^2}} + v_{20} \delta_{i,j} \quad (4.1)$$

where $\vec{V}$ is the correlation parameter vector $[v_1, v_2, \ldots, v_{20}]$.

4.2.2 Validation and Model Error

The validation methodology for this research is based on a methodology proposed by Ontiveros et. al. (Ontiveros, Cartillier, & Modarres, 2010) which makes use of the reciprocal measurement errors associated between the experiential and true lengths Equation (3.8) and the model true lengths Equation (3.9). These measurement errors $E_{a',e}$ and $E_{a',m}$ both fall into a distribution which then translate to a combined effect measurement error $E_{a',t}$,

$$a = E_{a',e} a_e = E_{a',m} a_m \Longrightarrow \frac{E_{a',m}}{E_{a',e}} = \frac{a_e}{a_m} = E_{a',t} \quad (4.2)$$

which would also fall into that same distribution. Originally this distribution was assumed to be lognormal (Ontiveros, Cartillier, & Modarres, 2010). However, it was hypothesized that there was another distribution that modeled measurement error better. The log-logistic distribution was tested against the lognormal distribution as shown in Figure 4-1.

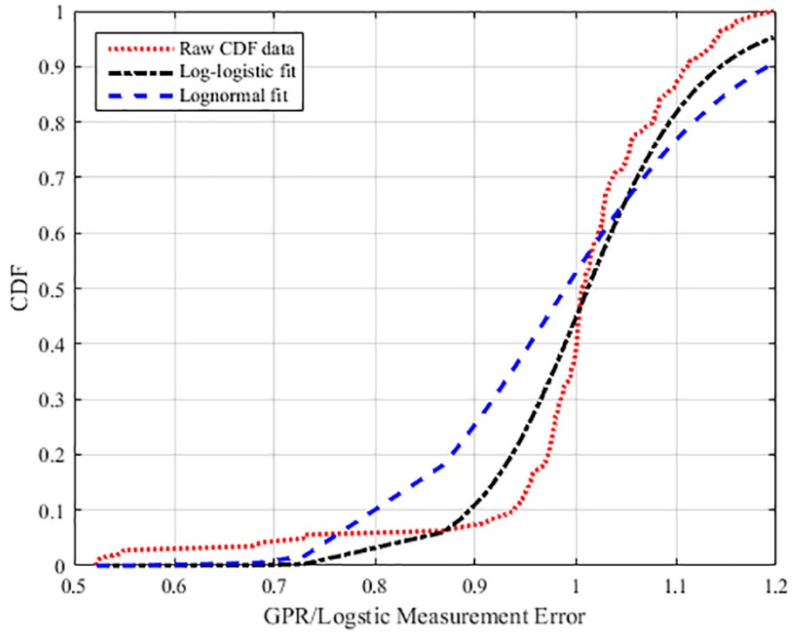

Figure 4-1: Comparison of the non-parametric CDF of measurement error (of a GPR/Logistic CPD model) against the parametric CDF of the lognormal distribution and the log-logistic distribution

The fitness test of the two showed that the log-logistic distribution has a higher average coefficient of determination ($R^2 = 0.93$) than the lognormal distribution ($R^2 = 0.30$) when checking against the raw measurement error distribution. As a result, this research adopts a log-logistic format of the distributions for the measurement errors $E_{a',e}$, $E_{a',m}$, and $E_{a',t}$,

$$f(E_{a',e}) \sim LOGLOGIST(\mu_e, \sigma_e) \tag{4.3}$$

$$f(E_{a',m}) \sim LOGLOGIST(\mu_m, \sigma_m) \tag{4.4}$$

$$f\left(E_{a\prime,t}\right) \sim LOGLOGIST\left(\mu_m - \mu_e, \sqrt{\sigma_m^2 + \sigma_e^2}\right) \tag{4.5}$$

where μ_e and σ_e are the experimental log-logistic parameters for $E_{a\prime,e}$ and μ_m and σ_m are the model log-logistic parameters for $E_{a\prime,m}$. This holds true for the following likelihood function from which Bayesian parameter estimation is used to estimate the model parameters μ_m and σ_m from a total of n validation points for each CPD model pair.

$$l\left(\frac{a_e}{a_m}, \mu_e, \sigma_e | \mu_m, \sigma_m\right) = \prod_{i=1}^{n} \frac{1}{\left(\frac{a_{e,i}}{a_{m,i}}\right)\sqrt{\sigma_m^2 + \sigma_e^2}} \frac{\exp\left[\frac{\ln\left(\frac{a_{e,i}}{a_{m,i}}\right) - (\mu_m - \mu_e)}{\sqrt{\sigma_m^2 + \sigma_e^2}}\right]}{\left\{1 + \exp\left[\frac{\ln\left(\frac{a_{e,i}}{a_{m,i}}\right) - (\mu_m - \mu_e)}{\sqrt{\sigma_m^2 + \sigma_e^2}}\right]\right\}^2} \tag{4.6}$$

The forty measurements described in Chapter 3 Section 3.4.3.5 (see Figure 3-8) and their conjoining model and experimental lengths are used to obtain the experimental log-logistic parameters by way of Bayesian parameter estimation using the likelihood function,

$$l\left(\frac{a_i}{a_{e,i}} | \mu_e, \sigma_e\right) = \prod_{i=1}^{40} \frac{1}{\left(\frac{a_i}{a_{e,i}}\right)\sigma_e} \frac{\exp\left[\frac{\ln\left(\frac{a_i}{a_{e,i}}\right) - \mu_e}{\sigma_e}\right]}{\left\{1 + \exp\left[\frac{\ln\left(\frac{a_i}{a_{e,i}}\right) - \mu_e}{\sigma_e}\right]\right\}^2} \tag{4.7}$$

based on Equation (4.3). Bayesian parameter estimation of Equation (4.7) produced a μ_e of 0.074 and a σ_e of 0.071 as the experimental log-logistic parameters.

4.3 Recursive Bayesian Estimation

The crack length data used for the GPR CPD model analysis is often limited in quantity. As stated in Chapter 3 Section 3.4.1, many of the photos used for data extraction are subject to motion blur so an average of twenty CP data may be obtained. Regardless, this data is sufficient in processing the GPR CPD model. However, for the PF/GPR CPD model, the full CP path can be extracted from a few CP data in addition to the AE signals cumulative count C and cumulative amplitude A, and the methodology described in Chapter 2 Section 2.2.2.4. This results in as many CP data as there are AE data, where for the Group 3 specimens can number between four hundred to twenty thousand data.

4.3.1 AE Intensity Based Crack Propagation

The approximation for the CP data extraction under PF is as follows. First an initial set of particles ($N_p = 1{,}000$) is generated based on the lognormal crack length distribution at the first cycle of detection. This distribution is based on a set of measurement estimates being taken at the first cycle of detection.

The next step is propagation of each particle based on the assumed propagation model of the crack. Since the previous study shows that the GPR CP model was the more realistic of the first three models of this research (Smith & Modarres, 2017) this will be the assumed CP model for this step of the PF analysis. Each particle is propagated to the next cycle by way of the following,

$$a_r^i = a_{r-1}^i + \Delta a_{r-1} \times \exp[NORRND(1,0,\sigma_{noise})] \tag{4.8}$$

where exp$[NORRND(1,0,\sigma_{noise})]$ is a random noise parameter selected for each particle i, and σ_{noise} is the noise standard deviation extracted from parameter g_{23} of the GPR CP parameter set $\vec{A}$ (see Chapter 3 Section 3.4.5) which represents noise. As stated in Chapter Section 3.4.6, this noise σ_{noise} may also be a source of uncertainty in this particular CP model.

The third step of the PF analysis is computing the updated distribution and weight of each of the propagated particles. The updated distribution is based on the two AE signals which can also be defined as AE Intensity Equation (2.7). Both signals as well as the intensity follow a power trend based on the AE data available, so the updated distribution may be defined as a lognormal distribution where,

$$\ln\left(\frac{A_r C_r}{A_0}\right) = \frac{1}{\beta}\ln(a_r) - \frac{1}{\beta}\ln(\alpha) + NOR(0,\sigma_v) \tag{4.9}$$

Equation (4.40) is acquired by taking the natural log of both sides of Equation (2.9) and replacing AE intensity with Equation (2.7). With the log-mean μ_r defined as $\frac{1}{\beta}\ln(a_r) - \frac{1}{\beta}\ln(\alpha)$ and the log-standard deviation defined as σ_v, the updated distribution is stated as,

$$\pi(C_r, A_r|a_r) = \frac{1}{A_r C_r \sigma_v \sqrt{2\pi}} \times \exp\left\{-\frac{1}{2}\left[\frac{\ln\left(\frac{A_r C_r}{A_0}\right) - \mu_r}{\sigma_v}\right]^2\right\} \tag{4.10}$$

and the weights are updated directly by Eq. (2.27).

The final step is adjusting for the degeneracy problem that was addressed in Chapter 2 Section 2.2.2.4. This is done by resampling to generate a new set of particles based on the updated weight. This is done by checking if the probability $\Pr\left(x_r^{i*} = x_r^j\right)$ is equal to w_r^j.

If this condition is true then the weight for this step r and particle j is reset to $\frac{1}{N_p}$. Then the total weights for step r are normalized by $\vec{w}_r = \vec{w}_r / \sum \vec{w}_r$ so that the sum of the weights is equal to 1.

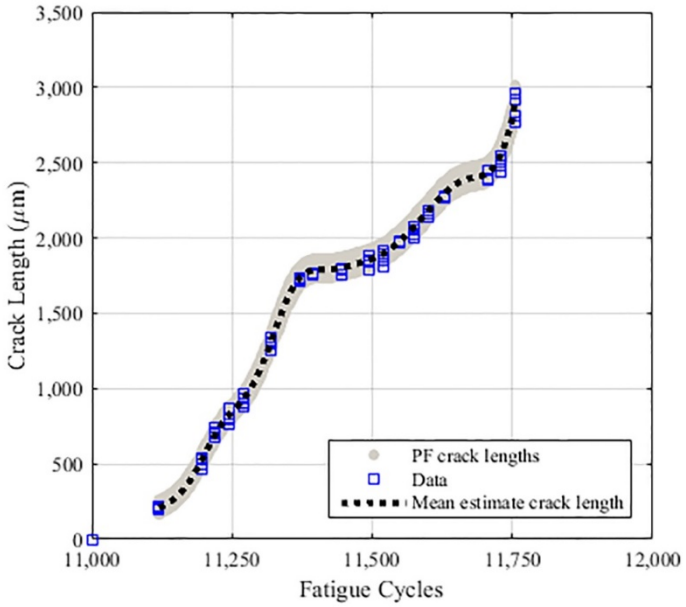

Figure 4-2: The PF analysis on raw crack length data from dog-bone specimen 5A3

This process is repeated for all steps r until all the AE data has been used. Figure 4-2 provides an example of the output from the PF updating process. The mean PF propagation curve is what is used for the PF/GPR CPD analysis.

4.4 Summary

The topics and methodologies covered in this chapter deals primarily with the relation between the CSFs and the CP and validation of the models and CSFs chosen. This relation has been proven to predict probable CP based on a known CSF-to-CPD correlation within a reasonable modeling error range (Smith, Modarres, & Droguett, 2017). Proof of this came from a utilization of the Ontiveros et. al methodology (Ontiveros, Cartillier, &

Modarres, 2010) where model error is estimated. A new development to this particular methodology is the finding that there are other distributions besides the lognormal distribution that better represent model error. For this research it is the log-logistic distribution, but for other research this might be a completely different distribution.

The CSF-to-CPD PF/GPR CP methodology where the GPR CP step model used for the procedure. Because of the wealth of AE data, the PF/GPR CP model is able to be estimated to a higher degree of completeness than the GPR CP model which is limited to the few detected crack measurements. All of the methodologies covered are used for the validation phase of this research and whose results are covered in the next chapter of this dissertation.

Chapter 5: Results and Discussion

5.1 Overview

The final step in this research is the completed analysis of the specimen data using the models and methodologies defined from Chapter 2 through Chapter 5. The first step is performing the Bayesian analysis on each specimen on each of the pertinent CPD model pairs. These results are then divided into training and validation data in order to check for the relations that exist between the CSFs and the CPD model posterior parameters. In addition to this the model error is computed and checked for correctness in the specimen data set aside for validation. Finally, through the model error results as well as the CSF-to-CPD correlation, an end-of-life analysis is performed which will estimate the approximate cycle of failure.

5.2 Bayesian Parameter Analysis

The Bayesian analysis for the CPD parameters is done by a routine designed in MATLAB (Mathworks, 2015). This routine (see Appendix B.1) makes use of the Metropolis-Hastings (MH) (Hastings, 1970) Sampling command[4] to process the Bayesian inference defined in Equation (3.7). The routine also makes use of a MATLAB software package called GPML (Rasmussen, Nickisch, & Williams, 2015) for the GPR (Chapter 2 Section 2.2.2.3) and PF/GPR (Chapter 2 Section 2.2.2.4) CP models and for the validation of the posterior CP parameters with respect to the CSFs. The target parameter set $\vec{\Theta}$ for updating has been defined as,

[4] The MATLAB command *mhsample* performs the Metropolis-Hastings algorithm https://au.mathworks.com/help/stats/mhsample.html

$$\vec{\Theta} = \left[\vec{A}, \vec{B}, P_{FD}\right]^T \quad (5.1)$$

in Chapter 2.

The posterior CPD model parameters are obtained for two training sets of data (defined in the next section) which are used for the correlation methodology (see Chapter 4 Section 4.2.1) for the validation sets. After the validation and model error methodology (Chapter 4 Section 4.2.2) is performed, the following results are gathered: The posterior CP and POD curves, the model error between the different CPD models, and the estimated end-of-life for the specimens.

5.2.1 Training and Validation Data Sets

The specimen data is divided into a training data set and a validation data set where a larger proportion of the data is reserved for the training data set. In analyzing the GPR CPD models, the specimens are divided as follows:

- Training Set: DB3, DB4, DB5, DB6, DB7, DB15, 1A2, 1B3, 5A2, 5A3, 5A4, 5A6, 5A8, 5A9, 5A20, 5A22, 5A23
- Validation Set: 5A10, 5A21, 5A24, 5A26

Meanwhile in analyzing the PF/GPR CPD models, the specimens are divided as follows:

- Training Set: 5A2, 5A3, 5A4, 5A6, 5A8, 5A9, 5A20, 5A22, 5A23
- Validation Set: 5A10, 5A21, 5A24, 5A26

The number of training data per specimen for the Bayesian analyses for the CPD models varies as shown in Table 5-1.

Table 5-1: Count of training data ($D = 1$) per specimen per CPD model

	Training Data	
	GPR	**PF/GPR**
DB3	30	N/A
DB4	16	N/A
DB5	45	N/A
DB6	7	N/A
DB7	108	N/A
DB15	14	N/A
1A2	20	N/A
1B3	42	N/A
5A2	20	650
5A3	64	226
5A4	28	225
5A6	28	225
5A8	72	225
5A9	24	225
5A20	48	225
5A22	32	225
5A23	20	225

In the case of the GPR CPD models, data was limited to the detected data only as explained in Chapter 3 Section 3.4.3. For the PF/GPR CPD models, a larger amount of data could be generated per sample. However, this too had to be limited because the GPML MATLAB code (Rasmussen, Nickisch, & Williams, 2015) has an upper limit to how much training data it can process at a time.

5.2.2 Posterior Distribution for Sets 1 and 2

The Bayesian analysis for the first two specimen sets take the priors defined in Chapter 3 Section 3.2.6. It was discovered that the hyper-parameters for the CP models throughout the test results don't show much difference from one result to another. In the case of specimen DB7 for example, the standard deviation between the GPR CP model hyper-parameters A_1 and A_2 are 0.0318 and 0.414, respectively. This is further exemplified by

the mean CP models for specimen DB7 presented in Figure 5-1, where the difference is highly minimal.

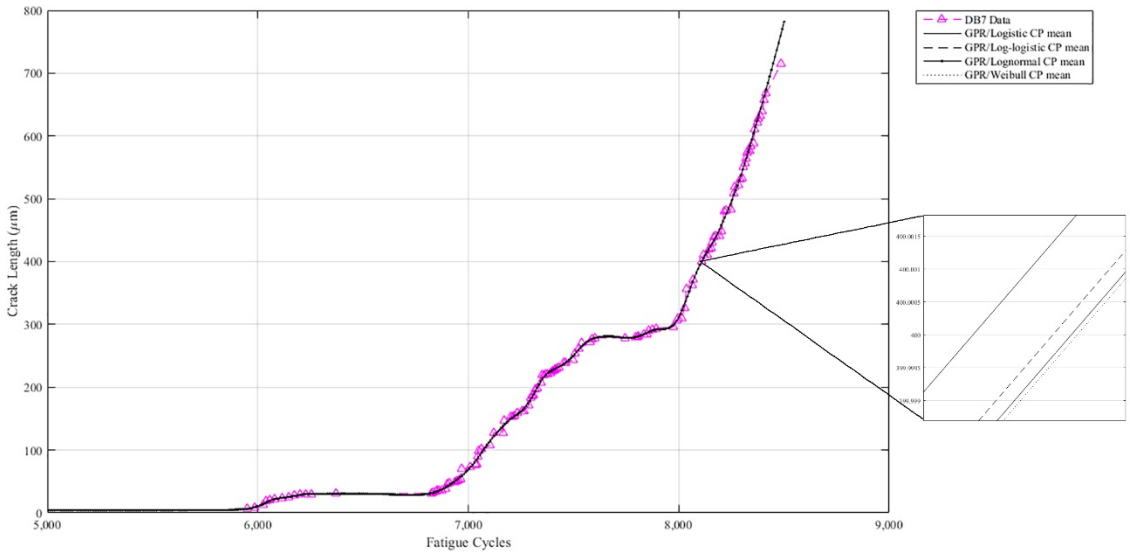

Figure 5-1: The CP posterior models for each of the four GPR CPD model pairs for specimen DB7

Thus the mean value of the posterior CPD parameter sets $\vec{\Theta}$ are presented in Table 5-2 and Table 5-3 and pictorially in Figure 5-2 and Figure 5-3.

