©2021

Changxi Wang

ALL RIGHTS RESERVED

# **RELIABILITY ESTIMATION AND STOCHASTIC MODELING OF**

# **CORROSION GROWTH**

by

Changxi Wang

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Industrial and Systems Engineering

Written under the direction of

Elsayed A. Elsayed

And approved by

New Brunswick, New Jersey January, 2021

# **ABSTRACT OF THE DISSERTATION**

# RELIABILITY ESTIMATION AND STOCHASTIC MODELING OF CORROSION GROWTH By CHANGXI WANG Dissertation Director:

Elsayed A. Elsayed

Corrosion and corrosion-related problems are the major factors leading to the agerelated structural degradation of infrastructures such as pipelines and pressure vessels. Corrosion defects may result in severe damages such as thickness penetration, fatigue cracks, brittle fracture, rupture and burst. Quantifying the growth of corrosion is critically important for the risk and reliability analysis of structures, planning for corrosion mitigation, repairs and determination of time intervals for the corrosion inspections and monitoring.

Pitting corrosion growth has been a focus of research, and a depth threshold of corrosion defect has been widely used for estimating the residual life of structures being monitored. However, corrosion volume loss of materials may also lead to failures such as pipeline bursts, which is more harmful but overlooked. In this dissertation, we develop a degradation model that characterizes both corrosion maximum depth growth, corrosion volume growth and corrosion propagation. We propose an improved inverse Gaussian (IIG) process to model the corrosion depth growth and demonstrate that it captures the dependency between the corrosion growth rate and the corrosion depth. We develop a corrosion pit volume growth model assuming that both the corrosion pit growth in the depth directions and radial directions follow IIG processes. Compared with other existing corrosion models, the proposed model captures the phenomenon where a critical amount of volume loss of materials leads to the failure of a component even though the corrosion defect's depth has not reached its failure threshold. A physics-based model that incorporates factors including the spatial and size distributions of the material particles and the influence of corroded pits is developed to capture the corrosion propagation.

Degradation branching stochastic models are developed to describe the corrosion pit propagation. They are general models and can be applied to cracks in materials and systems that consist of multiple units where the degradation of one unit may affect adjacent units and the failure occurs when the total degradation reaches an unacceptable amount. The models capture the phenomenon that a growing degradation branch may initiate new branches when a certain criterion is met, where the criterion may be a branch's degradation amount threshold or other physical processes. The effect of the random branching angles and the random number of branches initiated in each branching on the total degradation is investigated, where the physics-based models are proposed to describe the relationship between the branching angles and the total degradation. The branching continues until the total degradation of all the branches reaches a threshold. Statistics of the degradation branching processes such as the mean and the variance of the total degradation, the expected number of branches initiated, the reliability and distribution of residual life are obtained.

The measurements of corrosion growth are continuously monitored and recorded. The rapid development of sensing and computing technologies has enabled the use of multiple sensors to monitor the degradation indicator (or its surrogate) of a component simultaneously. However, there are challenges in integrating the measurements from multiple sensors. First, missing data arises due to data transmission failures and manipulation errors. We propose a variety of stochastic bridges to deal with the missing data. Second, different sensors may capture different aspects of the degradation process and may be sensitive in different stages of the degradation process. We propose a non-parametric model that assigns a sensor's weight (contribution) based on its performance in the previous time instants. It utilizes a moving time window to determine the switching of the sensors between clusters with time so that the weights are adjusted accordingly. The advantage of the proposed approach is that no specific distribution of degradation data or underlying degradation models are required.

#### ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my advisor Professor Elsayed A. Elsayed for his guidance, patience, inspiration, encouragement and support through my Ph.D. study over the years. His knowledge and personality have had a profound influence on my academic career and life. He has set me an example of excellence as a mentor, instructor and role model. My sincere thanks are extended to Professor Susan Albin, Professor Javier Cabrera and Professor Hoang Pham for serving on my committee and providing me with valuable suggestions that improve the quality of my dissertation.

My special thanks go to Professor Tie Gang from Harbin Institute of Technology for his insightful academic instructions and the experimental set-up of this research.

I also take this opportunity to thank my wife Ting Wang. It is her academic suggestion, mental support and love that encourage and accompany me through the challenging, interesting and unforgettable journey. I would also like to thank my parents, parentsin-law and grandparents for their understanding and selfless support on all my decisions.