Table 5-2: The mean GPR propagation model posterior hyper parameter values for specimen Sets 1 and 2

Hyper parameter	DB3	DB4	DB5	DB6	DB7	DB15	1B3	1A2
g_1	1.02	0.98	1.00	0.04	1.05	1.00	0.01	0.08
g_2	501	642	164	7.20×10^6	1.97×10^{12}	115	374	977
g_3	2.46×10^7	4.46×10^3	202	1.90×10^4	5.80×10^9	2.30×10^3	1.07×10^5	3.62×10^5
g_4	1.46×10^4	2.75×10^5	590	6.02×10^6	1.39×10^{12}	6.67×10^3	1.34×10^6	8.34×10^5
g_5	1.00	1.15	1.00	2.31	1.00	1.00	12.2	5.56
g_6	0.60	20.5	1.00	71.7	0.66	1.00	3.60×10^3	24.9
g_7	1.76×10^3	90.0	2.62	2.59×10^3	7.57×10^3	4.07	1.07×10^4	1.07×10^3
g_8	0.95	3.82	1.00	31.89	0.87	1.00	149	25.3
g_9	11.5	27.3	1.01	496	4.97	0.97	931	167
g_{10}	0.96	4.90	1.00	88.4	0.83	1.00	281	19.2
g_{11}	1.08	21.3	57.31	1.00	2.55×10^4	36.8	1.00	0.08
$g_{12} \dots g_{19}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
g_{20}	1.02×10^{-4}	1.75×10^{-8}	0.42	0.02	7.82×10^{-7}	0.44	0.08	4.63×10^{-3}
g_{21}	3.19	23.9	0.99	39.9	13.7	0.13	0.06	86.6
g_{22}	1.77	1.62	0.01	5.90	1.94	0.58	0.57	2.90
g_{23}	1.02×10^{-4}	1.80×10^{-8}	0.07	0.02	0.01	0.03	0.08	0.03

Table 5-3: The mean POD model posterior hyper parameter values for specimen Sets 1 and 2 based on GPR CP model

CPD	Hyper parameter	DB3	DB4	DB5	DB6	DB7	DB15	1B3	1A2
GPR/Logistic	η_0	0.12	2.28	0.12	0.12	0.17	3.24	2.13	2.11
	η_1	0.20	−0.14	0.20	0.20	0.19	−0.12	−0.0023	−0.074
	P_{FD}	0.21	0.29	0.20	0.20	0.09	0.13	0.040	0.060
GPR/Log-logistic	β_0	0.12	2.28	0.12	0.12	0.17	3.24	2.13	2.11
	β_1	0.20	−0.14	0.20	0.20	0.19	−0.12	−0.0023	−0.074
	P_{FD}	0.21	0.29	0.20	0.20	0.09	0.13	0.040	0.060
GPR/Lognormal	ζ_0	0.14	0.65	0.12	0.12	0.12	0.51	−0.15	0.41
	ζ_1	0.21	1.27	0.20	0.20	0.20	0.81	0.67	1.04
	P_{FD}	0.38	0.67	0.20	0.20	0.20	0.67	0.042	0.063
GPR/Weibull	α	0.12	1.09	0.12	0.12	0.12	8.30	3.03	6.83
	β	0.17	0.34	0.20	0.20	0.21	0.53	0.49	0.63
	P_{FD}	0.25	0.72	0.20	0.20	0.20	0.41	0.058	0.064

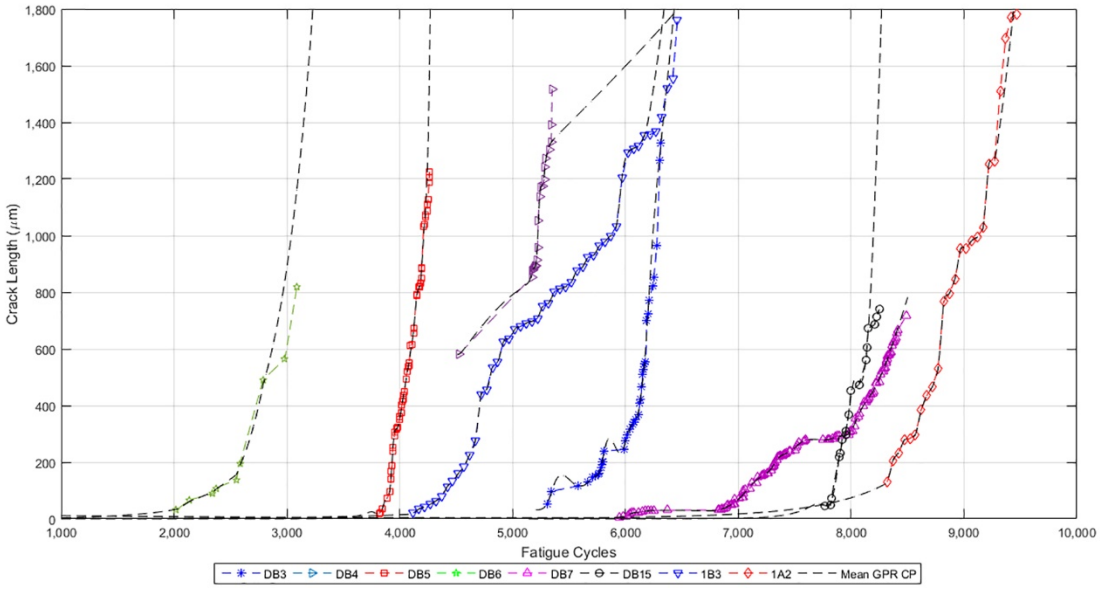

Figure 5-2: The GPR CP posterior mean models for specimen Sets 1 and 2

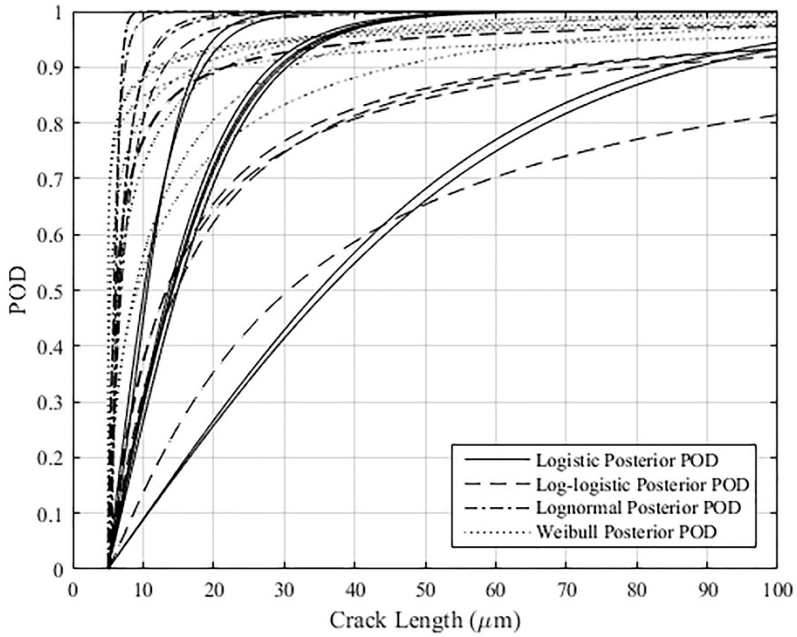

Figure 5-3: The POD posterior mean models for specimen Sets 1 and 2

5.2.3 Prior Distribution for Set 3

Obtaining the prior distribution for the third set of specimens involves an intensive goodness-of-fit analysis to determine which density functions best represents the behavior

of the posterior distributions of parameters from the first and second sets of the specimens from the previous section. The process involves first grouping all the like-posterior parameters from the $\vec{\Theta}$ sets (for example g_1 with g_1 and g_2 with g_2). Then from each parameter group the non-parametric cumulative distribution is obtained and compared to a set of likely density functions as seen in the example in Figure 5-4.

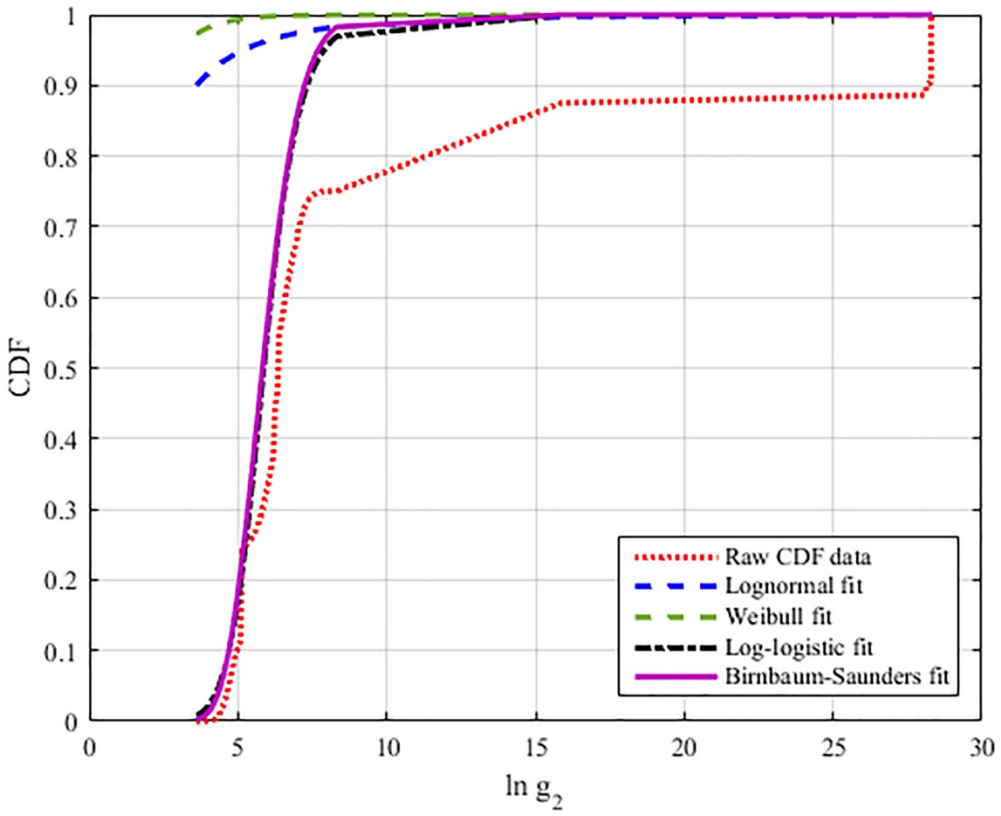

Figure 5-4: Non-parametric CDF of GPR/Logistic CPD parameter ln g_2 and parametric CDF distribution fits

Selection of the appropriate distribution was based on two things. The first is the given range of the parameter groups. For instance, in the example in Figure the GPR/Logistic CPD parameter ln g_2 ranges in the positive region so distributions such as lognormal,

Weibull, log-logistic, and Birnbaum-Saunders would be considered. While another parameter (GPR/Logistic CPD parameter $\ln g_4$) that ranges between the positive and negative region would consider the normal or logistic distributions as options. Second, the coefficient of determination R^2 is computed for each option and the density function that yields the highest value is selected as the prior distribution fit for the second group. The proposed distributions in Figure 5-4 for example have the R^2 values:

- $R^2_{Birnbaum-Saunders} = 0.616$
- $R^2_{log-logistic} = 0.660$
- $R^2_{lognormal} = -5.32$[5]
- $R^2_{Weibull} = -6.11$[5]

Thus based on the values, the log-logistic would be the best fit for that CPD parameter. Each CPD model therefore has its own unique prior distribution. The second group prior distributions for the thirteen specimens fitting specimen Set 3 are shown in Table 5-3, Table 5-4, Table 5-5, and Table 5-6 by CPD model.

[5] The general definition of the coefficient of determination is used where a negative value means that the data is a poor fit to the proposed model https://en.wikipedia.org/wiki/Coefficient_of_determination

Table 5-4: Prior distribution for GPR/logistic and PF/GPR/logistic CPD Model

Parameter	Distribution	Hyper-Parameter 1	Hyper-Parameter 2
η_0	Birnbaum-Saunders	0.084	0.96
η_1	Lognormal	-0.47	1.25
$ln\, g_1$	Logistic	-1.15	1.12
$ln\, g_2$	Log-logistic	1.98	0.32
$ln\, g_3$	Extreme Value	14.94	5.94
$ln\, g_4$	Gamma	5.55	2.45
$ln\, g_5$	Normal	0.56	1.12
$ln\, g_6$	Logistic	1.68	1.51
$ln\, g_7$	Gamma	2.09	2.70
$ln\, g_8$	Logistic	1.37	1.22
$ln\, g_9$	Logistic	2.89	1.59
$ln\, g_{10}$	Logistic	1.26	1.27
$ln\, g_{11}$	Logistic	1.35	2.08
$ln\, g_{12}$	Extreme Value	3.19×10^{-9}	4.11×10^{-9}
$ln\, g_{13}$	Extreme Value	-4.38×10^{-10}	1.69×10^{-9}
$ln\, g_{14}$	Extreme Value	-4.02×10^{-10}	1.70×10^{-9}
$ln\, g_{15}$	Extreme Value	-5.58×10^{-10}	1.62×10^{-9}
$ln\, g_{16}$	Extreme Value	1.02×10^{-9}	1.62×10^{-9}
$ln\, g_{17}$	Extreme Value	5.34×10^{-10}	2.40×10^{-9}
$ln\, g_{18}$	Extreme Value	2.51×10^{-9}	3.72×10^{-9}
$ln\, g_{19}$	Extreme Value	2.64×10^{-9}	4.00×10^{-9}
$ln\, g_{20}$	Extreme Value	-4.31	4.62
$ln\, g_{21}$	Normal	1.23	2.63
$ln\, g_{22}$	Normal	-0.080	1.70
$ln\, g_{23}$	Logistic	-5.16	2.77
P_{FD}	Beta	0.44	2.80

Table 5-5: Prior distribution for GPR/log-logistic and PF/GPR/log-logistic CPD Model

Parameter	Distribution	Hyper-Parameter 1	Hyper-Parameter 2
β_0	Lognormal	-0.60	1.46
β_1	Extreme Value	0.15	0.16
$ln\ g_1$	Extreme Value	-0.32	1.62
$ln\ g_2$	Log-logistic	1.97	0.32
$ln\ g_3$	Extreme Value	14.4	5.85
$ln\ g_4$	Gamma	5.42	2.49
$ln\ g_5$	Normal	0.69	1.11
$ln\ g_6$	Logistic	1.68	1.51
$ln\ g_7$	Gamma	2.15	2.59
$ln\ g_8$	Logistic	1.31	1.25
$ln\ g_9$	Logistic	3.27	1.78
$ln\ g_{10}$	Logistic	1.60	1.45
$ln\ g_{11}$	Extreme Value	3.95	4.33
$ln\ g_{12}$	Extreme Value	2.06×10^{-9}	2.54×10^{-9}
$ln\ g_{13}$	Extreme Value	-9.35×10^{-10}	2.60×10^{-9}
$ln\ g_{14}$	Extreme Value	3.51×10^{-9}	6.05×10^{-9}
$ln\ g_{15}$	Extreme Value	-2.15×10^{-12}	2.10×10^{-9}
$ln\ g_{16}$	Extreme Value	1.08×10^{-9}	2.23×10^{-9}
$ln\ g_{17}$	Extreme Value	1.45×10^{-9}	2.96×10^{-9}
$ln\ g_{18}$	Extreme Value	1.92×10^{-9}	2.82×10^{-9}
$ln\ g_{19}$	Extreme Value	3.20×10^{-9}	5.55×10^{-9}
$ln\ g_{20}$	Extreme Value	-4.21	4.46
$ln\ g_{21}$	Extreme Value	2.47	2.06
$ln\ g_{22}$	Normal	-0.08	1.70
$ln\ g_{23}$	Extreme Value	-4.07	2.78
P_{FD}	Beta	0.47	2.70

Table 5-6: Prior distribution for GPR/lognormal and PF/GPR/lognormal CPD Model

Parameter	Distribution	Hyper-Parameter 1	Hyper-Parameter 2
ζ_0	Logistic	0.20	0.16
ζ_1	Weibull	0.63	1.28
$ln\, g_1$	Extreme Value	-0.41	1.55
$ln\, g_2$	Log-logistic	1.97	0.32
$ln\, g_3$	Gamma	5.78	1.90
$ln\, g_4$	Log-logistic	2.49	0.26
$ln\, g_5$	Logistic	0.45	0.52
$ln\, g_6$	Logistic	1.97	1.82
$ln\, g_7$	Gamma	2.21	2.36
$ln\, g_8$	Logistic	1.35	1.18
$ln\, g_9$	Logistic	2.52	1.52
$ln\, g_{10}$	Logistic	1.23	1.33
$ln\, g_{11}$	Logistic	2.05	2.22
$ln\, g_{12}$	Extreme Value	-1.13×10^{-11}	3.19×10^{-10}
$ln\, g_{13}$	Extreme Value	3.26×10^{-10}	6.38×10^{-10}
$ln\, g_{14}$	Extreme Value	1.34×10^{-10}	3.73×10^{-10}
$ln\, g_{15}$	Extreme Value	9.41×10^{-11}	2.81×10^{-10}
$ln\, g_{16}$	Extreme Value	3.23×10^{-10}	4.68×10^{-10}
$ln\, g_{17}$	Extreme Value	1.85×10^{-10}	2.88×10^{-10}
$ln\, g_{18}$	Extreme Value	1.14×10^{-10}	2.84×10^{-10}
$ln\, g_{19}$	Extreme Value	2.86×10^{-10}	3.81×10^{-10}
$ln\, g_{20}$	Extreme Value	-4.00	4.14
$ln\, g_{21}$	Extreme Value	2.53	1.96
$ln\, g_{22}$	Extreme Value	0.64	1.15
$ln\, g_{23}$	Extreme Value	-3.95	2.39
P_{FD}	Beta	0.54	1.32

Table 5-7: Prior distribution for GPR/Weibull and PF/GPR/Weibull CPD Model

Parameter	Distribution	Hyper-Parameter 1	Hyper-Parameter 2
α	Lognormal	-0.47	1.80
β	Birnbaum-Saunders	0.30	0.59
$ln\, g_1$	Logistic	-1.07	1.14
$ln\, g_2$	Log-logistic	1.97	0.32
$ln\, g_3$	Extreme Value	14.6	5.83
$ln\, g_4$	Gamma	5.45	2.48
$ln\, g_5$	Normal	0.67	1.10
$ln\, g_6$	Logistic	1.43	1.49
$ln\, g_7$	Extreme Value	7.04	2.57
$ln\, g_8$	Logistic	1.29	1.22
$ln\, g_9$	Normal	3.05	2.63
$ln\, g_{10}$	Normal	1.91	2.50
$ln\, g_{11}$	Normal	1.86	3.83
$ln\, g_{12}$	Extreme Value	2.46×10^{-9}	3.52×10^{-9}
$ln\, g_{13}$	Extreme Value	-5.01×10^{-11}	1.66×10^{-9}
$ln\, g_{14}$	Extreme Value	3.44×10^{-9}	5.47×10^{-9}
$ln\, g_{15}$	Extreme Value	1.73×10^{-9}	5.21×10^{-9}
$ln\, g_{16}$	Extreme Value	1.86×10^{-9}	3.75×10^{-9}
$ln\, g_{17}$	Extreme Value	6.66×10^{-10}	1.82×10^{-9}
$ln\, g_{18}$	Extreme Value	1.47×10^{-9}	3.31×10^{-9}
$ln\, g_{19}$	Extreme Value	4.94×10^{-9}	7.82×10^{-9}
$ln\, g_{20}$	Normal	-7.08	6.12
$ln\, g_{21}$	Normal	1.23	2.65
$ln\, g_{22}$	Normal	-0.16	1.83
$ln\, g_{23}$	Normal	-6.19	5.22
P_{FD}	Beta	0.53	1.56

As a point of reference to the first group prior (Figure 3-14), the second prior distribution for the CD models are presented in Figure 5-5.

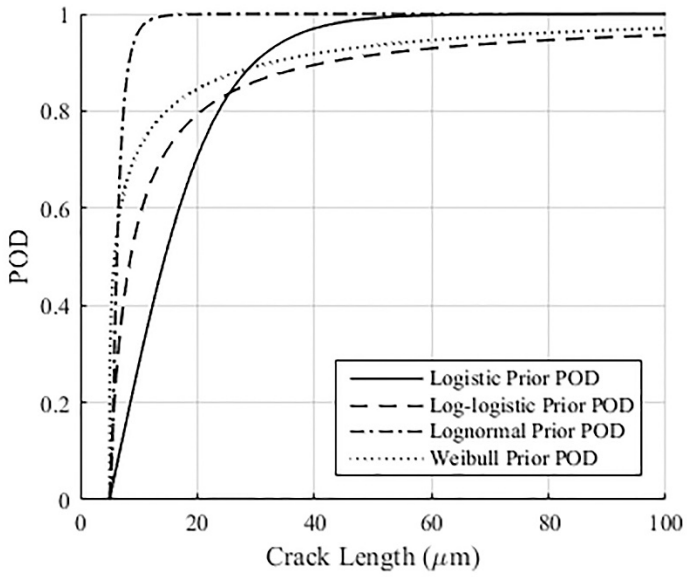

Figure 5-5: The second group prior logistic, log-logistic, lognormal, and Weibull POD mean curves for GPR and PF/GPR CP models

5.2.4 Posterior Results from CPD Analysis

Similarly to Section 5.2.2, the Bayesian posterior distributions for the third set of specimen tests are laid out in the following tables and figures where Table 5-8, Table 5-10, and Figure 5-6 represent the results from GPR based CPDs and Table 5-9, Table 5-11, and Figure 5-7 represent the results from the PF/GPR based CPDs. These posterior results only include the training parameter sets and models however.

Table 5-8: The mean GPR propagation model posterior hyper parameter values for specimen Set 3

Hyper parameter	**5A2**	**5A3**	**5A4**	**5A6**	**5A8**	**5A9**	**5A20**	**5A22**	**5A23**
g_1	0.236	0.347	0.289	0.239	0.334	0.627	0.869	0.357	0.296
g_2	681	1.24×10^7	6.25×10^7	443	416	7.76×10^3	2.45×10^3	1.91×10^3	9.10×10^4
g_3	2.91×10^6	2.38×10^5	3.55×10^5	$4.90E \times 10^6$	3.49×10^6	8.68×10^3	4.51×10^6	6.72×10^4	4.80×10^5
g_4	1.76×10^6	1.51×10^7	3.75×10^5	1.33×10^6	3.53×10^6	4.11×10^5	2.09×10^5	4.27×10^4	2.72×10^7
g_5	1.86	1.74	1.65	1.86	2.00	1.45	1.16	1.56	30.4
g_6	12.7	5.04	8.36	10.7	5.30	2.26	1.78	5.76	10.1
g_7	4.62×10^3	227	271	1.02×10^3	380	100	272	335	678
g_8	7.90	3.55	3.38	4.08	3.38	2.04	1.18	5.32	5.44
g_9	76.6	15.9	29.3	20.9	29.6	8.12	1.70	18.1	79.2
g_{10}	13.2	4.79	5.96	11.0	6.80	1.45	3.90	4.32	21.4
g_{11}	5.14	5.19	113	1.25	1.11×10^3	109	0.875	4.18	1.61
$g_{12} \dots g_{19}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
g_{20}	1.43×10^{-4}	0.0118	0.0325	9.26×10^{-5}	0.0746	0.112	4.95×10^{-4}	2.47×10^{-4}	0.0191
g_{21}	9.52	16.3	97.3	3.54	64.5	4.33	58.1	90.3	198
g_{22}	0.581	1.49	1.11	0.078	0.772	0.711	0.465	1.07	2.32
g_{23}	0.0437	0.0153	0.0145	0.0251	0.0126	5.25×10^{-3}	5.42×10^{-3}	9.61×10^{-3}	0.0125

Table 5-9: The mean PF/GPR propagation model posterior hyper parameter values for specimen Set 3

Hyper parameter	5A2	5A3	5A4	5A6	5A8	5A9	5A20	5A22	5A23
g_1	1.05×10^3	5.17×10^8	4.56×10^4	1.17×10^6	649	83.2	699	646	7.51×10^3
g_2	1.34×10^5	3.50×10^6	5.45×10^5	1.13×10^5	5.28×10^5	9.46×10^4	4.65×10^4	9.42×10^3	2.81×10^6
g_3	1.29×10^6	2.79×10^6	5.16×10^4	2.39×10^5	9.51×10^5	7.66×10^6	3.84×10^5	3.03×10^5	1.56×10^7
g_4	1.85	1.40	1.56	2.29	4.52	1.42	1.20	1.42	1.16
g_5	6.05	8.58	5.26	3.39	122	2.44	26.0	7.64	21.1
g_6	1.33×10^3	101	404	792	194	124	582	1.18×10^3	457
g_7	3.81	8.39	3.24	6.75	3.97	2.36	1.75	2.57	3.97
g_8	43.4	26.6	37.9	13.9	30.1	1.38	3.28	13.9	319
g_9	5.38	7.13	3.11	11.6	5.20	2.86	1.51	4.36	4.30
g_{10}	333	23.1	7.38×10^3	0.822	2.99	2.91	2.95	1.19×10^3	1.02
$g_{11} \cdots g_{19}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
g_{20}	0.302	1.50×10^{-4}	0.181	4.56×10^{-6}	1.29×10^{-9}	0.0513	2.16×10^{-5}	5.93	0.0358
g_{21}	24.9	53.6	32.5	391	19.7	2.59	79.7	19.5	3.94
g_{22}	0.610	3.45	0.215	0.598	0.333	0.577	1.16	0.283	3.00
g_{23}	0.0276	9.75×10^{-3}	0.0177	6.65×10^{-3}	0.0154	0.0273	4.52×10^{-3}	8.94×10^{-3}	2.83×10^{-3}

Table 5-10: The mean POD model posterior hyper parameter values for specimen Set 3 based on GPR CP model

CPD	Hyper parameter	5A2	5A3	5A4	5A6	5A8	5A9	5A20	5A22	5A23
GPR/Logistic	η_0	0.069	0.091	0.13	0.062	0.065	0.14	0.058	0.072	0.066
	η_1	0.60	1.10	0.63	0.64	0.54	0.53	0.34	0.29	0.48
	P_{FD}	0.43	0.34	0.18	0.63	0.81	0.14	0.21	0.20	0.19
GPR/Log-logistic	β_0	0.42	0.89	3.62	0.49	1.59	0.97	0.77	0.58	0.53
	β_1	0.06	0.10	0.10	0.10	0.10	0.09	0.05	0.07	0.12
	P_{FD}	0.63	0.39	0.29	0.81	0.91	0.27	0.33	0.28	0.29
GPR/Lognormal	ζ_0	0.49	0.34	0.37	0.25	0.37	0.34	0.32	0.16	0.10
	ζ_1	0.56	0.59	0.81	2.01	5.61	0.50	0.60	0.32	0.54
	P_{FD}	0.78	0.64	0.72	0.89	0.96	0.62	0.37	0.55	0.96
GPR/Weibull	α	0.33	0.33	1.08	0.17	0.36	1.20	0.83	0.62	0.17
	β	0.33	0.32	0.38	0.24	0.36	0.28	0.37	0.33	0.40
	P_{FD}	0.76	0.59	0.86	0.90	0.96	0.54	0.38	0.47	0.96

Table 5-11: The mean POD model posterior hyper parameter values for specimen Set 3 based on PFGPR CP model

CPD	Hyper parameter	5A2	5A3	5A4	5A6	5A8	5A9	5A20	5A22	5A23
PF/GPR/Logistic	η_0	0.074	0.036	0.084	0.091	0.056	0.080	0.08	0.11	0.080
	η_1	0.62	0.28	0.38	0.20	0.90	0.65	0.56	1.35	1.67
	P_{FD}	0.56	0.30	0.48	0.91	0.80	0.16	0.19	0.15	0.88
PF/GPR/Log-logistic	β_0	0.60	2.30	1.36	0.71	0.78	1.52	0.30	0.36	1.52
	β_1	0.04	0.09	−0.02	0.12	0.03	0.06	−0.01	0.08	1.22
	P_{FD}	0.63	0.15	0.30	0.76	0.89	0.30	0.17	0.18	0.67
PF/GPR/Lognormal	ζ_0	0.23	0.44	0.18	0.18	0.23	0.21	0.19	0.21	0.21
	ζ_1	7.43	0.56	0.53	3.45	4.96	0.61	0.50	0.58	0.35
	P_{FD}	0.78	0.72	0.80	0.92	0.96	0.54	0.38	0.39	0.95
PF/GPR/Weibull	α	0.91	1.81	0.18	3.61	0.25	1.25	0.27	0.61	0.47
	β	0.38	0.34	0.62	0.38	0.31	0.39	0.32	0.36	0.36
	P_{FD}	0.81	0.25	0.78	0.85	0.96	0.47	0.43	0.38	0.96

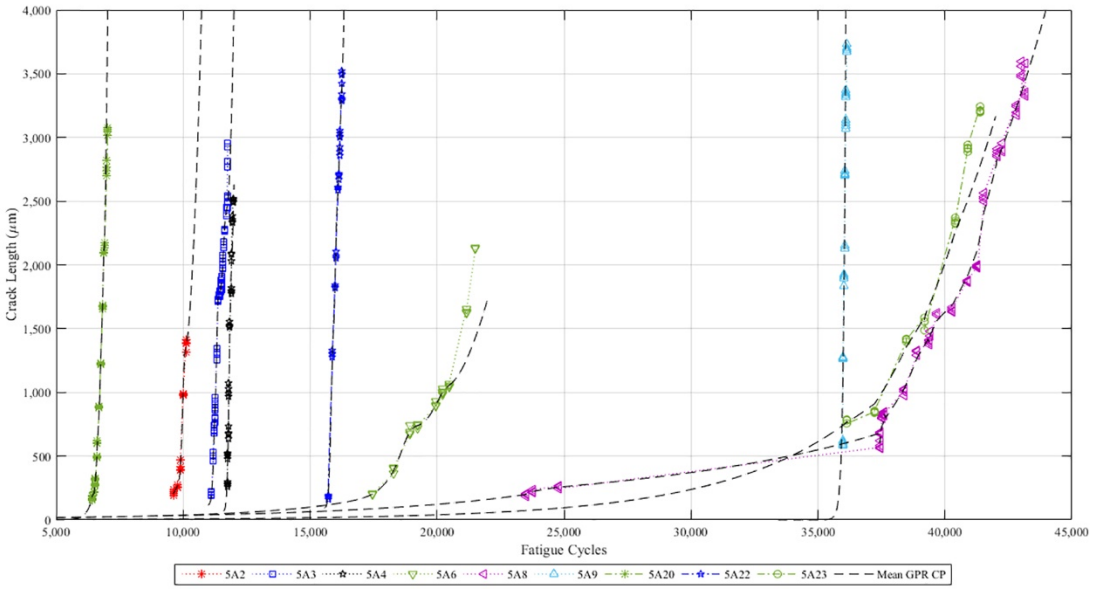

Figure 5-6: The GPR CP posterior mean models for specimen Set 3

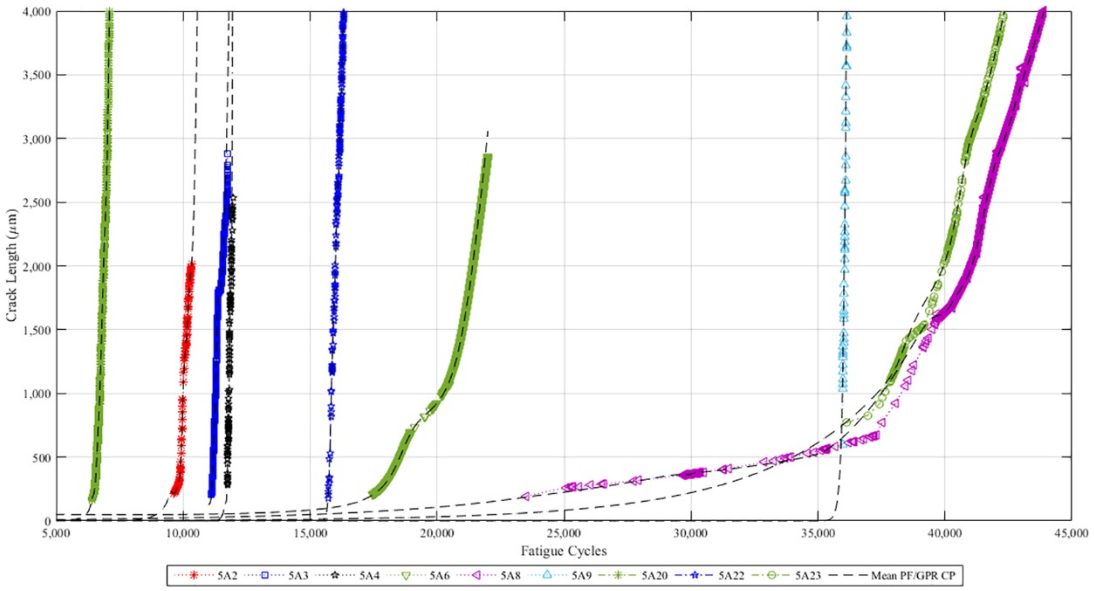

Figure 5-7: The PF/GPR CP posterior mean models for specimen Set 3

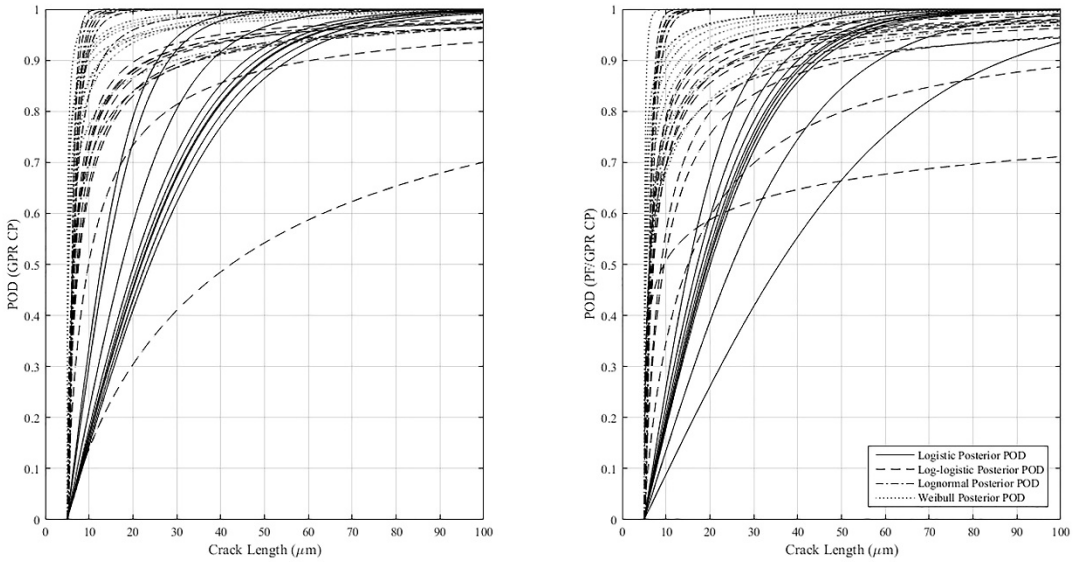

Figure 5-8: The POD posterior mean models for specimen Set 3

5.3 Crack Shaping Factor Correlation Analysis and Model Validation

The posterior parameters obtained from the Bayesian parameter analysis defined in Section 5.2 will now be used for the correlation and validation stage in this section.

5.3.1 Crack Shaping Factor-to-Crack Propagation and Detection Correlation

The CSF-to-CPD correlation step is performed according to the methodology defined in Chapter 4 Section 4.2.1. This is necessary for the validation step because the correlation parameter set $\vec{V}$ for each CPD parameter will estimate the GPR and PF/GPR parameters for the validation specimens. The CSF vs. CPD parameter plots generated (such as the one shown in Figure 5-8) shows that the kernel function Equation (4.1) designed for the methodology is effective.

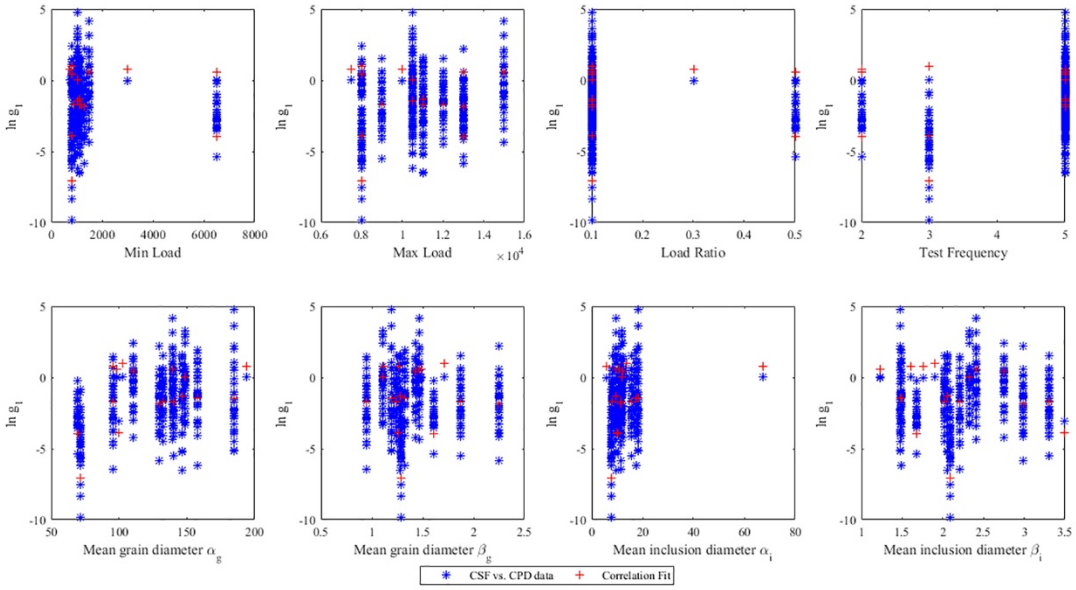

Figure 5-9: CSF vs. GPR/Logistic CPD parameter $\ln g_1$

The complete collection of correlation parameter sets are comprised of eight 20×25 size correlation matrices representing the eight CPD model sets under study (GPR or PFGPR/Logistic, Log-logistic, Lognormal, or Weibull). These correlation parameter sets are used to estimate the model parameters for the validation specimens. Table 5-12 through

Table 5-19 show the estimated CPDs parameters (P_{FD} estimate omitted for the results) for these validation specimens.

Table 5-12: The estimated GPR/Logistic CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
η_0	0.0716	0.0818	0.0628	0.0818
η_1	0.508	0.745	0.805	0.620
$ln\, g_1$	-1.01	-1.05	-2.17	-1.45
$ln\, g_2$	8.64	7.96	7.85	26.13
$ln\, g_3$	15.8	12.1	11.6	16.5
$ln\, g_4$	12.2	15.0	13.8	13.6
$ln\, g_5$	0.562	0.511	0.613	0.511
$ln\, g_6$	1.24	1.30	2.67	1.78
$ln\, g_7$	5.33	5.47	5.24	5.45
$ln\, g_8$	1.11	1.16	2.39	1.59
$ln\, g_9$	2.83	2.91	5.02	3.65
$ln\, g_{10}$	1.17	1.22	2.51	1.67
$ln\, g_{11}$	0.942	0.942	0.942	0.942
$ln\, g_{12}$	-9.20×10^{-10}	-1.16×10^{-10}	-9.82×10^{-11}	5.38×10^{-10}
$ln\, g_{13}$	-4.27×10^{-10}	-1.13×10^{-9}	-1.20×10^{-9}	4.28×10^{-10}
$ln\, g_{14}$	1.62×10^{-10}	-1.35×10^{-9}	3.33×10^{-10}	-1.50×10^{-9}
$ln\, g_{15}$	-5.92×10^{-10}	-1.56×10^{-9}	-1.22×10^{-9}	-3.39×10^{-10}
$ln\, g_{16}$	-8.07×10^{-10}	4.12×10^{-10}	3.78×10^{-10}	4.78×10^{-10}
$ln\, g_{17}$	1.81×10^{-9}	-5.58×10^{-10}	-1.86×10^{-10}	1.70×10^{-10}
$ln\, g_{18}$	5.07×10^{-10}	8.05×10^{-10}	-1.29×10^{-9}	1.08×10^{-9}
$ln\, g_{19}$	1.45×10^{-9}	-3.21×10^{-10}	3.02×10^{-10}	-1.03×10^{-9}
$ln\, g_{20}$	-7.19	-7.74	-6.55	-12.67
$ln\, g_{21}$	2.96	2.96	2.96	2.96
$ln\, g_{22}$	-1.73×10^{-4}	-1.50×10^{-4}	-1.61×10^{-4}	-1.66×10^{-4}
$ln\, g_{23}$	-3.97	-4.52	-6.02	-5.28

Table 5-13: The estimated GPR/Log-logistic CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
β_0	1.68	1.60	1.26	1.60
β_1	0.0788	0.0754	0.106	0.0755
$ln\, g_1$	-1.10	-1.15	-2.36	-1.57
$ln\, g_2$	7.32	7.56	6.90	7.65
$ln\, g_3$	11.7	11.8	11.7	11.7
$ln\, g_4$	13.0	14.6	13.0	14.7
$ln\, g_5$	0.622	0.625	0.627	0.623
$ln\, g_6$	0.79	1.24	3.50	1.86
$ln\, g_7$	6.13	7.14	7.36	6.49
$ln\, g_8$	1.15	1.20	2.48	1.65
$ln\, g_9$	2.40	2.51	5.17	3.44
$ln\, g_{10}$	1.20	1.26	2.59	1.73
$ln\, g_{11}$	0.900	0.913	0.873	1.621
$ln\, g_{12}$	-6.21×10^{-10}	-7.01×10^{-10}	-2.85×10^{-9}	2.21×10^{-10}
$ln\, g_{13}$	-1.27×10^{-11}	-2.08×10^{-9}	-7.31×10^{-10}	-1.98×10^{-9}
$ln\, g_{14}$	-2.61×10^{-9}	2.02×10^{-9}	1.00×10^{-9}	2.32×10^{-9}
$ln\, g_{15}$	-1.72×10^{-9}	5.71×10^{-11}	1.43×10^{-9}	-1.37×10^{-9}
$ln\, g_{16}$	5.75×10^{-10}	5.22×10^{-10}	-6.12×10^{-10}	1.29×10^{-9}
$ln\, g_{17}$	3.58×10^{-10}	2.67×10^{-10}	4.24×10^{-10}	-2.64×10^{-9}
$ln\, g_{18}$	-6.50×10^{-10}	7.48×10^{-10}	7.65×10^{-10}	1.51×10^{-10}
$ln\, g_{19}$	-1.34×10^{-9}	1.28×10^{-9}	1.37×10^{-9}	-5.23×10^{-10}
$ln\, g_{20}$	-6.94	-7.10	-6.78	-9.21
$ln\, g_{21}$	1.94	2.02	4.17	2.78
$ln\, g_{22}$	-3.05×10^{-3}	-1.54×10^{-7}	-1.82×10^{-1}	-1.56×10^{-9}
$ln\, g_{23}$	-4.55	-4.90	-5.82	-5.37