# LIST OF TABLES

<b>Table 2.1</b> Experimentally observed branching angles in fracture mechanics
specimens26
<b>Table 3.1</b> The time to failure prediction accuracy and sum of squared errors66
<b>Table 4.1</b> The information of numerical calculation of integrals
<b>Table 5.1</b> The different cases of branching patterns in the case study
Table 5.2 The calculation steps in Case 1 130
Table 5.3 The calculation steps in Case 2 132
Table 5.4 The calculation steps in Case 3
Table 6.1 Descriptions of the sensors
Table 6.2 Mean squared errors of interpolated values 173
<b>Table 6.3</b> Example of sensors data observed during five-time instants
<b>Table 6.4</b> Summary of parameters used in the case study
<b>Table 6.5</b> MSE of individual sensors and the integrated paths
Table 6.6 Comparison results of absolute percentage error by using each sensor
and integrated paths with different percentage of actual remaining lifetime

# LIST OF FIGURES

Figure 1.1 Three-dimensional X-ray micro-tomographic image of a corrosion pit
by Horner et al. (2011) (notice the morphologies of the corrosion growth) .5
Figure 2.1 Schematics of the common forms of corrosion (Davis, 2000)10
Figure 2.2 A graphical representation of definitions associated with the
branching tree
Figure 2.3 The rotations of a crack surface
Figure 2.4 (a) Definitions of the angular location and direction of branching
from a flaw adjacent to the primary fracture (b) The branching direction as a
function of fracture velocity (Rabinovitch et al., 2011)
Figure 2.5 (a) Infrared image of the failed membranes of Polymer Electrolyte
Membrane (PEM) Fuel Cell. The bright regions represent leakage caused by
cracks (b) Cross-sectional views of branched PEM cracks (Singh et al.,
2017)
Figure 2.6 Crack propagation and branching in a fuel cell membrane
(Thompson, 2007)
Figure 3.1 (a) Degradation versus time (b) degradation versus starting crack
length (c) degradation increments versus time (d) degradation increments
versus starting crack length42
<b>Figure 3.2</b> (a) The predicted 90% probability interval and the other nine units (b)
The failure probability curves for IG and IIG44

<b>Figure 3.3</b> (a) The corrosion depth growth data (b) The expectation and 95%
confidence interval with the IIG model45
Figure 3.4 Failure probability curve for IG and IIG47
Figure 3.5 Use the regular IG model to fit IIG data when (a) data are censored at
t = 7 (c) data are censored at $t = 11$ , use the IIG model to fit IIG data when
(b) data are censored at $t = 7$ (d) data are censored at $t = 11$ 49
<b>Figure 3.6</b> Use the IIG model to fit IG data when (a) data are censored at $t = 12$
(b) data are censored at $t = 8$ , use the IG model to fit IG data when (c) data
are censored at $t = 12$ (d) data are censored at $t = 8$
Figure 3.7 Schematic diagram of the corrosion pit growth
Figure 3.8 Corrosion area growth
Figure 3.9 Corrosion data and predicted mean corrosion growth path61
<b>Figure 3.10</b> Selected simulated corrosion pit 3-D images from $t = 3$ to $t = 9$ 61
Figure 3.11 Predicted volume growth and maximum depth growth62
Figure 3.12 Reliability when considering different failure modes
Figure 3.13 Corrosion loss prediction comparison using different models64
Figure 3.14 Corrosion depth and volume prediction when data are censored at
t = 10
Figure 3.15 Reliability when considering both depth growth and volume growth
Figure 3.16 Particles and pits (Liao <i>et al.</i> , 1998)68

Figure 5.17 Thea data for particles on the ES and ET surfaces of common 7075
T6 aluminum alloy with estimated CDFs (Harlow, 2012)69
Figure 3.18 The overlap of the pits71
<b>Figure 3.19</b> The coalescence of the two pits at $t_2$
Figure 3.20 The particles in a material
Figure 3.21 Selected figures of the pits over time
Figure 3.22 (a) Total volume growth of the pits (b) Reliability of the unit with
multiple pits for threshold80
Figure 4.1 Schematic diagram of the DBP85
Figure 4.2 The CDFs of IG distribution and exponential distribution
Figure 4.3 Selected figures of the contours of DGS images of crack growth and
branching process in a soda-lime glass specimen (Dondeti et al., 2019)97
Figure 4.4 Selected photos of PE fringes of crack growth and branching process
in a soda-lime glass specimen (Dondeti et al., 2019)97
Figure 4.5 The crack growth data of the three specimens used in the case study
Figure 4.6 The apparent SIFs of cracks over time
Figure 4.7 The estimated PDF of time to branch based on SIF growth and
observed times of branching of specimens100
Figure 4.8 The actual and predicted total crack length