Table 5-14: The estimated GPR/Lognormal CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
ζ_0	0.252	0.251	0.252	0.251
ζ_1	1.00	0.974	0.501	0.768
$ln\, g_1$	-0.97	-1.01	-2.09	-1.39
$ln\, g_2$	8.46	8.62	8.46	8.54
$ln\, g_3$	11.3	11.6	11.4	11.3
$ln\, g_4$	12.8	13.3	13.0	13.0
$ln\, g_5$	0.572	0.597	1.230	0.820
$ln\, g_6$	1.30	1.35	2.79	1.86
$ln\, g_7$	5.04	5.04	5.05	5.03
$ln\, g_8$	0.99	1.04	2.14	1.42
$ln\, g_9$	1.85	1.93	3.99	2.66
$ln\, g_{10}$	1.14	1.19	2.45	1.64
$ln\, g_{11}$	0.419	0.425	0.407	0.755
$ln\, g_{12}$	-2.75×10^{-10}	-4.50×10^{-11}	-1.25×10^{-11}	-2.11×10^{-10}
$ln\, g_{13}$	-1.57×10^{-10}	-1.52×10^{-10}	-2.22×10^{-10}	1.55×10^{-10}
$ln\, g_{14}$	1.79×10^{-10}	-3.66×10^{-10}	-1.27×10^{-10}	-2.57×10^{-10}
$ln\, g_{15}$	-2.54×10^{-10}	2.28×10^{-10}	6.68×10^{-11}	1.63×10^{-10}
$ln\, g_{16}$	9.60×10^{-11}	1.44×10^{-10}	-1.02×10^{-12}	-1.28×10^{-10}
$ln\, g_{17}$	2.08×10^{-10}	2.50×10^{-11}	-5.38×10^{-11}	-4.36×10^{-11}
$ln\, g_{18}$	-6.45×10^{-11}	2.00×10^{-10}	-4.12×10^{-11}	1.78×10^{-10}
$ln\, g_{19}$	-2.20×10^{-11}	-2.82×10^{-10}	6.59×10^{-12}	-6.52×10^{-10}
$ln\, g_{20}$	-8.90	-8.90	-8.90	-8.90
$ln\, g_{21}$	1.73	1.81	3.73	2.48
$ln\, g_{22}$	-0.542	-0.487	-0.559	-0.487
$ln\, g_{23}$	-6.99	-7.01	-7.28	-7.10

Table 5-15: The estimated GPR/Weibull CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
α	0.939	0.921	0.988	0.917
β	0.267	0.341	0.388	0.298
$ln\, g_1$	-0.90	-0.93	-1.81	-1.24
$ln\, g_2$	9.80	9.00	7.67	9.06
$ln\, g_3$	12.0	12.3	11.6	17.9
$ln\, g_4$	14.7	15.4	13.2	15.4
$ln\, g_5$	0.476	0.476	0.477	0.476
$ln\, g_6$	1.46	1.52	3.17	2.10
$ln\, g_7$	6.08	6.10	6.09	6.69
$ln\, g_8$	0.95	0.99	2.04	1.36
$ln\, g_9$	2.16	2.26	4.65	3.10
$ln\, g_{10}$	1.54	1.61	3.31	2.21
$ln\, g_{11}$	1.13	1.13	1.13	1.13
$ln\, g_{12}$	-3.04×10^{-9}	1.39×10^{-9}	9.77×10^{-10}	-1.22×10^{-9}
$ln\, g_{13}$	1.40×10^{-9}	-1.92×10^{-9}	-1.58×10^{-9}	-1.81×10^{-9}
$ln\, g_{14}$	1.72×10^{-9}	1.11×10^{-9}	-7.09×10^{-10}	2.86×10^{-9}
$ln\, g_{15}$	5.59×10^{-11}	1.22×10^{-9}	-3.86×10^{-10}	3.59×10^{-9}
$ln\, g_{16}$	-3.32×10^{-10}	1.39×10^{-10}	-8.16×10^{-10}	1.44×10^{-9}
$ln\, g_{17}$	2.35×10^{-10}	-1.94×10^{-9}	-2.23×10^{-9}	-2.73×10^{-9}
$ln\, g_{18}$	-1.30×10^{-9}	-2.77×10^{-10}	1.09×10^{-9}	5.79×10^{-10}
$ln\, g_{19}$	-3.73×10^{-10}	2.24×10^{-9}	8.03×10^{-10}	2.03×10^{-9}
$ln\, g_{20}$	-8.48	-8.48	-8.48	-8.48
$ln\, g_{21}$	1.59	1.66	3.42	2.28
$ln\, g_{22}$	-0.403	-0.294	-0.452	-0.294
$ln\, g_{23}$	-6.76	-6.76	-6.76	-6.76

Table 5-16: The estimated PF/GPR/Logistic CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
η_0	0.0679	0.0654	0.0698	0.0656
η_1	0.828	0.828	0.828	0.828
$ln\, g_1$	-0.95	-0.96	-0.96	-0.95
$ln\, g_2$	8.07	8.07	8.07	8.07
$ln\, g_3$	8.8	8.8	8.8	8.8
$ln\, g_4$	9.3	14.3	13.3	11.4
$ln\, g_5$	0.997	0.558	-0.087	0.500
$ln\, g_6$	1.53	1.54	1.54	1.54
$ln\, g_7$	6.66	6.66	6.66	6.66
$ln\, g_8$	1.43	1.43	1.43	1.43
$ln\, g_9$	2.88	2.88	2.88	2.88
$ln\, g_{10}$	1.15	1.64	0.84	1.09
$ln\, g_{11}$	2.38	2.17	2.60	2.17
$ln\, g_{12}$	-3.00×10^{-9}	-3.06×10^{-10}	-3.03×10^{-9}	7.28×10^{-10}
$ln\, g_{13}$	-1.37×10^{-9}	3.20×10^{-10}	-2.15×10^{-9}	-8.59×10^{-10}
$ln\, g_{14}$	2.91×10^{-11}	-8.44×10^{-10}	-4.66×10^{-10}	-8.15×10^{-10}
$ln\, g_{15}$	8.44×10^{-10}	-1.02×10^{-9}	-3.49×10^{-9}	-1.19×10^{-9}
$ln\, g_{16}$	4.80×10^{-10}	2.37×10^{-9}	-3.49×10^{-10}	3.86×10^{-10}
$ln\, g_{17}$	1.64×10^{-9}	1.30×10^{-9}	1.63×10^{-9}	1.00×10^{-9}
$ln\, g_{18}$	2.47×10^{-9}	7.86×10^{-10}	3.67×10^{-9}	6.40×10^{-10}
$ln\, g_{19}$	1.82×10^{-9}	1.64×10^{-9}	$8.73E-10$	3.55×10^{-9}
$ln\, g_{20}$	-11.12	-10.11	-12.13	-10.11
$ln\, g_{21}$	3.24	2.94	3.53	2.94
$ln\, g_{22}$	-0.857	-0.856	-0.857	-0.858
$ln\, g_{23}$	-3.93	-4.21	-5.13	-4.62

Table 5-17: The estimated PF/GPR/Log-logistic CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
β_0	1.12	1.11	1.14	0.690
β_1	-0.005	0.121	0.472	0.309
$ln\, g_1$	-0.91	-0.95	-1.96	-1.31
$ln\, g_2$	6.12	6.32	6.00	6.30
$ln\, g_3$	9.8	9.8	9.8	9.8
$ln\, g_4$	12.9	12.9	12.9	12.9
$ln\, g_5$	0.463	0.463	0.463	0.463
$ln\, g_6$	0.68	0.71	1.47	0.98
$ln\, g_7$	4.67	4.24	5.09	4.24
$ln\, g_8$	1.94	1.08	0.73	1.53
$ln\, g_9$	1.88	1.96	4.04	2.69
$ln\, g_{10}$	1.40	1.36	0.80	0.44
$ln\, g_{11}$	1.64	1.49	1.79	1.49
$ln\, g_{12}$	1.13×10^{-9}	-1.86×10^{-9}	-1.01×10^{-9}	6.80×10^{-10}
$ln\, g_{13}$	-9.31×10^{-10}	1.19×10^{-9}	-2.07×10^{-9}	3.93×10^{-10}
$ln\, g_{14}$	1.21×10^{-9}	5.38×10^{-10}	-8.00×10^{-10}	2.43×10^{-9}
$ln\, g_{15}$	-1.21×10^{-9}	-1.32×10^{-9}	8.73×10^{-11}	-3.97×10^{-9}
$ln\, g_{16}$	4.95×10^{-10}	3.29×10^{-9}	-4.07×10^{-10}	3.01×10^{-9}
$ln\, g_{17}$	-8.29×10^{-10}	3.49×10^{-10}	-3.17×10^{-9}	3.20×10^{-10}
$ln\, g_{18}$	-1.14×10^{-9}	-1.57×10^{-9}	1.31×10^{-9}	7.57×10^{-10}
$ln\, g_{19}$	3.20×10^{-10}	1.43×10^{-9}	-9.90×10^{-10}	2.10×10^{-9}
$ln\, g_{20}$	-5.80	-8.07	-5.77	4.30
$ln\, g_{21}$	3.81	2.95	-1.88	2.45
$ln\, g_{22}$	-1.74×10^{-7}	-1.66×10^{-7}	-1.85×10^{-7}	-1.65×10^{-7}
$ln\, g_{23}$	-4.05	-4.11	-6.71	-5.06

Table 5-18: The estimated PF/GPR/Lognormal CPD parameters for the validation specimens

Parameter	**5A10**	**5A20**	**5A24**	**5A26**
ζ_0	0.224	0.224	0.224	0.224
ζ_1	2.41	2.39	1.91	2.20
$ln\, g_1$	-0.92	-0.94	-0.89	-1.66
$ln\, g_2$	5.44	9.76	9.67	7.55
$ln\, g_3$	10.8	10.8	10.8	10.8
$ln\, g_4$	12.1	12.7	13.4	13.2
$ln\, g_5$	0.155	0.154	0.154	0.155
$ln\, g_6$	2.11	2.54	1.83	2.50
$ln\, g_7$	5.08	4.60	5.62	4.60
$ln\, g_8$	1.19	1.19	1.19	1.19
$ln\, g_9$	1.19	1.21	1.15	2.14
$ln\, g_{10}$	0.98	0.98	0.98	0.98
$ln\, g_{11}$	2.29	0.33	2.25	1.24
$ln\, g_{12}$	-6.26×10^{-10}	-4.22×10^{-10}	-4.15×10^{-10}	-3.06×10^{-10}
$ln\, g_{13}$	-2.10×10^{-9}	-1.16×10^{-9}	2.33×10^{-10}	-2.10×10^{-9}
$ln\, g_{14}$	-3.92×10^{-10}	-1.04×10^{-10}	-1.70×10^{-10}	-5.84×10^{-10}
$ln\, g_{15}$	-2.76×10^{-10}	-3.35×10^{-10}	-6.55×10^{-11}	-3.13×10^{-10}
$ln\, g_{16}$	-5.97×10^{-10}	-2.84×10^{-10}	-5.75×10^{-10}	-5.09×10^{-10}
$ln\, g_{17}$	-1.16×10^{-10}	-3.49×10^{-10}	-2.91×10^{-10}	-1.31×10^{-10}
$ln\, g_{18}$	-1.80×10^{-9}	-2.16×10^{-9}	-1.64×10^{-9}	-1.80×10^{-9}
$ln\, g_{19}$	2.73×10^{-10}	4.37×10^{-10}	1.22×10^{-10}	3.44×10^{-10}
$ln\, g_{20}$	-9.17	-12.6	-3.86	-10.97
$ln\, g_{21}$	2.89	2.67	3.11	2.67
$ln\, g_{22}$	-0.249	-0.244	-0.269	-0.366
$ln\, g_{23}$	-4.66	-4.68	-4.83	-4.63

Table 5-19: The estimated PF/GPR/Weibull CPD parameters for the validation specimens

Parameter	5A10	5A20	5A24	5A26
α	1.55	1.53	1.58	0.90
β	0.403	0.388	0.406	0.391
$ln\, g_1$	-0.86	-0.86	-0.86	-0.86
$ln\, g_2$	7.23	7.23	7.23	7.23
$ln\, g_3$	9.8	13.4	12.4	8.0
$ln\, g_4$	13.2	13.2	12.6	13.0
$ln\, g_5$	0.453	0.562	0.380	0.553
$ln\, g_6$	1.50	1.86	1.26	1.83
$ln\, g_7$	5.09	5.43	4.86	5.53
$ln\, g_8$	0.62	0.75	0.53	0.80
$ln\, g_9$	2.40	2.40	2.40	2.40
$ln\, g_{10}$	1.46	1.55	1.41	1.55
$ln\, g_{11}$	2.21	2.21	2.18	3.73
$ln\, g_{12}$	-4.37×10^{-10}	-1.44×10^{-9}	1.30×10^{-9}	-1.31×10^{-9}
$ln\, g_{13}$	-1.75×10^{-10}	-1.60×10^{-9}	1.19×10^{-9}	-4.35×10^{-9}
$ln\, g_{14}$	-1.30×10^{-9}	1.25×10^{-9}	-3.12×10^{-9}	-4.07×10^{-10}
$ln\, g_{15}$	4.66×10^{-10}	1.86×10^{-9}	-6.64×10^{-9}	-1.43×10^{-9}
$ln\, g_{16}$	5.60×10^{-10}	-3.24×10^{-9}	2.52×10^{-9}	-2.07×10^{-9}
$ln\, g_{17}$	-6.04×10^{-10}	-2.91×10^{-11}	2.11×10^{-10}	3.93×10^{-10}
$ln\, g_{18}$	2.23×10^{-9}	-1.90×10^{-9}	3.75×10^{-9}	1.02×10^{-10}
$ln\, g_{19}$	-3.20×10^{-10}	2.02×10^{-9}	-3.06×10^{-10}	1.31×10^{-9}
$ln\, g_{20}$	-8.55	-10.62	-7.18	-10.44
$ln\, g_{21}$	1.49	1.49	1.49	1.49
$ln\, g_{22}$	-2.08	-1.45	-1.38	-1.74
$ln\, g_{23}$	-4.81	-4.61	-5.07	-4.60

5.3.2 Model Error and Validation Analysis

The validation analysis method described in Chapter 4 Section 4.2.2 is performed on one-hundred-forty-four validation points from the validation specimens listed in Section 5.2.1. Under the PF/GPR CP model, two-to-four-hundred validation points are extracted for the validation analysis. The results of the Bayesian estimations for μ_m, σ_m, and $E_{a\prime,t}$ of Equation (4.7) are presented in Table 5-20.

Table 5-20: Posterior mean and standard deviation for model log-logistic parameters and measurement error $\boldsymbol{E'_{a,t}}$

CPD Model	ME Parameter	PF/GPR		GPR	
		Mean	**SD**	**Mean**	**SD**
Logistic	μ_m	0.069	0.014	0.069	0.022
	σ_m	4.56×10^{-4}	4.58×10^{-4}	1.37×10^{-3}	1.38×10^{-3}
	$E_{a',t}$	1.071	0.027	1.072	0.046
Log-logistic	μ_m	0.067	0.019	0.069	0.022
	σ_m	7.83×10^{-4}	7.95×10^{-4}	1.35×10^{-3}	1.36×10^{-3}
	$E_{a',t}$	1.070	0.037	1.071	0.046
Lognormal	μ_m	0.069	0.019	0.068	0.023
	σ_m	8.54×10^{-4}	8.64×10^{-4}	1.94×10^{-3}	1.94×10^{-3}
	$E_{a',t}$	1.071	0.037	1.071	0.053
Weibull	μ_m	0.068	0.019	0.068	0.023
	σ_m	2.34×10^{-3}	2.29×10^{-4}	1.76×10^{-3}	1.78×10^{-3}
	$E_{a',t}$	1.072	0.056	1.072	0.051

These results then directly translate to the 95% confidence intervals for the overall model error between the different CPD models in addition to the median estimate presented in Table 5-21.

Table 5-21: The model error confidence bounds between the different CPD models

CPD Model	PF/GPR			GPR		
	2.50%	**50%**	**97.50%**	**2.50%**	**50%**	**97.50%**
Logistic	1.8%	7.11%	12.7%	1.8%	7.1%	17.1%
Log-logistic	0.2%	6.9%	14.7%	1.9%	7.0%	17.0%
Lognormal	0.3%	7.10%	14.9%	3.5%	7.0%	18.5%
Weibull	3.7%	7.11%	19.2%	2.8%	7.0%	18.1%

The primary findings from this part of the results is in improvement in model error from previous findings in this research (Smith & Modarres, 2017). The same validation procedure performed on specimen Sets 1 and 2 produced an average relative model error of 4% and a model error precision between 55% and 57% under the GPR CP model (Smith & Modarres, 2017). This result was already documented as a significant advantage that the GPR CP model has over other models including the AE (77–85%) and log-linear (58–59%) CP models (Smith & Modarres, 2017; Keshtgar, 2013).

From the recent analysis, the average GPR relative model error increased to 7.0% which is still an acceptable number for good model error. The GPR model error precision however, showed drastic improvement to between 15.0% and 15.4%. For the PF/GPR CP model these values improved further to an average PF/GPR relative model error of 7.1% and a PF/GPR model error precision range between 10.9% and 15.4%.

These findings lend to the previous conclusion that the GPR CP model is among the most realistic representations of propagation. However, the findings from the newly tested PF/GPR CP model shows that the fourth CP model may be more realistic than even the GPR CP model. This is likely a result of the stated advantage (Chapter 2 Section 2.2.2.4) that the PF/GPR CP model has over the GPR CP model. The PF/GPR is tied to nine CSF data as well as two AE data resulting in a CP model that relates propagation to eleven input. This further validates the other finding from the previous study which is that model input directly correlates to model error precision and CP realism (Smith & Modarres, 2017). Additionally the PF/GPR CP model takes advantage of much more input/output data due to the abundance of AE data that is gathered for the second set of tests. Finally, the CPD model pair with the smallest model error precision range is the PF/GPR/logistic model pair whereas the pair with the smallest model error average is the PF/GPR/log-logistic pair. This finding is similar to a previous analysis of the first two specimen set data as a standalone set (the GPR/log-logistic was originally the smallest for model error precision range and average) (Smith & Modarres, 2017).

5.3.3 End-of-Life Analysis

The results from the previous two sections of Section 5.3 are used in the end-of-life analysis for this research where model correction is used by applying the model error correction to the estimated CPD models.

5.3.3.1 Methodology for End-of-Life Analysis

The "end-of-life" for this research is defined as the time when a specimen's crack reaches the critical length a_{cr}. For simplicity this is when critical plane stress intensity factor or fracture toughness K_{IC} is reached where,

$$K_{IC} \geq f(g)\sigma\sqrt{\pi a_{cr}} = f(g)\frac{S_{max}}{K_f}\sqrt{\pi a_{cr}} \tag{5.2}$$

K_f is the fatigue notch factor, S_{max} is the maximum testing load on the specimen (see Tables 3 and 4), and $f(g)$ is the geometry correction factor,

$$f(g) = 1.12 - 0.231\left(\frac{a+r}{D}\right) + 10.55\left(\frac{a+r}{D}\right)^2 - 21.72\left(\frac{a+r}{D}\right)^3 + 30.39\left(\frac{a+r}{D}\right)^4 \tag{5.3}$$

The eFatigue calculator (eFatigue LLC, 2017) is used to approximate the K_f value from the theoretical stress concentration factor K_t which is 2.61 for the validation specimens. The K_f value is calculated from,

$$K_f = 1 + \frac{K_t}{\left[1 + \frac{0.025}{r}\left(\frac{2{,}070\text{ MPa}}{S_u}\right)^{1.8}\right]} \tag{5.4}$$

where r is in millimeters and an Al 7075-T6 ultimate strength S_u of 560 MPa (MakeItFrom.com, 2016) is used resulting in a K_f of 2.2696. The fracture toughness K_{IC} is represented as a distribution $LOGN(3.18,0.063)$ MPa$\sqrt{\text{m}}$ because of the variability in source estimates for this value (Clinton Aluminum & Stainless Steel, 2014; CRP

MECCANICA S.r.l., 2017). A separate MATLAB routine is used to match a range of K_{IC} values to a range of computed stress intensity values which are functions of the estimated true CP curves. This is the procedure that generates the end-of-life distribution for the specimens.

5.3.3.2 End-of-Life Analysis Results

Figure 5-10 through Figure 5-13 present both the estimated true CP based on Table 5-20 and the estimated end-of-life distributions for the four validation specimens.

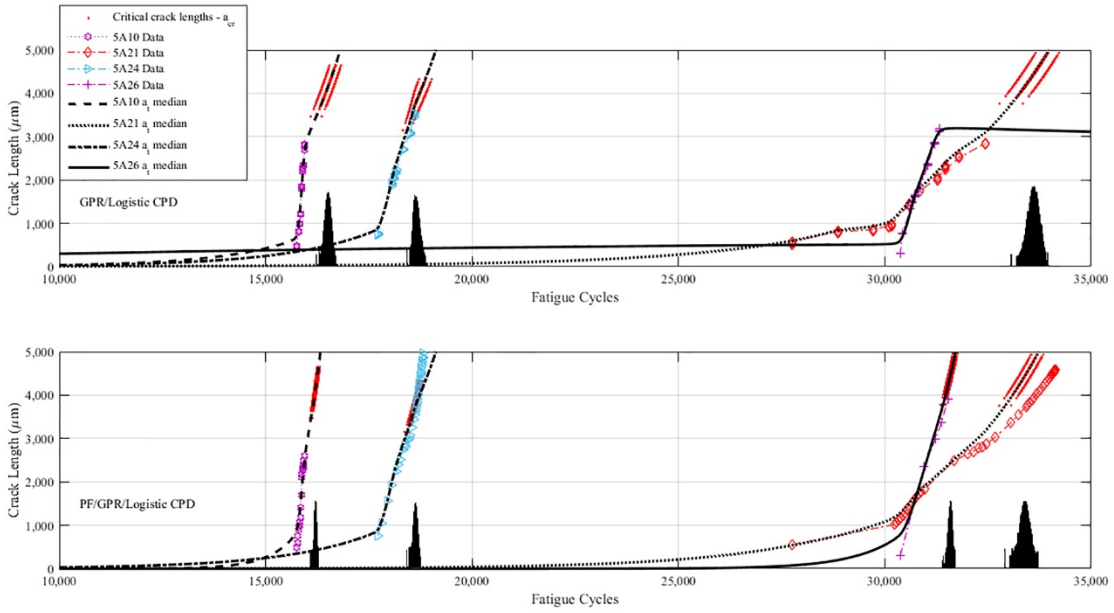

Figure 5-10: The validation median true crack length estimate and life distributions for logistic-based CPD models

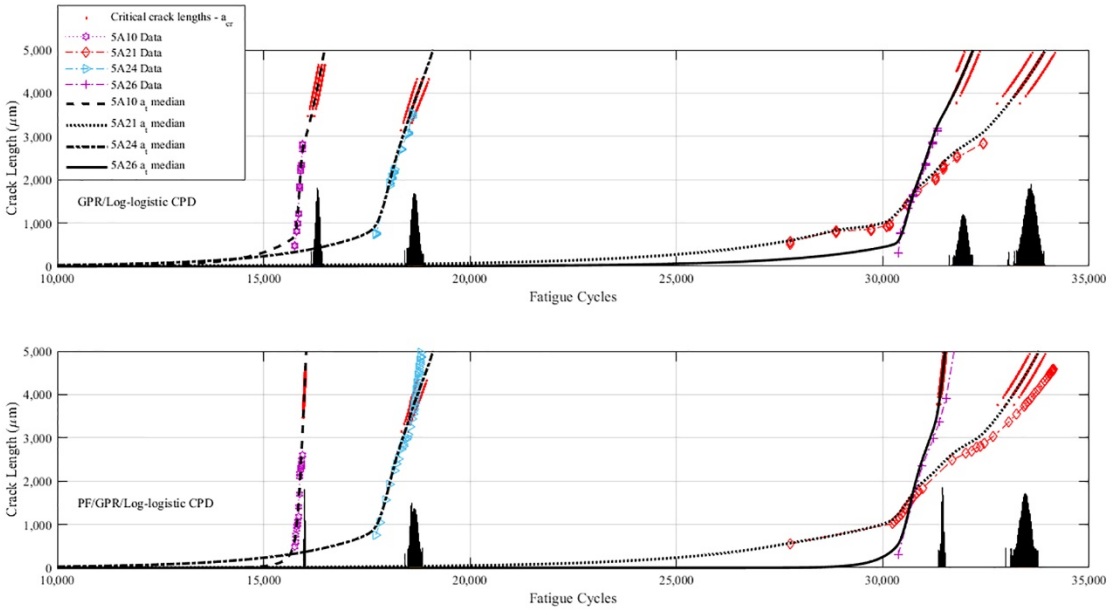

Figure 5-11: The validation median true crack length estimate and life distributions for log-logistic-based CPD models

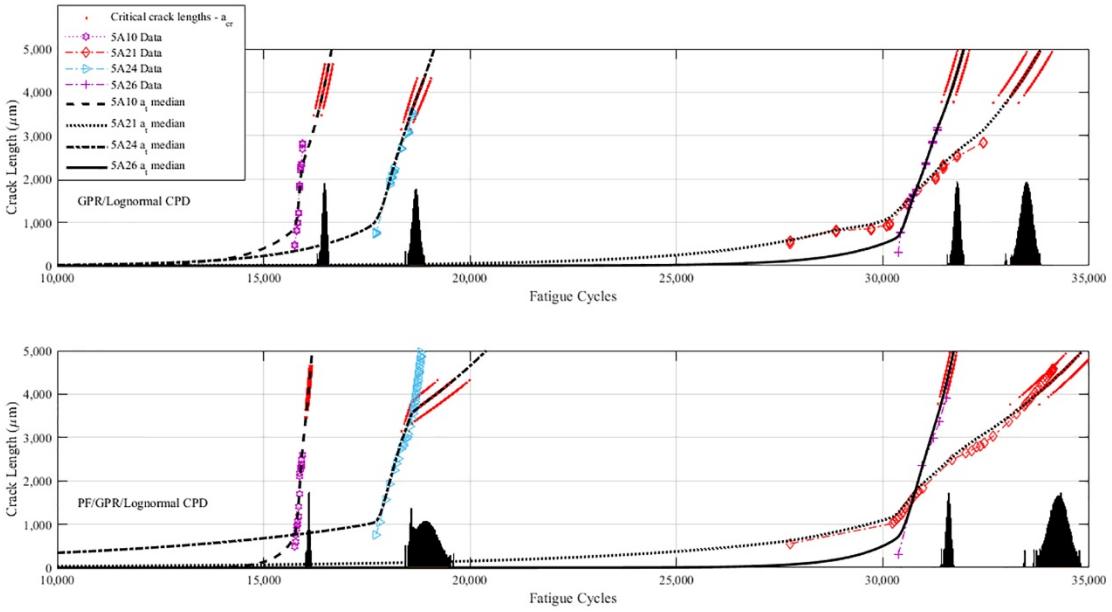

Figure 5-12: The validation median true crack length estimate and life distributions for lognormal-based CPD models

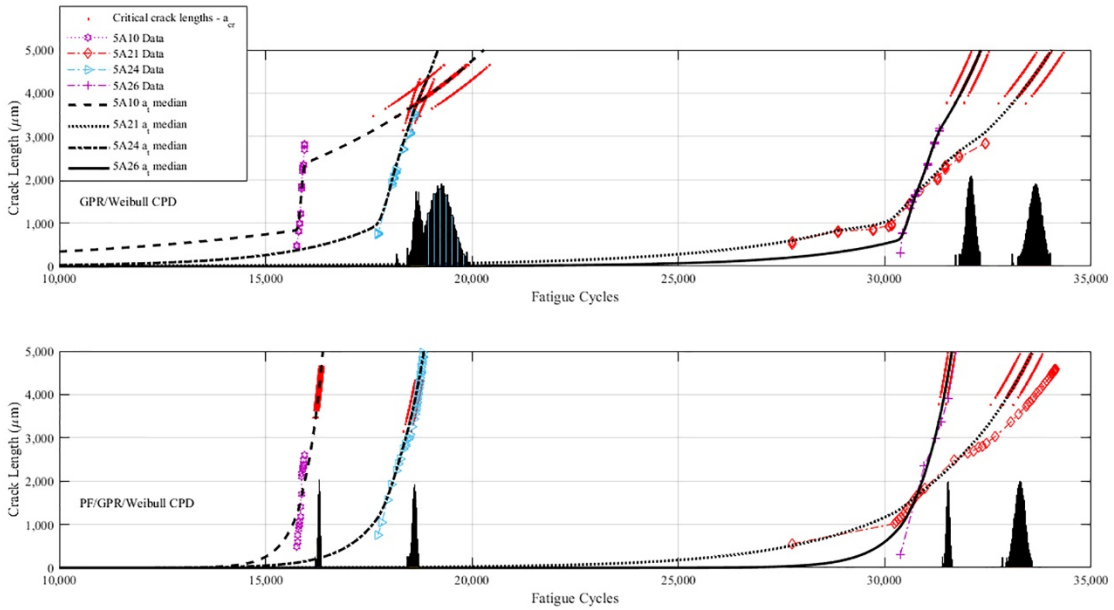

Figure 5-13: The validation median true crack length estimate and life distributions for Weibull-based CPD models

From the non-parametric distributions the mean-cycles-to-failure (MCTF) are extracted and presented in Table 5-22 and Table 5-23.

Table 5-22: The estimated CPD specific MCTFs in Fatigue Cycles (GPR based)

CPD Model	**MCTF (Fatigue Cycles)**			
	Specimen 5A10	**Specimen 5A21**	**Specimen 5A24**	**Specimen 5A26**
GPR/Logistic	1.65×10^4	3.36×10^4	1.87×10^4	N/A
GPR/Log-logistic	1.63×10^4	3.36×10^4	1.87×10^4	3.19×10^4
GPR/Lognormal	1.65×10^4	3.35×10^4	1.87×10^4	3.18×10^4
GPR/Weibull	1.92×10^4	3.36×10^4	1.87×10^4	3.21×10^4

Table 5-23: The estimated CPD specific MCTFs in Fatigue Cycles (PF/GPR based)

CPD Model	MCTF (Fatigue Cycles)			
	Specimen 5A10	**Specimen 5A21**	**Specimen 5A24**	**Specimen 5A26**
PF/GPR/Logistic	1.62×10^4	3.34×10^4	1.86×10^4	3.16×10^4
PF/GPR/Log-logistic	1.60×10^4	3.34×10^4	1.86×10^4	3.14×10^4
PF/GPR/Lognormal	1.61×10^4	3.42×10^4	1.89×10^4	3.16×10^4
PF/GPR/Weibull	1.63×10^4	3.33×10^4	1.86×10^4	3.15×10^4

As a point of comparison, the actual CTFs for specimens 5A10, 5A21, 5A24, and 5A26 are 1.59×10^4, 3.41×10^4, 1.95×10^4, and 3.24×10^4 fatigue cycles respectively. The one deviation that occurs in the predicted CP results as well as the MCTF is the GPR/logistic estimate for specimen 5A26 as seen in Figure 5-10 where the growth is not strictly increasing. This lends further credence to the finding that including the AE data as part of the GPR CP model contributes to the realism of the overall model. However, it is also likely a result of the PF/GPR model having more data to represent CP than the GPR model. The deviation in the case of the GPR/log-logistic estimate for specimen 5A26 may be a result of a shortage of CP data. In effect, the results of the PF/GPR CPD models further validate that in its addition of AE output. Where the GPR CP model is limited in the number of detected CP data, the PF/GPR CP model implements the abundant AE data in generating additional CP data for a more thorough and effective Bayesian analysis and a better CP estimate and end-of-life estimate. There are however, two notable drawbacks to this model as it applies to the CPD Bayesian methodology. The first is that the GPML routine (Rasmussen, Nickisch, & Williams, 2015) was limited on processing PF CP data for the Bayesian analysis, so there is an upper limit as to how many input/output data the code can process. Second, is that the PF/GPR CP model proved to be computationally expensive than the GPR model which itself is expensive.

Another finding based on this is that PF/GPR CPD models usually predict a more conservative estimate for MCTF than the GPR CPD models. The only times when the GPR CPD model is more conservative is for two specimens under the GPR/lognormal CPD model as seen by Table 5-22 and Table 5-23. These instances likely occur as a result of the correlated estimates for the CP curves. However, the CPD model is developed as one joint model. Therefore these instances may also be connected with the POD models. As seen in Figure 5-14, the correlated estimates for the posterior POD have the lognormal POD as the most liberal of the four POD models and both logistic and log-logistic exhibit the most conservative behavior.

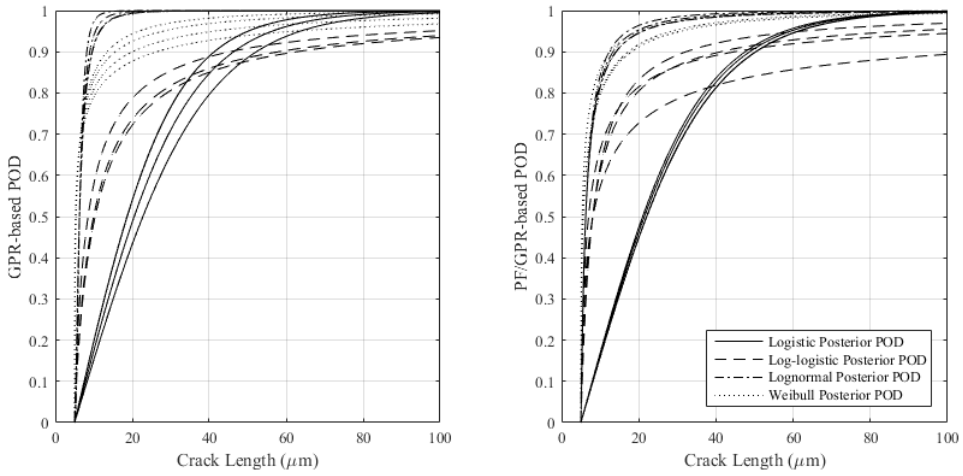

Figure 5-14: The validation median POD curves for the GPR and PF/GPR CPD models ($a_{lth} = 5\ \mu m$)

That is, it is more likely that cracks will be detected early in the crack's life-cycle, rather than more conservative POD models (logistic and log-logistic) where cracks are more likely to be detected later. With this, it is concluded that combination of relevant and high quantity CSFs and conservative POD model result in a conservative MCTF estimate.

5.4 Summary

In this chapter, the results from the Bayesian analysis of the CPD model and CSF correlations are discussed. The Bayesian parameter analysis of the CPD likelihood model has produced the updated posterior CPD models for the training specimens which are then used for designing the correlation between CPD models and CSFs. Through this correlation as well as model error correction, the CPD models and the RUL for validation specimens with their own unique CSFs are estimated. Although the procedure for the methodology was computationally expensive due to the complexity of GPR modeling, the end-of-life estimates for the validation specimens are very close to the actual end-of-life cycles.

Chapter 6: Conclusions, Contributions, and Future Work

6.1 Summary

The area of PHM has seen many advances in the study of CPD research. Among them is the correlation between AE signals to fatigue markers (Keshtgar, 2013), an enhanced correlation model that relates CP to a series of CSFs through a multivariate Gaussian relationship (Mohanty, Chattopadhyay, & Peralta, 2011), and the furtherance of identification of uncertainties in CPD modeling (Sankararaman, Ling, Shantz, & Mahadevan, 2009). The methodology described in this dissertation further validates the usefulness and effectiveness of a Bayesian analysis methodology performed on a joint-CPD model that is composed of a CP and a detection model.

For this research a set of CPD models were chosen and designed for the testing of the joint-CPD model. This included four POD models that are prevalent in the PHM field and two variations of the GPR CP model. The first variant was a direct correlation of CP to CSFs while the second variant implemented a PF technique that included AE indices as a means of modeling CP. Through PF generation of a CP-path based on AE data, the PF/GPR CP model successfully combined elements of the two CPD models tested in this research: the AE CP model and the GPR CP model. Fatigue test data gathered from previous research (Keshtgar, 2013; Smith & Modarres, 2017) provided the initial data for the methodology and modeling and this was later updated by new fatigue test data and AE data (Sauerbrunn, 2016; Smith, Modarres, & Droguett, 2017). This data was preconditioned through measurement error correction based on high-magnification analysis of the final crack and in-test documentation of CP.

The data was put through a rigorous Bayesian analysis methodology that was produced via a series of MATLAB routines developed and obtained for this research. The methodology updated a series of joint-CPD models and then the posterior models were put through a validation and RUL methodology that along with measurement error correction, successfully predicted the CP, POD, and RUL based on a CSF/CPD parameter correlation methodology. The effectiveness of this methodology is a result of the overabundance of data that was made available, especially for the PF/GPR based CPD models where data sets were built from over twenty-thousand AE data sets per specimen. While the increase in data contributed to a model with less uncertainties, it also results in a much more expensive Bayesian analysis through the data processing software used for the research. Despite this, the onset of this new methodology opens up the possibilities of combining this methodology with other Bayesian methodologies, such as the one proposed by Rabiei et. al. (Rabiei, Droguett, & Modarres, 2016), in order to produce a more effective Bayesian analysis of RUL estimation.

6.2 Principal Contributions

During the course of this research, a number of significant findings were discovered thus the list of contributions to CPD modeling, RUL estimation, and the PHM field is presented as follows:

1. A new approach was proposed which groups models (crack propagation and detection) into a single integrated model for a singular Bayesian analysis.
2. A set of correlations was developed between CPD model uncertainty and a concept introduced in this research as CSFs that can be used to assess CP. These correlations can be used to predict the CP, POD, RUL of specimens with a unique CSF set.

3. The effect of the POD of sensor data on the CP mode was considered, resulting in the need for a joint-CPD model.
4. It was discovered that with regard to a validation methodology used by Ontiveros et al. (Ontiveros, Cartillier, & Modarres, 2010), the modeling of measurement error is not specifically modeled as a lognormal distribution at all times. For this research, it was found that the log-logistic distribution is a better model for measurement error, but this proves that the distribution used for measurement error is not restricted to one distribution.
5. An improved application of the GPR CP model was designed in which a path-wise CP model captures the true crack path and fits it to a large set of CSFs.
6. A new CP model was developed that combines elements of GPR CP modeling based on CSF-to-CP relation and PF techniques in which AE indices were utilized for the modeling.

6.3 Recommendation for Future Work

The following are some recommendations that can improve on this methodology:

- As mentioned in Chapter 3 Section 3.4.4, the loading condition CSFs were treated as constant values because the variation was considered to be minimal. However, the material CSFs were treated as variable because of the variety of grain and inclusion sizes on each specimen. That said, for future application of this methodology there should be a consideration for the variation of loading CSFs in a similar fashion as the material grain and inclusion diameters in this study.

- One of the drawbacks to using big data is that the methodology is considerably expensive computationally. Likely this is a result of the data set size as well as the processing speed of the MATLAB software used for development of the code. There are two recommendations based on this:
    - The MATLAB routine should be reexamined and revised in an effort to minimize the processing time and make it less expensive.
    - The methodology should be made available in additional programming languages such as OpenBUGS, R, and C++ as an alternative to the MATLAB routine.
- The demonstration of the methodology in this dissertation only considered nine CSFs in the GPR modeling. For future studies, additional CSFs should be checked to see their effect on the CPD model uncertainty. This should be followed up by a sensitivity analysis of all CSFs in order to determine which CSFs have more of an effect on CP and POD.
- The demonstration of the methodology used GPR based CP models to demonstrate its usefulness; but GPR has a known issue of overfitting the input and output data that it is correlating. Therefore it is suggested that an alternative model, Artificial Neural Network (ANN) (McCulloch & Pitts, 1943; Kleene, 1956), be explored as it also serves as an effective machine learning model.

Appendix A: Fatigue Crack Data

This appendix lists only the fatigue crack lengths at the cycles where they were detected. Cycles prior to this did not exhibit any crack detections.