Figure 4.9 The reliability of the unit/system subject to crack growth and branching......102 Figure 5.3 Schematic diagram of the degradation of the original branch and Figure 5.4 (a) Definitions of the angular location and direction of branching from a flaw adjacent to the primary fracture (b) The branching direction as a function of fracture velocity (Rabinovitch et al., 2011) ......116 **Figure 5.5** The growth of tree k after time t, the bold solid lines represent Figure 5.6 The linear fit of the crack propagation rate—branching angle data 128 Figure 5.10 The comparison of the PDF of total degradation obtained by Figure 5.12 The PDFs of the total degradation over time obtained by numerical 

Figure 5.13 The means and variances obtained with simulations and numerical
solutions144
Figure 5.14 The reliability of the system obtained by numerical solution and
simulations144
Figure 6.1 Comparison of the piecewise linear mean function and linear mean
function149
<b>Figure 6.2</b> Estimate the parameter $c_1$ by considering the observed data150
Figure 6.3 The plot of the variance of a nonlinear Brownian bridge152
Figure 6.4 The schematic diagram of a Gamma Bridge154
Figure 6.5 The variance plot of a Gamma bridge158
Figure 6.6 The interpolated data with the nonlinear Brownian Bridge161
Figure 6.7 The interpolated data with the nonlinear Gamma Bridge162
Figure 6.8 Interpolated data with the inverse Gaussian Bridge
Figure 6.9 The corrosion growth prediction with data interpolated by the
nonlinear Brownian Bridge163
Figure 6.10 The corrosion growth prediction with data interpolated by the
nonlinear Gamma Bridge164
Figure 6.11 Corrosion growth prediction with data interpolated by the inverse
Gaussian Bridge164
Figure 6.12 Failure time prediction confidence intervals with 40% missing data
interpolated from the bridges166

Figure 6.13 Failure time prediction confidence intervals with 20% missing data
interpolated from the bridges166
Figure 6.14 Failure time prediction confidence intervals with 10% missing data
interpolated from the bridges167
Figure 6.15 SSPE using three bridges when different proportions of data are
missing168
Figure 6.16 The complete dataset
Figure 6.17 The missing data pattern
Figure 6.18 An interpolated missing dataset171
<b>Figure 6.19</b> The variance plot
Figure 6.20 Mean interpolated data by stochastic bridges (dashed lines denote
interpolated data)
Figure 6.21 Mean interpolated data by modified bootstrap and <i>k</i> NNI (dashed
lines denote interpolated data)174
Figure 6.22 Cluster switching example (sensors 2, 3 and 6 switched clusters at
time 2)177
Figure 6.23 (a) Weights of the sensors by variance (b) Weights of the sensors by
cluster shifting181
Figure 6.24 Final normalized weights of the sensors
Figure 6.25 Integrated paths and mean of the incomplete data

# **TABLE OF CONTENTS**

ABSTRA	ACT OF	THE DISSERTATIONii
ACKNO	WLED	GEMENTv
LIST OF	TABL	ESvi
LIST OF	FIGU	RESvii
CHAPTE	ER 1	INTRODUCTION
1.1	The	e Motivations of the Research1
1.2	Pro	blem Definition
	1.2.1	Stochastic Modeling of Corrosion Growth
	1.2.2	Stochastic Modeling of Degradation Branching Processes
	1.2.3	Missing Data Interpolation7
	1.2.4	Multiple Sensor Data Interpolation and Integration7
1.3	Org	ganization of the Dissertation8
CHAPTE	ER 2	LITERATURE REVIEW
2.1	Lite	erature Review on Degradation Models9
	2.1.1	General Corrosion Concepts
	2.1.2	Physics-based Corrosion Degradation Models13
	2.1.3	Deterministic Corrosion Degradation Models15
	2.1.4	Stochastic Corrosion Degradation Models17
	2.1.5	Degradation Propagation and Stochastic Branching Processes21