Appendix A.1 Specimen DB3

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
5307	41	**6090**	270
5346	75	**6117**	283
5580	89	**6126**	314
5670	99	**6135**	324
5709	113	**6144**	358
5736	117	**6153**	391
5763	122	**6162**	404
5772	124	**6171**	422
5781	141	**6180**	425
5790	155	**6189**	538
5799	155	**6198**	557
5808	184	**6207**	594
5988	189	**6243**	632
5997	212	**6252**	656
6015	228	**6279**	743
6036	243	**6306**	972
6063	256	**6315**	1019
6072	261		

Appendix A.2 Specimen DB4

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
4527	447	5238	875
5169	656	5250	901
5172	674	5265	903
5184	676	5280	922
5187	682	5283	954
5196	686	5295	978
5208	687	5322	1003
5211	702	5343	1023
5223	737	5346	1070
5226	809	5349	1167

Appendix A.3 Specimen DB5

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
3816	14	4082	415
3830	18	4084	423
3844	26	4096	471
3890	57	4106	473
3914	74	4124	504
3922	109	4126	518
3924	130	4148	606
3934	145	4154	608
3936	185	4166	628
3938	193	4178	630
3952	226	4180	631
3956	235	4186	639
3970	243	4190	652
3972	246	4192	680
3982	248	4194	682
3998	269	4212	794
4002	279	4220	798
4012	289	4222	799
4018	308	4228	820
4026	320	4230	825
4030	325	4242	835
4032	336	4244	853
4044	346	4252	867
4054	378	4260	914
4056	380	4262	928
4064	399	4264	942
4078	414		

Appendix A.4 Specimen DB6

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
2019	24	2583	149
2136	51	2790	378
2343	69	2982	434
2370	83	3090	631
2553	106		

Appendix A.5 Specimen DB7

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
5950	5.1	7232	122.1	7998	238.1
5984	6.4	7258	124.4	8014	238.1
6024	9.4	7264	124.6	8028	251.3
6040	14.3	7288	131.2	8036	273.4
6056	15.8	7298	140.4	8062	278.5
6078	16.9	7304	142.8	8070	285.5
6116	17.6	7312	144.0	8108	307.4
6148	18.9	7320	150.8	8120	315.7
6176	21.0	7328	152.9	8136	315.9
6200	22.7	7344	159.1	8144	322.9
6226	22.9	7350	168.5	8152	323.9
6256	22.9	7360	169.0	8160	330.2
6370	23.5	7368	169.7	8168	337.2
6826	23.9	7376	170.2	8178	338.8
6834	25.2	7392	171.2	8196	339.3
6846	25.9	7400	173.7	8202	344.7
6854	26.7	7408	174.7	8216	369.3
6860	26.9	7414	176.1	8224	369.6
6876	27.2	7422	176.5	8248	370.8
6892	29.7	7430	178.3	8260	391.3
6900	35.1	7454	183.7	8268	399.0
6908	36.5	7462	183.8	8286	400.9
6936	37.9	7500	186.8	8294	408.5
6944	38.4	7508	194.9	8300	409.7
6956	39.9	7524	200.4	8308	423.4
6964	40.9	7540	208.3	8314	427.6
6970	53.3	7580	209.3	8318	433.7

7010	56.5	7588	212.8	8324	441.3
7036	58.4	7600	214.1	8328	443.4
7042	59.3	7746	214.3	8338	443.6
7044	68.4	7794	214.4	8344	448.9
7048	76.0	7802	215.3	8354	452.7
7064	78.4	7810	215.7	8362	469.8
7104	83.1	7834	217.8	8372	478.2
7120	97.9	7850	218.4	8378	482.2
7164	98.2	7858	222.6	8388	486.2
7170	112.5	7874	224.2	8398	491.7
7204	117.1	7894	225.9	8406	505.6
7212	117.7	7976	228.0	8416	513.4
7220	118.2	7990	236.0	8486	550.1

Appendix A.6 Specimen DB15

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
7770	36.4	7998	347
7818	36.8	8070	364
7824	56	8136	431
7892	167	8138	466
7906	180	8152	518
7918	214	8208	527
7952	228	8224	548
7954	238	8250	568
7978	283		

Appendix A.7 Specimen 1A2

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
8320	100	8920	651
8370	159	8970	733
8420	180	9020	734
8470	214	9070	755
8520	218	9120	764
8570	229	9170	790
8620	295	9220	963
8670	334	9270	972
8720	359	9320	1160
8770	407	9370	1306
8820	589	9420	1362
8870	611	9470	1370

Appendix A.8 Specimen 1B3

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
4120	16	5320	585
4170	27	5370	616
4220	31	5420	624
4270	39	5470	630
4320	49	5520	642
4370	61	5570	674
4420	87	5620	683
4470	102	5670	710
4520	124	5720	715
4570	142	5770	741
4620	175	5820	753
4670	213	5870	768
4720	337	5920	794
4770	351	5970	927
4820	411	6020	994
4870	427	6070	1006
4920	481	6120	1013
4970	490	6170	1042
5020	514	6220	1044
5070	522	6270	1051
5120	531	6320	1090
5170	536	6370	1169
5220	544	6420	1196
5270	577	6470	1354

Appendix A.9 Specimen 5A2

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
9645	181	9995	761
9645	173	9995	760
9645	146	10120	1084
9645	162	10120	1061
9770	211	10120	1071
9770	198	10120	1010
9770	201	10245	1385
9770	195	10245	1402
9895	319	10245	1388
9895	298	10245	1385
9895	301	10350	1558
9895	359	10350	1516
9995	758	10350	1502
9995	749	10350	1501

Appendix A.10 Specimen 5A3

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
11120	164	11495	1374
11120	155	11495	1425
11120	170	11495	1418
11120	148	11495	1445
11195	385	11520	1455
11195	355	11520	1471
11195	409	11520	1423
11195	411	11520	1390
11220	569	11550	1521
11220	537	11550	1519
11220	543	11550	1515
11220	523	11550	1518
11245	665	11575	1598
11245	618	11575	1580
11245	612	11575	1540
11245	590	11575	1556
11270	739	11600	1644
11270	677	11600	1675
11270	715	11600	1678
11270	694	11600	1661
11320	1001	11630	1751
11320	1009	11630	1740
11320	1032	11630	1749
11320	966	11630	1751
11370	1330	11705	1881
11370	1323	11705	1887
11370	1316	11705	1842
11370	1326	11705	1834
11395	1354	11730	1912
11395	1354	11730	1878
11395	1347	11730	1941
11395	1356	11730	1955
11445	1348	11755	2163
11445	1379	11755	2131
11445	1377	11755	2248
11445	1379	11755	2273

Appendix A.11 Specimen 5A4

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
11730	231	11835	1165
11730	198	11835	1172
11730	213	11885	1374
11730	220	11885	1366
11755	394	11885	1383
11755	403	11885	1396
11755	387	11910	1563
11755	367	11910	1600
11780	564	11910	1607
11780	527	11910	1605
11780	516	11935	1832
11780	490	11935	1794
11805	825	11935	1813
11805	788	11935	1797
11805	743	11960	1930
11805	759	11960	1919
11835	1196	11960	1914
11835	1159	11960	1937

Appendix A.12 Specimen 5A6

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
17460	158	19965	688
17460	158	19965	714
17460	152	20215	756
17460	159	20215	788
18285	279	20215	773
18285	283	20215	770
18285	309	20470	801
18285	315	20470	800
18940	517	20470	808
18940	529	20470	818
18940	528	21145	1270
18940	573	21145	1251
19240	550	21145	1255
19240	576	21145	1246
19240	571	21520	1643
19240	569	21520	1639
19965	687	21520	1634
19965	694	21520	1642

Appendix A.13 Specimen 5A8

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
23510	157	39705	1239
23510	156	39705	1238
23510	143	39705	1236
23510	141	39705	1249
23760	181	40310	1255
23760	164	40310	1269
23760	172	40310	1261
23760	182	40310	1285
24790	204	40885	1444
24790	201	40885	1447
24790	197	40885	1434
24790	189	40885	1440
37445	438	41290	1533
37445	433	41290	1535
37445	447	41290	1518
37445	438	41290	1526
37470	530	41540	1919
37470	479	41540	1936
37470	525	41540	1950
37470	517	41540	1971
37520	616	42065	2192
37520	629	42065	2206
37520	634	42065	2239
37520	610	42065	2224
37620	635	42245	2234
37620	649	42245	2223
37620	631	42245	2231
37620	643	42245	2274
38400	782	42820	2506
38400	749	42820	2495
38400	770	42820	2460
38400	792	42820	2443
38905	994	43170	2752
38905	1016	43170	2566
38905	996	43170	2578
38905	1021	43170	2557
39380	1067	43020	2738
39380	1082	43020	2766
39380	1058	43020	2685

39380	1083	43020	2676
39405	1109	43375	3029
39405	1137	43375	2999
39405	1138	43375	2975
39405	1110	43375	3018

Appendix A.14 Specimen 5A9

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
35740	397	36040	1652
35740	383	36040	1634
35740	384	36065	2108
35740	409	36065	2078
35965	464	36065	2087
35965	476	36065	2079
35965	452	36090	2360
35965	456	36090	2411
35990	981	36090	2377
35990	967	36090	2397
35990	979	36115	2592
35990	979	36115	2552
36015	1413	36115	2563
36015	1457	36115	2577
36015	1468	36140	2821
36015	1477	36140	2836
36040	1638	36140	2844
36040	1652	36140	2869

Appendix A.15 Specimen 5A10

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
15765	380	15865	1423
15765	358	15865	1410
15765	381	15865	1379
15765	353	15865	1368
15790	626	15890	1689
15790	608	15890	1725
15790	634	15890	1720
15790	640	15890	1760
15815	756	15915	1805
15815	765	15915	1802
15815	764	15915	1797
15815	751	15915	1799
15840	919	15940	2166
15840	935	15940	2149
15840	936	15940	2167
15840	933	15940	2074

Appendix A.16 Specimen 5A20

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
6400	130	6700	684
6400	123	6700	682
6400	120	6700	677
6400	126	6700	679
6425	149	6750	945
6425	142	6750	945
6425	142	6750	938
6425	146	6750	939
6500	166	6825	1296
6500	169	6825	1281
6500	165	6825	1271
6500	166	6825	1287
6525	212	6875	1616
6525	210	6875	1610
6525	215	6875	1610
6525	203	6875	1609
6550	241	6900	1641
6550	249	6900	1642
6550	247	6900	1656
6550	251	6900	1669
6600	384	7000	2131
6600	374	7000	2167
6600	374	7000	2103
6600	380	7000	2074
6625	469	7025	2354
6625	479	7025	2363
6625	466	7025	2322
6625	463	7025	2342

Appendix A.17 Specimen 5A21

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
27760	393	30595	1104
27760	429	30595	1075
27760	440	30845	1346
27760	443	30845	1346
28865	630	30845	1336
28865	598	30845	1334
28865	601	31270	1579
28865	618	31270	1540
29720	659	31270	1542
29720	637	31270	1537
29720	655	31470	1784
29720	666	31470	1777
30095	719	31470	1762
30095	700	31470	1725
30095	705	31795	1927
30095	721	31795	1936
30170	751	31795	1973
30170	734	31795	1926
30170	726	32425	2188
30170	758	32425	2171
30595	1112	32425	2166
30595	1093	32425	2179

Appendix A.18 Specimen 5A22

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
15700	145	16130	2053
15700	125	16130	2087
15700	134	16130	2078
15700	147	16130	2085
15880	979	16155	2248
15880	997	16155	2224
15880	1003	16155	2219
15880	1022	16155	2199
15980	1417	16180	2349
15980	1400	16180	2329
15980	1405	16180	2305
15980	1397	16180	2316
16005	1575	16230	2571
16005	1592	16230	2536
16005	1619	16230	2526
16005	1617	16230	2539
16105	1988	16255	2631
16105	1999	16255	2689
16105	2008	16255	2685
16105	2007	16255	2708

Appendix A.19 Specimen 5A23

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
36120	602	39185	1146
36120	582	39185	1145
36120	603	40410	1821
36120	579	40410	1791
37230	652	40410	1785
37230	646	40410	1806
37230	652	40885	2223
37230	659	40885	2264
38485	1076	40885	2236
38485	1094	40885	2247
38485	1072	41390	2496
38485	1088	41390	2463
39185	1221	41390	2457
39185	1199	41390	2472

Appendix A.20 Specimen 5A24

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
17720	572	18170	1726
17720	587	18170	1728
17720	599	18170	1727
17720	560	18170	1728
18045	1437	18325	2081
18045	1460	18325	2067
18045	1429	18325	2080
18045	1481	18325	2088
18095	1560	18500	2365
18095	1580	18500	2346
18095	1558	18500	2361
18095	1554	18500	2379
18145	1692	18600	2714
18145	1670	18600	2683
18145	1674	18600	2646
18145	1657	18600	2724

Appendix A.21 Specimen 5A26

Fatigue Cycles	Detected Crack Length (µm)	Fatigue Cycles	Detected Crack Length (µm)
6075	239	6145	1259
6075	236	6145	1246
6075	228	6160	1357
6075	226	6160	1354
6085	575	6160	1324
6085	574	6160	1325
6085	586	6210	1801
6085	595	6210	1793
6125	1031	6210	1819
6125	1025	6210	1817
6125	1031	6240	2200
6125	1023	6240	2173
6135	1159	6240	2190
6135	1139	6240	2195
6135	1132	6265	2409
6135	1149	6265	2392
6145	1261	6265	2456
6145	1264	6265	2418

Appendix B: MATLAB Codes

Appendix B.1 Full CPD Bayesian Analysis Code for GPR Analysis

The code provided is the Bayesian analysis for the GPR/Log-logistic CPD model where the specimen of interest is specimen 5A2 (Sauerbrunn, 2016). To use the routine, it is necessary to download the GPML MATLAB package (Rasmussen, Nickisch, & Williams, 2015) and link it to the run file on line 18. The kernel function also has to be defined as both an outside-routine function (Appendix B.2) and as an in-routine function. The load command loads files that will have to be generated by the user or requested of the author.

```
% ====================================================================
% Full POD Likelihood
% POD Processing Method - Hit-or-Miss
% POD CDF Definition - Log-logistic
% Crack growth PDF Definition - MVNormal Kernel 4 (K4)
% ====================================================================
% by Reuel Smith
% ====================================================================
% 23 GPR parameters plus 2 logistic POD parameters
% Parameter Definitions
% x(1) - Logistic POD alpha1 parameter
% x(2) - Logistic POD alpha2 parameter
% x(3)-x(25) - the log of the GPR parameters
% x(26) - the false call or false positive probability parameter
% (D=1|a<alow)
clear
clc
run('C:\Users\ReuelS\Desktop\POD Research\gpml-matlab-v3.6\startup.m')

% ====================================================
% Signal Response (Total Data) Based POD Parameters
% ====================================================
load('DBdata','Data_5A2','init_a')
load('PriordistGPR','ParamsetGPRloglogistic')
load('measurementerror','ME')

%% Constants
% ====================================================
significance=0.05;
alow = 5;           % lower threshold of crack detection (micrometers)
Ea = ME(1);             % 50% measurement error

nsamples=20000;     %Number of samples (higher number of samples
increases the acceptance rate)
```

```
K = 1000;           %this value can remain as it is even if model
parameters or model change
M = 10;            % This parameter controls the size of the new markov
chain which omits M-1 out of M values.  This will curb the effect of
autocorrelation
n=26;                % Number of parameters
priors = ParamsetGPRloglogistic;
% =====================================================
% Kernel 4 covariance function determination
% =====================================================
D = 9;
covfuncPOST =
{@covSum,{@covConst,@covLINard,@covSEard,@covNNone,@covNoise}};
% mean function determination
meanfuncPOST = @meanZero;
% likelihood function determination
likfuncPOST = @likGauss;
% intial guess of hyperparameters of likelihood function
hypPOST.lik = log(0.1);

% =====================================================
% Gather the inputs (X) and the outputs (Y) for DB 5A2
% =====================================================
% X and Y will be taken from Sample 5A2
Xfull = Data_5A2(:,3:11);       % All Cycle data
Yfull = Data_5A2(:,1);          % All Crack Length data (um)

% Detected Data D = 1
X_5A2 = Xfull(20:39,:);
Y_5A2 = Yfull(20:39)./Ea;
Z_5A2 = X_5A2;
Z_5A2b = [linspace(1,11000,1001)',ones(1001,1)*Xfull(1,2:end)];
XtrainD1 = X_5A2;
YtrainD1 = Y_5A2;
XnewD1 = Xfull(40:end,:);
YnewD1 = Yfull(40:end)./Ea;
XnewD1test = XnewD1;
YnewD1test = YnewD1;

% Undetected Data D = 0
XnewD0 = Xfull(1:19,:);

% Data input for prior
Znew = XnewD1;
Znewb = [linspace(1,11000,1001)',ones(1001,1)*Xfull(1,2:end)];
%% GPML Analysis
% =====================================================
[DB5A2nlml, DB5A2result, DB5A2newm, DB5A2news2,DB5A2K,DB5A2f]=
kernel7(XtrainD1,YtrainD1,Z_5A2,significance);
init_param=[2 -0.2 0.2 10 5.5 6 0.2 0.2 3.4 -2.5 3.5 -2 5.95 0 0 0 0 0
0 0 0 0.65 1.5 1.35 -2.75 0.2];   % initial parameter guess
init_param(3:25) = log(DB5A2result)';
% ================================
% NEW DB DATA
% ================================
% This establishes our prior data for the crack growth curve
```

```
[DBnew_nlml, DBnew_result, DBnew_newm, DBnew_news2,DBnew_K,DBnew_f]=
kernel7(XtrainD1,YtrainD1,Znew,significance);
[DBnew_nlmlb, DBnew_resultb, DBnew_newmb,
DBnew_news2b,DBnew_Kb,DBnew_fb]=
kernel7(XtrainD1,YtrainD1,Znewb,significance);

% convert to lognormal parameters
mu_y = log(DBnew_newm./sqrt(1+DBnew_news2./(DBnew_newm.^2)));
SD_y = sqrt(log(1+DBnew_news2./(DBnew_newm.^2)));
%% MHSAMPLING AND LIKELIHOOD SETUP
% =====================================================
prop_sig=eye(n);            % sigma for the proposed PDF
% The log-logistic POD CDF function
loglogisticpodcdf=@(x2,mu,logsig) 1./(1+exp(-((log(x2-alow)-
mu)./exp(logsig))));

% Kernel 4 calculation
KN = @(x,XN) covConst(x(3),XN)+...
    covLINard([x(4) x(5) x(6) x(7) x(8) x(9) x(10) x(11)
x(12)]',XN)+...
    covSEard([x(13) x(14) x(15) x(16) x(17) x(18) x(19) x(20) x(21)
x(22)]',XN)+...
    covNNone([x(23) x(24)]',XN)+covNoise(x(25), XN);
kNP1 = @(x,XN,XNP1) covConst(x(3),XN,XNP1)+...
    covLINard([x(4) x(5) x(6) x(7) x(8) x(9) x(10) x(11)
x(12)]',XN,XNP1)+...
    covSEard([x(13) x(14) x(15) x(16) x(17) x(18) x(19) x(20) x(21)
x(22)]',XN,XNP1)+...
    covNNone([x(23) x(24)]',XN,XNP1)+covNoise(x(25), XN,XNP1);

% Mean and Variance of new test data
MU_K = @(x,XN,YN,XNP1) kNP1(x,XN,XNP1)'*(KN(x,XN)\log(YN));
S2_K = @(x,XN,XNP1) diag(KN(x,XNP1)-
kNP1(x,XN,XNP1)'*(KN(x,XN)\kNP1(x,XN,XNP1)));

pdfD1 = @(x,XN,YN,XNP1,YNP1)(1./(1+exp(-((log(YNP1-alow)-
x(1))./exp(x(2)))))).*1./(YNP1.*sqrt(2.*pi.*S2_K(x,XN,XNP1))).*exp(-
((log(YNP1)-MU_K(x,XN,YN,XNP1)).^2)./(2.*S2_K(x,XN,XNP1)));
pdfD0b = @(x,MU_K1,S2_K1)(1-quadgk(@(a)(1-x(26)).*(1./(1+exp(-((log(a-
alow)-x(1))./exp(x(2)))))).*1./(a.*sqrt(2.*pi.*S2_K1)).*exp(-((log(a)-
MU_K1).^2)./(2.*S2_K1)),alow,Inf));
pdfD0 = @(x,XN,YN,XNP1) arrayfun(@(q1,q2)
pdfD0b(x,q1,q2),MU_K(x,XN,YN,XNP1),S2_K(x,XN,XNP1));

PODpriors = @(x) pdf(priors(1,1),x(1))*pdf(priors(2,1),x(2));
GPRpriors = @(x)
pdf(priors(3,1),x(3))*pdf(priors(4,1),x(4))*pdf(priors(5,1),x(5))*...

pdf(priors(6,1),x(6))*pdf(priors(7,1),x(7))*pdf(priors(8,1),x(8))*...

pdf(priors(9,1),x(9))*pdf(priors(10,1),x(10))*pdf(priors(11,1),x(11))*.
..

pdf(priors(12,1),x(12))*pdf(priors(13,1),x(13))*pdf(priors(14,1),x(14))
*...
```

```
pdf(priors(15,1),x(15))*pdf(priors(16,1),x(16))*pdf(priors(17,1),x(17))
*...

pdf(priors(18,1),x(18))*pdf(priors(19,1),x(19))*pdf(priors(20,1),x(20))
*...

pdf(priors(21,1),x(21))*pdf(priors(22,1),x(22))*pdf(priors(23,1),x(23))
*...
    pdf(priors(24,1),x(24))*pdf(priors(25,1),x(25));
Pfcprior = @(x) pdf(priors(26,1),x(26));
pdf = @(x)
PODpriors(x)*GPRpriors(x)*Pfcprior(x)*prod(pdfD1(x,XtrainD1,YtrainD1,Xn
ewD1,YnewD1))*prod(pdfD0(x,XtrainD1,YtrainD1,XnewD0));

% define proposal distribution and r.n. generator
proppdf = @(x,y) mvnpdf(x,y,prop_sig);
proprnd = @(y) [normrnd(y(1),0.01),normrnd(y(2),0.01),...
    normrnd(y(3),0.1),normrnd(y(4),0.1),normrnd(y(5),0.1),...
    normrnd(y(6),0.1),normrnd(y(7),0.1),normrnd(y(8),0.1),...
    normrnd(y(9),0.1),normrnd(y(10),0.1),normrnd(y(11),0.1),...
    normrnd(y(12),0.1),normrnd(y(13),0.1),normrnd(y(14),1e-11),...
    normrnd(y(15),1e-11),normrnd(y(16),1e-11),normrnd(y(17),1e-11),...
    normrnd(y(18),1e-11),normrnd(y(19),1e-11),normrnd(y(20),1e-11),...
    normrnd(y(21),1e-11),normrnd(y(22),0.001),normrnd(y(23),0.1),...
    normrnd(y(24),0.1),normrnd(y(25),0.001),normrnd(y(26),0.1)];

 [result,accept] =
mhsample(init_param,nsamples,'pdf',pdf,'proppdf',proppdf,'proprnd',prop
rnd,'burnin',K,'thin',M);

% ============================================================
% Calculation of the autocorrelation values for lagged values
% ============================================================
AC = zeros(n,1);
lag = 1;
for i=1:n
    me = mean(result(:,i));
    v = var(result(:,i));
    m2 = result(:,i)-me;
    ACfactor = zeros(nsamples-lag,1);
    for j=1:(nsamples-lag)
        ACfactor(j) = m2(j)*m2(j+lag);
    end
    AC(i) = 1/(v*(nsamples-lag))*sum(ACfactor);
end
%% POSTERIOR PARAMETERS AND CALCULATIONS
% =======================================
[F1,x1] = ecdf(result(:,1));[F2,x2] = ecdf(result(:,2));
[F3,x3] = ecdf(result(:,3));[F4,x4] = ecdf(result(:,4));
[F5,x5] = ecdf(result(:,5));[F6,x6] = ecdf(result(:,6));
[F7,x7] = ecdf(result(:,7));[F8,x8] = ecdf(result(:,8));
[F9,x9] = ecdf(result(:,9));[F10,x10] = ecdf(result(:,10));
[F11,x11] = ecdf(result(:,11));[F12,x12] = ecdf(result(:,12));
[F13,x13] = ecdf(result(:,13));[F14,x14] = ecdf(result(:,14));
[F15,x15] = ecdf(result(:,15));[F16,x16] = ecdf(result(:,16));
```

```
[F17,x17] = ecdf(result(:,17));[F18,x18] = ecdf(result(:,18));
[F19,x19] = ecdf(result(:,19));[F20,x20] = ecdf(result(:,20));
[F21,x21] = ecdf(result(:,21));[F22,x22] = ecdf(result(:,22));
[F23,x23] = ecdf(result(:,23));[F24,x24] = ecdf(result(:,24));
[F25,x25] = ecdf(result(:,25));[F26,x26] = ecdf(result(:,26));
pd1 = fitdist(x1,'Lognormal');pd2 = fitdist(x2,'ExtremeValue');
pd3 = fitdist(x3,'ExtremeValue');pd4 = fitdist(x4,'Loglogistic');
pd5 = fitdist(x5,'ExtremeValue');pd6 = fitdist(x6,'Gamma');
pd7 = fitdist(x7,'Normal');pd8 = fitdist(x8,'Logistic');
pd9 = fitdist(x9,'Gamma');pd10 = fitdist(x10,'Logistic');
pd11 = fitdist(x11,'Logistic');pd12 = fitdist(x12,'Logistic');
pd13 = fitdist(x13,'ExtremeValue');pd14 = fitdist(x14,'ExtremeValue');
pd15 = fitdist(x15,'ExtremeValue');pd16 = fitdist(x16,'ExtremeValue');
pd17 = fitdist(x17,'ExtremeValue');pd18 = fitdist(x18,'ExtremeValue');
pd19 = fitdist(x19,'ExtremeValue');pd20 = fitdist(x20,'ExtremeValue');
pd21 = fitdist(x21,'ExtremeValue');pd22 = fitdist(x22,'ExtremeValue');
pd23 = fitdist(x23,'ExtremeValue');pd24 = fitdist(x24,'Normal');
pd25 = fitdist(x25,'ExtremeValue');pd26 = fitdist(x26,'Beta');
% POD POSTERIOR PARAMETERS
[b1muPOST,b1SDPOST] = normfit(result(:,1));
[b2muPOST,b2SDPOST] = normfit(result(:,2));
PODMCerror = [b1SDPOST;b2SDPOST]./sqrt(nsamples);

% GPR POSTERIOR PARAMETERS
for i=1:n-3
    GPRPOST(i,1)=mean(result(:,i+2));
    GPRPOST(i,2)=std(result(:,i+2));
    GPRMCerror(i,1) = std(result(:,i+2))/sqrt(nsamples);
end

hypPOST.cov =
[mean(result(:,3));mean(result(:,4));mean(result(:,5));mean(result(:,6)
);...

mean(result(:,7));mean(result(:,8));mean(result(:,9));mean(result(:,10)
);...

mean(result(:,11));mean(result(:,12));mean(result(:,13));mean(result(:,
14));...

mean(result(:,15));mean(result(:,16));mean(result(:,17));mean(result(:,
18));...

mean(result(:,19));mean(result(:,20));mean(result(:,21));mean(result(:,
22));...
    mean(result(:,23));mean(result(:,24));mean(result(:,25))];
% predictive output mean and std dev
[logmuPOST,logs2POST] = gp(hypPOST,@infExact, meanfuncPOST,
covfuncPOST, likfuncPOST, XtrainD1, log(YtrainD1), Znewb);
[logmuPOSTD1,logs2POSTD1] = gp(hypPOST,@infExact, meanfuncPOST,
covfuncPOST, likfuncPOST, XtrainD1, log(YtrainD1), XnewD1);

%transform back to strictly positive domain
muPOST = exp(logmuPOST+0.5*logs2POST);
s2POST = (exp(logs2POST)-1).*exp(2*logmuPOST+logs2POST);
```

```
fPOST = [exp(logmuPOST-norminv(1-
significance/2)*sqrt(logs2POST));flipdim(exp(logmuPOST+norminv(1-
significance/2)*sqrt(logs2POST)),1)];

muPOSTD1 = exp(logmuPOSTD1+0.5*logs2POSTD1);
s2POSTD1 = (exp(logs2POSTD1)-1).*exp(2*logmuPOSTD1+logs2POSTD1);
fPOSTD1 = [exp(logmuPOSTD1-norminv(1-
significance/2)*sqrt(logs2POSTD1));flipdim(exp(logmuPOSTD1+norminv(1-
significance/2)*sqrt(logs2POSTD1)),1)];

% FALSE CALL POSTERIOR PARAMETER
FCPOST = betafit(result(:,26));
[FCmuPOST,FCSDPOST] = normfit(result(:,26));
FCMCerror = FCSDPOST./sqrt(nsamples);

% ========================================
%% PLOTS
% ========================================
% The prior crack growth curve
% ========================================
figure(1)
fill([Znewb; flipdim(Znewb,1)], DBnew_fb, [7 7 2]/8)
hold on;
plot(Znewb, DBnew_newmb,'k:');
dat = plot(XtrainD1(:,1),YtrainD1,
'b*',XnewD1test(:,1),YnewD1test,'r*',XnewD0(:,1),zeros(length(XnewD0(:,
1)),1),'r+',Znewb(:,1), DBnew_newmb,'k:');
xlabel('Fatigue Cycles')
ylabel('Crack Length (\mum)')
hold off
xlim([8000,10500])
ylim([0,2200])
legend(dat,'D = 1 Training','D = 1 Test','D = 0','Prior Growth
Curve','location','Northwest')

% ========================================
% the posterior crack growth curve
% ========================================
figure(2)
fill([Znewb; flipdim(Znewb,1)], fPOST, [7 7 2]/8)
hold on;
plot(Znewb(:,1), muPOST,'k:');
datb = plot(XtrainD1(:,1),YtrainD1,
'b*',XnewD1test(:,1),YnewD1test,'r*',XnewD0(:,1),zeros(length(XnewD0(:,
1)),1),'r+',Znewb(:,1), muPOST,'k:');
xlabel('Fatigue Cycles')
ylabel('True Crack Length (\mum)')
hold off
xlim([8000,10500])
ylim([0,2200])
legend(datb,'D = 1 Training','D = 1 Test','D = 0','Posterior Growth
Curve','location','Northwest')

% ========================================
% Prior and Posterior POD Curves
% ========================================
```

```
atrue=linspace(alow,240,1000);
[bet0]=icdf(priors(1,1),[.025 .50 .975]);
[bet1]=icdf(priors(2,1),[.025 .50 .975]);
[bet0POST]=icdf(pd1,[.025 .50 .975]);
[bet1POST]=icdf(pd2,[.025 .50 .975]);
logist_lowPRIOR=loglogisticpodcdf(atrue,bet0(1),bet1(1));
logist_meanPRIOR=loglogisticpodcdf(atrue,bet0(2),bet1(2));
logist_highPRIOR=loglogisticpodcdf(atrue,bet0(3),bet1(3));
logist_lowPOST=loglogisticpodcdf(atrue,bet0POST(1),bet1POST(1));
logist_meanPOST=loglogisticpodcdf(atrue,bet0POST(2),bet1POST(2));
logist_highPOST=loglogisticpodcdf(atrue,bet0POST(3),bet1POST(3));

figure(3)
plot(atrue,logist_meanPRIOR,'b-
',atrue,logist_lowPRIOR,'b:',atrue,logist_highPRIOR,'b--')
title('Prior Logistic POD CDF')
xlabel('True crack size (\mum)')
ylabel('POD CDF')
legend('mean','2.5% bound','97.5% bound')
ylim([0 1])

figure(4)
plot(atrue,logist_meanPOST,'b-
',atrue,logist_lowPOST,'b:',atrue,logist_highPOST,'b--')
title('Posterior Logistic POD CDF')
xlabel('True crack size (\mum)')
ylabel('POD CDF')
legend('mean','2.5% bound','97.5% bound')
ylim([0 1])

% ================================================================
% Posterior Parameter Distributions
% ================================================================
figure(5)
subplot(5,6,1)
hist(result(:,1),100)
xlabel('Loglogistic POD \beta_0 parameter')
ylabel('Frequency')
subplot(5,6,2)
hist(result(:,2),100)
xlabel('Loglogistic POD \beta_1 parameter')
ylabel('Frequency')
subplot(5,6,7)
hist(result(:,3),100)
xlabel('GPR log(A_1) parameter')
ylabel('Frequency')
subplot(5,6,8)
hist(result(:,4),100)
xlabel('GPR log(A_2) parameter')
ylabel('Frequency')
subplot(5,6,9)
hist(result(:,5),100)
xlabel('GPR log(A_3) parameter')
ylabel('Frequency')
subplot(5,6,10)
hist(result(:,6),100)
```

```
xlabel('GPR log(A_4) parameter')
ylabel('Frequency')
subplot(5,6,11)
hist(result(:,7),100)
xlabel('GPR log(A_5) parameter')
ylabel('Frequency')
subplot(5,6,12)
hist(result(:,8),100)
xlabel('GPR log(A_6) parameter')
ylabel('Frequency')
subplot(5,6,13)
hist(result(:,9),100)
xlabel('GPR log(A_7) parameter')
ylabel('Frequency')
subplot(5,6,14)
hist(result(:,10),100)
xlabel('GPR log(A_8) parameter')
ylabel('Frequency')
subplot(5,6,15)
hist(result(:,11),100)
xlabel('GPR log(A_9) parameter')
ylabel('Frequency')
subplot(5,6,16)
hist(result(:,12),100)
xlabel('GPR log(A_1_0) parameter')
ylabel('Frequency')
subplot(5,6,17)
hist(result(:,13),100)
xlabel('GPR log(A_1_1) parameter')
ylabel('Frequency')
subplot(5,6,18)
hist(result(:,14),100)
xlabel('GPR log(A_1_2) parameter')
ylabel('Frequency')
subplot(5,6,19)
hist(result(:,15),100)
xlabel('GPR log(A_1_3) parameter')
ylabel('Frequency')
subplot(5,6,20)
hist(result(:,16),100)
xlabel('GPR log(A_1_4) parameter')
ylabel('Frequency')
subplot(5,6,21)
hist(result(:,17),100)
xlabel('GPR log(A_1_5) parameter')
ylabel('Frequency')
subplot(5,6,22)
hist(result(:,18),100)
xlabel('GPR log(A_1_6) parameter')
ylabel('Frequency')
subplot(5,6,23)
hist(result(:,19),100)
xlabel('GPR log(A_1_7) parameter')
ylabel('Frequency')
subplot(5,6,24)
hist(result(:,20),100)
xlabel('GPR log(A_1_8) parameter')
```

```
ylabel('Frequency')
subplot(5,6,25)
hist(result(:,21),100)
xlabel('GPR log(A_1_9) parameter')
ylabel('Frequency')
subplot(5,6,26)
hist(result(:,22),100)
xlabel('GPR log(A_2_0) parameter')
ylabel('Frequency')
subplot(5,6,27)
hist(result(:,23),100)
xlabel('GPR log(A_2_1) parameter')
ylabel('Frequency')
subplot(5,6,28)
hist(result(:,24),100)
xlabel('GPR log(A_2_2) parameter')
ylabel('Frequency')
subplot(5,6,29)
hist(result(:,25),100)
xlabel('GPR log(A_2_3) parameter')
ylabel('Frequency')
subplot(5,6,30)
hist(result(:,26),100)
xlabel('False call probability P_F_C parameter')
ylabel('Frequency')

figure(6)
subplot(5,6,1)
plot(x1,F1,'r--',x1,cdf(pd1,x1),'b:')
xlabel('Loglogistic POD \beta_0 parameter')
ylabel('Posterior CDF')
subplot(5,6,2)
plot(x2,F2,'r--',x2,cdf(pd2,x2),'b:')
xlabel('Loglogistic POD \beta_1 parameter')
ylabel('Posterior CDF')
subplot(5,6,7)
plot(x3,F3,'r--',x3,cdf(pd3,x3),'b:')
xlabel('GPR log(A_1) parameter')
ylabel('Posterior CDF')
subplot(5,6,8)
plot(x4,F4,'r--',x4,cdf(pd4,x4),'b:')
xlabel('GPR log(A_2) parameter')
ylabel('Posterior CDF')
subplot(5,6,9)
plot(x5,F5,'r--',x5,cdf(pd5,x5),'b:')
xlabel('GPR log(A_3) parameter')
ylabel('Posterior CDF')
subplot(5,6,10)
plot(x6,F6,'r--',x6,cdf(pd6,x6),'b:')
xlabel('GPR log(A_4) parameter')
ylabel('Posterior CDF')
subplot(5,6,11)
plot(x7,F7,'r--',x7,cdf(pd7,x7),'b:')
xlabel('GPR log(A_5) parameter')
ylabel('Posterior CDF')
subplot(5,6,12)
plot(x8,F8,'r--',x8,cdf(pd8,x8),'b:')
```

```
xlabel('GPR log(A_6) parameter')
ylabel('Posterior CDF')
subplot(5,6,13)
plot(x9,F9,'r--',x9,cdf(pd9,x9),'b:')
xlabel('GPR log(A_7) parameter')
ylabel('Posterior CDF')
subplot(5,6,14)
plot(x10,F10,'r--',x10,cdf(pd10,x10),'b:')
xlabel('GPR log(A_8) parameter')
ylabel('Posterior CDF')
subplot(5,6,15)
plot(x11,F11,'r--',x11,cdf(pd11,x11),'b:')
xlabel('GPR log(A_9) parameter')
ylabel('Posterior CDF')
subplot(5,6,16)
plot(x12,F12,'r--',x12,cdf(pd12,x12),'b:')
xlabel('GPR log(A_1_0) parameter')
ylabel('Posterior CDF')
subplot(5,6,17)
plot(x13,F13,'r--',x13,cdf(pd13,x13),'b:')
xlabel('GPR log(A_1_1) parameter')
ylabel('Posterior CDF')
subplot(5,6,18)
plot(x14,F14,'r--',x14,cdf(pd14,x14),'b:')
xlabel('GPR log(A_1_2) parameter')
ylabel('Posterior CDF')
subplot(5,6,19)
plot(x15,F15,'r--',x15,cdf(pd15,x15),'b:')
xlabel('GPR log(A_1_3) parameter')
ylabel('Posterior CDF')
subplot(5,6,20)
plot(x16,F16,'r--',x16,cdf(pd16,x16),'b:')
xlabel('GPR log(A_1_4) parameter')
ylabel('Posterior CDF')
subplot(5,6,21)
plot(x17,F17,'r--',x17,cdf(pd17,x17),'b:')
xlabel('GPR log(A_1_5) parameter')
ylabel('Posterior CDF')
subplot(5,6,22)
plot(x18,F18,'r--',x18,cdf(pd18,x18),'b:')
xlabel('GPR log(A_1_6) parameter')
ylabel('Posterior CDF')
subplot(5,6,23)
plot(x19,F19,'r--',x19,cdf(pd19,x19),'b:')
xlabel('GPR log(A_1_7) parameter')
ylabel('Posterior CDF')
subplot(5,6,24)
plot(x20,F20,'r--',x20,cdf(pd20,x20),'b:')
xlabel('GPR log(A_1_8) parameter')
ylabel('Posterior CDF')
subplot(5,6,25)
plot(x21,F21,'r--',x21,cdf(pd21,x21),'b:')
xlabel('GPR log(A_1_9) parameter')
ylabel('Posterior CDF')
subplot(5,6,26)
plot(x22,F22,'r--',x22,cdf(pd22,x22),'b:')
xlabel('GPR log(A_2_0) parameter')
```

```
ylabel('Posterior CDF')
subplot(5,6,27)
plot(x23,F23,'r--',x23,cdf(pd23,x23),'b:')
xlabel('GPR log(A_2_1) parameter')
ylabel('Posterior CDF')
subplot(5,6,28)
plot(x24,F24,'r--',x24,cdf(pd24,x24),'b:')
xlabel('GPR log(A_2_2) parameter')
ylabel('Posterior CDF')
subplot(5,6,29)
plot(x25,F25,'r--',x25,cdf(pd25,x25),'b:')
xlabel('GPR log(A_2_3) parameter')
ylabel('Posterior CDF')
subplot(5,6,30)
plot(x26,F26,'r--',x26,cdf(pd26,x26),'b:')
xlabel('False call probability P_F_C parameter')
ylabel('Posterior CDF')

figure(13)
plot(YnewD1,muPOSTD1,'r+')
xlabel('Test Crack lengths (\mum)')
ylabel('Model Prediction Crack lengths (\mum)')

save results_loglogistic_SR_K4_1B3
```

Appendix B.2 Kernel Function

This kernel function is the same as that of Equation (3.42) in Chapter 3 Section 3.4.5. It is recommended that the user designs their own kernel function using the options provided in the GPML MATLAB package designed by Rasmussen and available online (Rasmussen, Nickisch, & Williams, 2015). This MATLAB package must be loaded prior to use of this function. Original credit for this function belongs to Dr. Martin Wayne.

```
% ====================================================================
% Kernel Function
% ====================================================================
% Original function by Dr. Martin Wayne
% Modified by Reuel Smith
% ====================================================================
function [nlml, result, newm, news2,K,f]= kernel7(x,y,z,significance)

% define the D variable
D = size(x);
D = D(2);
% define z as linear space plus one step beyond data
% z = linspace(1,length(y)+50,length(y)+50)';

% transform to log scale
y = log(y);

% covariance function determination
covfunc =
{@covSum,{@covConst,@covLINard,@covSEard,@covNNone,@covNoise}};
% intial guess of hyper-parameters of covariance function
hyp.cov = [0;zeros(D,1);zeros(D,1);log(0.1);0;log(0.1);log(0.1)];

% mean function determination
meanfunc = @meanZero;
% intial guess of hyper-parameters of mean function
% hyp.mean = mean(y);

% likelihood function determination
likfunc = @likGauss;
% intial guess of hyper-parameters of likelihood function
hyp.lik = log(0.1);

% choose optimal hyper-parameters by minimizing log likelihood with
% conjugate gradient optimizer
hyp = minimize(hyp, @gp, -500, @infExact, meanfunc, covfunc, likfunc,
x, y);

% negative log likelihood for comparison purposes
nlml = gp(hyp, @infExact, meanfunc, covfunc, likfunc, x, y);
```

```
% predictive output mean and std dev
[m s2] = gp(hyp, @infExact, meanfunc, covfunc, likfunc, x, y, z);

%transform back to strictly positive domain
newm = exp(m+0.5*s2);
news2 = (exp(s2)-1).*exp(2*m+s2);

% 100*(1-significance) percent probability intervals
f = [exp(m-norminv(1-significance/2)*sqrt(s2));flipdim(exp(m+norminv(1-
significance/2)*sqrt(s2)),1)];

% plot results and actual data
% figure(2)
% fill([z; flipdim(z,1)], f, [7 7 2]/8)
% hold on;
% plot((z(:,4)), newm);
% % plot(exp(x(:,4)), exp(y), 'r+');
% hold off
% compute vector of hyper-parameters for model selection
result = [exp(hyp.cov)];

K = feval(covfunc{:}, hyp.cov, x);  % covariance matrix
```

Appendix B.3 Recursive Bayesian Particle Filter Data Generator

This routine generates the PF data for a given set of fatigue test specimen data and its AE data by way of the methodology defined in Chapter 4 Section 4.3. The routine also requires the GPML MATLAB code package as well as a linear extrapolation function (Appendix B.4). The latter function is needed to match the unknown true crack lengths with the known AE indices. This routine is best applied when the AE data is over 100 data. The specimen data set used for this example is specimen 5A2.

```
% ========================================================================
% Recursive Bayes Estimation Particle Filtered Data Generation
% ========================================================================
% by Reuel Smith
% ========================================================================
clear
clc
run('C:\Users\ReuelS\Desktop\POD Research\gpml-matlab-v3.6\startup.m')

% ========================================
% LOAD DATA
% ========================================
load('DBdata.mat','Data_5A2');
load('DBAEdata.mat','Data_AE_5A2');
load('measurementerror','ME')
% ========================================
% CONSTANTS AND DATA INITIALIZATION
% ========================================
N = 1000;                   % The number of particles at each time step
A0 = 35;                    % threshold amplitude in dB
Ea = ME(1);                 % 50% measurement error
significance=0.05;
% ========================================
% Renames the test data and isolates the cycle and crack data
testdat = Data_5A2;testN = testdat(:,3);testa =
testdat(:,1)./Ea;testCSF = testdat(:,3:11);
Init_a = testa(20:23);
testNc = testN(20:end);testac = testa(20:end);testCSFc =
testCSF(20:end,:);
for i = 1:length(testac)/4
    testai(i,1) = mean(testac((i-1)*4+1:i*4));
    testNi(i,1) = mean(testNc((i-1)*4+1:i*4));
end
testCSFi = [testNi ones(length(testNi),1)*testCSF(1,2:end)];
% ========================================
% Renames the AE test data and isolates the cycle, cumulative count (C)
and
% cumulative amplitude (A)
```

```
AEtestddat = Data_AE_5A2;AEtestN = AEtestddat(:,1);AEtestC =
AEtestddat(:,4);AEtestA = AEtestddat(:,5);AEtestCSFs = [AEtestN
ones(length(AEtestN),1)*testCSF(1,2:end)];
% ========================================
% Set the initial true crack length point and its distribution
Init_a_param = lognfit(Init_a);
a0 = lognrnd(Init_a_param(1),Init_a_param(2),[N,1]);
% Set initial cycle
N0 = testN(20).*ones(N,1);
% Set initial weight
w0 = (1/N).*ones(N,1);
% ========================================
% Set up input sets: cycles, cumulative counts, and cumulative
amplitude
% such that the entire range is covered.  Extrapolate what isn't
readily
% available
% ========================================

for i = 1:length(AEtestN)
    if AEtestN(i) > testNc(1)
        if AEtestN(i) == testNc(1)
            ZN = sort([AEtestN(i:end);testNc(end)]);
            ZC =
[AEtestC(i:end);extrapolater(AEtestN,AEtestC,testNc(end))];
            ZA =
[AEtestA(i:end);extrapolater(AEtestN,AEtestA,testNc(end))];
        else
            ZN = sort([testNc(1);AEtestN(i:end);testNc(end)]);
            ZC =
[extrapolator(AEtestN,AEtestC,testNc(1));AEtestC(i:end);extrapolater(AE
testN,AEtestC,testNc(end))];
            ZA =
[extrapolator(AEtestN,AEtestA,testNc(1));AEtestA(i:end);extrapolater(AE
testN,AEtestA,testNc(end))];
        end
        break;
    end
end
% The CSF set of the new set of data
ZCSF = [ZN ones(length(ZN),1)*testCSF(1,2:end)];
% Calculates the mean approximation of the true crack length path
[DBnlml, DBresult, DBnewm, DBnews2,DBK,DBf]=
kernel7(testCSFc,testac,ZCSF,significance);
% The standard deviation of the white noise
sigWN = sqrt(DBresult(end));
% Calculates the change in crack length between steps
testda = diff(DBnewm);
% Calculates the da/dN
testdadN = testda./diff(ZN);
% Calculate the AE intensity parameters
p = polyfit(log((ZC.*ZA)./A0),log(DBnewm),1);beta = p(1);alpha =
exp(p(2));sigv = mean(sqrt(DBnews2));
% ========================================
% SAMPLE PROPAGATION
% ========================================
% Propagate each of the N samples for k number of steps
```

```
% Number of steps - k
K = length(ZN);
% initialize the crack length, cycle, and weight matrices
aRBE = [a0 zeros(N,K-1)];NRBE = ones(N,1)*ZN';wRBE = [w0 zeros(N,K-
1)];anew = a0;wnew = w0;
ameanRBE = [mean(a0);zeros(K-1,1)];

for k = 2:K % NUMBER OF STEPS
    % update of the crack length for the next step
    aold = anew;
    anew = anew + testda(k-1).*exp(normrnd(0,sigWN,[N,1]));
    mu = (1/beta)*log(anew) - (1/beta)*log(alpha);
    % Distribution zK|xk
    pdistnew = (1/(ZC(k)*ZA(k)*sigv*sqrt(2*pi))).*exp(-
0.5.*(((log((ZC(k)*ZA(k))/A0) - mu)./sigv).^2));
    % Calculate the new weights for the next step
    wnew = wnew.*pdistnew;
    % Normalize the weights for the next step
    wnew = wnew./sum(wnew);
    p2 = lognfit(anew);
    for i = 2:N % NUMBER OF SAMPLES
        if logncdf(anew(i),p2(1),p2(2)) == wnew(i)
            anew(i) = aold(i) + testda(k-1)*exp(normrnd(0,sigWN));
            wnew(i) = 1/N;
        end
    end
    wnew = wnew./sum(wnew);
    aest = sum(anew.*wnew);
    ameanRBE(k) = aest;
    wRBE(:,k) = wnew;
    aRBE(:,k) = anew;
end
% ========================================
% PLOTS
% ========================================
figure(1)
% plot(testN,testa,'b*',N0,a0,'r.')
% plot(testN,testa,'b*',N0,a0,'r.',ZCSF(:,1),DBnewm,'b:')
plot(testN,testa,'b*',reshape(NRBE,[K*N,1]),reshape(aRBE,[K*N,1]),'g.',
ZN,ameanRBE,'k:')
xlabel('Cycles')
ylabel('Crack Length (\mum)')
legend('Test crack lengths (\mum)','RBE crack lengths (\mum)','Mean
estimate crack length (\mum)','location','Northwest')
xlim([9500,11000])

figure(2)
plot(AEtestN,AEtestC,'r.',AEtestN,AEtestA,'b.')
%
plot(AEtestN,AEtestC,'r.',AEtestN,AEtestA,'b.',AEtestN,Cnewm,'r:',AEtes
tN,Anewm,'b:')

save('Data_RBE_5A2','ZN','ZCSF','NRBE','aRBE','ameanRBE')
```

Appendix B.4 Linear Extrapolation

As stated in Appendix B3, the linear extrapolation function is designed to match unknown variables with known variables so that they are all set at the same time stamp or step. This tool works best when the AE index data numbers 100 or more.

```
% ====================================================================
% Linear Extrapolator
% ====================================================================
% by Reuel Smith
% ====================================================================
function Y = extrapolator(x,y,X)

for i = 2:length(x)
    if x(i-1) < X && x(i) > X
        low = i-1;
        hi = i;
    end
    if x(i-1) < X && x(i) < X
        low = i-1;
        hi = i;
    end
end
p = polyfit([x(low) x(hi)],[y(low) y(hi)],1);
Y = polyval(p,X);
```

Bibliography

Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002). A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, *50*(2), 174-188.

ASTM. (2015). *ASTM E976-15 Standard Guide for Determining the Reproducibility of Acoustic Emission Sensor Response.* West Conshohocken, PA: ASTM International.

ASTM E466-15. (2015). Standard Practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials. West Conshohocken, PA: ASTM International.

Bannantine, J. A., Comer, J. J., & Handrock, J. L. (1990). *Fundamentals of Matal Fatigue Analysis.* Upper Saddle River, NJ: Prentic-Hall Inc.

Bao, T., Peng, Y., Cong, P., & Wang, J. (2010). Analysis of Crack Propagation in Concrete Structures with Structural Information Entropy. *Science China Technological Sciences, 53*(7), 1943-1948.

Bayes, T. (1763). An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions, 53*, 370–418.

Bencala, K. E., & Seinfeld, J. H. (1976). On Frequency distributions of Air Pollutant Concentrations. *Atmospheric Environment, 10*, 941-950.

Berger, A., Melice, J. L., & Demuth, C. (1982). Statistical Distributions of Daily and High Atmospheric SO2 Concentrations. *Atmospheric Environment, 16*(12), 2863-2877.

Box, G. E. (1979). Robustness in the Strategy of Scientific Model Building. *Robustness in Statistics*, 201-236.

Chen, T., Morris, J., & Martin, E. (2007). Gaussian Process Regression for Multivariate Spectroscopic Calibration. *Chemometrics and Intelligent Laboratory Systems*, *87*, 59–71.

Clinton Aluminum & Stainless Steel. (2014, 12 22). *Grade-7075-Text-data - Grade-7075-Text-data.pdf*. Retrieved from Aluminum 7075-T6; 7075-T651: http://www.clintonaluminum.com/wp-content/uploads/2014/08/Grade-7075-Text-data.pdf

CRP MECCANICA S.r.l. (2017, 1 17). *aluminium-7075-t6-7075-t651.pdf*. Retrieved from Aluminum 7075-T6; 7075-T651: http://www.crpmeccanica.com/PDF/aluminium-7075-t6-7075-t651.pdf

Davidson, D. L., & Lankford, J. (2013). Fatigue Crack Growth in Metals and Alloys: Mechanisms and Mircomechanics. *International Materials Reviews*, 45-76.

Doucet, A., De Freitas, N., & Gordon, N. (2001). An Introduction to Sequential Monte Carlo Methods. In *Sequential Monte Carlo Methods in Practice* (pp. 3-14). New York: Springer.

Doucet, A., Godsill, S., & Andrieu, C. (2000). On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing, 10*, 197–208.

eFatigue LLC. (2017). *eFatigue - Fatigue analysis on the Web*. Retrieved from eFatigue - a trusted source for fatigue analysis: https://www.efatigue.com/

Ekengren, J. (2008). *Estimating inclusion content in high performance steels.* Universitetstryckeriet, Karlstad: Karlstad University Studies.

Forman, R. G., Kearney, V. E., & Eagle, R. M. (1967). Numerical Analysis of Crack Propagation in Cyclic Loaded Structures. *Journal of Basic Engineering, 89*, 459-464.

Georgiou, G. A. (2006). *Probability of Detection (PoD) Curves: Derivation, applications and limitations.* London: Crown.

Griffith, A. A. (1921). The Phenomena of Rupture and Flow in Solids. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 221*(582-593), 163-198.

Hanlon, T., Kwon, Y. N., & Suresh, S. (2003). Grain size effects on the fatigue response of nanocrystalline metals. *Scripta Materialia, 49*, 675–680.

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika, 57*(1), 97-109.

Hertzberg, R. W., Manson, J. A., & Skibo, M. D. (1975). Frequency Sensitivity of Fatigue Processes in Polymeric Solids. *Polymer Engineering & Science, 15*(4), 1548-2634.

Hertzberg, R. W., Manson, J. A., & Skibo, M. D. (1980). Fatigue Crack Propagation in Nylon 66 Blends. *Fracture Mechanics: Proceedings of the Twelfth National*

Symposium on Fracture Mechanics ASTM STP700 (pp. 49-64). Baltimore, MD: American Society for Testing and Materials.

Irwin, G. R. (1948). Fracture Dynamics. *Fracturing of Metals, ASM*, 152.

Irwin, G. R. (1957). Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate. *Journal of Applied Mechanics, 24*, 361-364.

Jakeman, A. J., & Taylor, J. A. (1985). A Hybrid ATDL-Gamma Distribution Model for PRedicting Urban Area Source Acid Gas Concentrations. *Atmospheric Environment, 19*, 1959-1967.

Jones, R., Peng, D., Huang, P., & Singh, R. R. (2015). Crack growth from naturally occurring material discontinuities in operational aircraft. *Procedia Engineering, 101*, 227–234.

Kappatos, A. V., & Dermatas, S. E. (2007). Feature Extraction for Crack Detection in Rain Conditions. *Journal of Nondestructive Evaluation, 26*, 57-70.

Keshtgar, A. (2013). *Accoustic Emission-Based Structural Health Management and Prognostics Subject to Small Fatigue Cracks.* College Park, MD: University of Maryland.

Kleene, S. C. (1956). Representation of Events in Nerve Nets and Finite Automata. *Annals of Mathematics Studies, 34*, 3–41.

Kujawski, D., & Ellyin, F. (1992). Crack Initiation and Total Fatigue Life of a Carbon Steel in Vacuum and Air. *Journal of testing and evaluation, 20*(6), 391–395.

MacKenzie, S. (2008). Overview of the Mechanisms of Failure in Heat Treated Steel Components. *Failure Analysis of Heat Treated Steel Components (ASM International)*, 43-86.

Mage, D. T. (1981). A Review of the Application of Probability Models for Aerometric Data. *Environmetrics 81*. Philadelphia.

MakeItFrom.com. (2016, 8 6). *7075-T6 Aluminum - MakeItFrom.com*. Retrieved from 7075-T6 Aluminum: http://www.makeitfrom.com/material-properties/7075-T6-Aluminum

Mathworks. (2015). MATLAB 2015b.

MathWorks. (2016). MATLAB R2016a.

McCulloch, W., & Pitts, W. (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics, 5*(4), 115–133.

McDonald, M., Zaman, K., & Mahadevan, S. (2009). Representation and First-Order Approximations for Propagation of Aleatory and Distribution Parameter Uncertainty. *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. Palm Springs, CA: American Institute of Aeronautics and Astronautics, Inc.

Mohanty, S., Chattopadhyay, A., & Peralta, P. (2011). Bayesian Statistic Based Multivariate Gaussian Process Approach for Offline/Online Fatigue Crack Growth Prediction. *Experimental Mechanics, 51*, 833–843.

Mohanty, S., Chattopadhyay, A., Peralta, P., Das, S., & Willhauck, C. (2007). *Fatigue Life Prediction Using Multivariate Gaussian Process.* AIAA.

Molent, L., Barter, S., & Jones, R. (2008). Some Practical Implications of Exponential Crack Growth. *Solid Mechanics and Its Applications*, 65-84.

Moore, C., & Doherty, J. (2005). Role of the calibration process in reducing model predictive error. *Water Resources Research, 41*, W05020.

Naderi, M., Kahirdeh, A., & Khonsari, M. M. (2012). Dissipated thermal energy and damage evolution of Glass/Epoxy using infrared thermography and acoustic emission. *Composites: Part B, 43*(3), 1613–1620.

National Aeronautics and Space Administration. (1971). *NASA Space Vehicle Design Criteria (Structures): Preliminary Criteria for the Fracture Control of Space Shuttle Structures.* NASA.

Nelson, H. G., Williams, D. P., & Tetelman, A. S. (1971). Embrittlement of a Ferrous Alloy in a Partially Dissociated Hydrogen Environment. *Metallurgical Transactions*, *2*, 953-959.

Niendorf, T., Rubitschek, F., Maier, H. J., Canadinc, D., & Karaman, I. (2010). On the fatigue crack growth–microstructure relationship in ultrafine-grained interstitial-free steel. *Journal of Materials Science, 45*(17), 4813-4821.

NIH. (2015, September 27). *ImageJ Version 1.50c.* Retrieved October 14, 2015, from http://imagej.nih.gov/ij/index.html

Ontiveros, V., Cartillier, A., & Modarres, M. (2010). An Integrated Methodology for Assessing Fire Simulation Code Uncertainty. *Nuclear Science and Engineering*, 179–201.

Ott, W. R. (1995). *Environmental Statistics and Data Analysis.* Boca Raton: Taylor & Francis.

Paris, P., & Erdogan, F. (1963). A Critical Analysis of Crack Propagation Laws. *Journal of Basic Engineering, 85*(2), 528-534.

Petrin Jr., C., Annis Jr, C., & Vukelich, S. I. (1993). *A Recommended Methodology to Quantify NDE/NDI Based on Aircraft Engine Experience.* Essex: Specialised Printing Services Limited.

Physical Acoustic Corporation. (2007). AEwin Software User's Manual. Princeton Junction, NJ.

Rabiei, E., Droguett, E. L., & Modarres, M. (2016). A Prognostics Approach Based on the Evolution of Damage Precursors using Dynamic Bayesian Networks. *Advances in Mechanical Engineering, 8*(9), 1-19.

Rabiei, M. (2011). *A Bayesian Framework for Structural Health Management Using Acoustic Emission Monitoring and Periodic Inspections.*

Rabiei, M., & Modarres, M. (2013). A Recursive Bayesian Framework for Structural Health Management Using Online Monitoring and Periodic Inspections. *Reliability Engineering & System Safety,* 154–164.

Randelius, M. (2008). *Influence of microstructure on fatigue and ductility properties of tool steels.* Stockholm, Sweden: Department of Materials Science and Engineering Royal Institute of Technology.

Rasmussen, C. E. (1996). *Evaluation of Gaussian Processes and other Methods for Non-Linear Regression.* Toronto: University of Toronto.

Rasmussen, C. E., Nickisch, H., & Williams, C. (2015 , July 7). *Documentation for GPML Matlab Code version 3.6.* Retrieved October 21, 2015, from http://www.gaussianprocess.org/gpml/code/matlab/doc/

Rountree, L. C. (1990). *The Radar Range Equation for the Detection of Steady Targets in Weibull Clutter.* Wright-Patterson Air Force Base: Air Force Institute of Technology.

Rummel, W. D., & Matzkanin, G. A. (1997). *Nondestructive Evaluation (NDE) Capabilities Data Book (3rd Edition).* Austin, TX: Nondeatructive Testing Information Analysis Center (NTIAC).

Rummel, W. D., & Matzkanin, G. A. (1997). *Non-Destructive Evaluation (NDE) Capabilities Data Book, 3rd Ed.* Nondestructive Testing Information Analysis Center.

Rusk, D. (2011). *Model Development Plan: Hinge Inspection Reliability.*

Sankararaman, S., Ling, Y., Shantz, C., & Mahadevan, S. (2009). Uncertainty Quantification in Fatigue Damage Prognosis. *Annual Conference of the Prognostics and Health Management Society.* San Diego.

Sauerbrunn, C. M. (2016). *Evaluation of Information Entropy from Acoustic Emission Waveforms as a Fatigue Damage Metric for Al7075-T6.* College Park, MD: University of Maryland.

Schleher, D. C. (1976). Radar Detection in Weibull Clutter. *IEEE Transactions on Aerospace and Electronic Systems, AES-12*(6), 736-743.

Sekine, M., & Mao, Y. (1990). *Weibull Radar Clutter.* London: Peter Peregrinus Ltd.

Singh, K. P., Warsono, & Bartolucci, A. A. (1997). Generalized Log-Logistic Model for Analysis of Environmental Pollutant Data. *MODSIM 97 IMACS Proceedings.* Hobart.

Skibo, M. D. (1977). *PhD Dissertation.* Bethlehem, PA: Lehigh University.

Smith, R., & Modarres, M. (2017). Small Crack Fatigue Growth and Detection Modeling with Uncertainty and Acoustic Emission Application. *Contributions to Statistics,* [In publication].

Smith, R., Modarres, M., & Droguett, E. L. (2017). Small Crack Fatigue Propagation Detection Modeling with Applications to Recursive Bayesian Estimation. *International Journal of Prognostics and Health Management,* (in review).

Thornton, W. C., & Tiffany, C. F. (1970). *Fracture Control of Metallic Pressure Vessels.* National Aeronautics and Space Administration (NASA).

Wacholder, S., Chanock, S., Garcia-Closas, M., El Ghormli, L., & Rothman, N. (2004). Assessing the Probability That a Positive Report is False: An Approach for

Molecular Epidemiology Studies. *Journal of the National Cancer Institute, 96*(6), 434-442.

Walker, K. (1970). The Effect of Stress Ratio during Crack Propagation and Fatigue for 2024-T3 and 7075-T6 Aluminum. *Effects of Environment and Complex Load History on Fatigue Life ASTM STP 462*, 1-14.

Williams, D. P. (1973). A New Criterion for Failure of Materials by Environment-Induced Cracking. *International Journal of Fracture, 9*(1), 63-74.

Yuan, X., Mao, D., & Pandey, M. D. (2009). A Bayesian approach to modeling and predicting pitting flaws in steam generator tubes. *Reliability Engineering & System Safety, 94*(11), 1838-1847.