2.2	Lite	erature Review on Missing Degradation Data Interpolation	30
2.3	Lite	erature Review on Data Integration	35
CHAPTER	3	CORROSION GROWTH MODELING	37
3.1	Intr	roduction	37
3.2	Co	rrosion Depth Growth Modeling	37
3.	.2.1	Inverse Gaussian Process	37
3.	.2.2	Improved Inverse Gaussian Process	39
3.	.2.3	Remaining Life Prediction and Failure Probability Estimation	40
3.	.2.4	Model Validation and Numerical Studies	42
3.	.2.5	IIG Model Robustness Evaluation	47
3.3	Co	rrosion Volume Growth Modeling	51
3.	.3.1	Volume Growth of the Corroded Pixels	52
3.	.3.2	Volume Growth of the Newly Corroded Pixels	53
3.	.3.3	Total Incremental Volume	55
3.	.3.4	Incremental Volume Under Stresses	56
3.	.3.5	Reliability Estimation and Remaining Life Prediction	57
3.	.3.6	Case Study	59
3.4	Co	rrosion Propagation Modeling	66
3.	.4.1	Characterization of the Particles	67
3.	.4.2	Corrosion Pits Initiation	69
3.	.4.3	Corrosion Growth of Multiple Pits	71

	3.4.4	Multiple Pits Initiation and Growth: Case Study78
3.5	Su	mmary
СНАРТЕ	ER 4	DEGRADATION BRANCHING MODELING
4.1	Int	roduction82
4.2	Th	e Degradation Branching Process Model83
	4.2.1	Properties of the DBP83
	4.2.2	Special Case of the DBP: Branching Brownian Motion Process91
	4.2.3	Case Study95
4.3	Su	mmary
CHAPTE	ER 5	GENERAL DIRECTIONAL DEGRADATION BRANCHING
PROCES	SES	
5.1	Int	roduction104
5.2	No	menclature
5.3	Ge	neral Directional Degradation Branching Process109
	5.3.1	Definitions and Properties of the GDDBP109
	5.3.2	Probabilistic Properties of the GDDBP116
5.4	Ca	se Study127
	5.4.1	Assumptions and Parameters Used in the Case Study127
	5.4.2	Total Degradation Estimation
	5.4.3	Comparison of the Simulations and the Theoretical Results140
5.5	Su	mmary145

CHAPTER 6	CORROSION MONITORING AND PREDICTION	146
6.1 I	ntroduction	146
6.1 N	Missing Data Interpolation Models	147
6.1.	1 Nonlinear Brownian Bridge	147
6.1.2	2 Nonlinear Gamma Bridge	153
6.1	3 Inverse Gaussian Bridge	159
6.1.4	4 Models Validation	160
6.2 N	Aultiple Sensors Data Integration	174
6.2.	1 Clustering of the Sensors' Observations	175
6.2.2	2 Sensors' Data Integration	177
6.2.3	3 Case Study	
6.3 S	Summary	
CHAPTER 7	CONCLUSIONS AND FUTURE RESEARCH	187
7.1 C	Conclusions	
7.2 F	Future Research	190
APPENDIX A	A PARAMETERS ESTIMATION IN IIG AND IG PROCESSES	S 191
APPENDIX B	B THE CRACK GROWTH DATA	194
APPENDIX C	C CORROSION VOLUME GROWTH MODELING BASED O	N
IMPROVED	GAMMA PROCESS	195
C.1 Corre	osion Pit Area Growth	198
C.2 Volu	me Growth of the Corroded Pixels	201

C.3 The Incremental Volume with Covariates
C.4 Reliability and Remaining Life Prediction
C.5 Parameters Estimation
C.6 True Area Growth and Recommendations to Select Number of Sectors211
C.7 Case Study214
C.8 Conclusions
APPENDIX D THE CORROSION VOLUME THRESHOLD DERIVATION227
APPENDIX E PARAMETERS ESTIMATION IN THE CORROSION VOLUME
GROWTH PROCESS BASED ON IIG PROCESS
APPENDIX F PROOFS OF THE DEGRADATION BRANCHING EQUATIONS
F.1 The PDF of the Total Degradation, the Reliability and the Expected Number
of Branches
F.2 The Expectation of the Total Degradation and the Mean Residual Life235
F.3 The Variance of the Total Degradation238
F.4 The Expectation and Variance of the Branches
F.5 The Mean and Variance of the Total Degradation where Time to Branch
Follows an Exponential Distribution245
F.6 The Mean and Variance of the Branches where Time to Branch Follows an
Exponential Distribution251
APPENDIX G MODIFIED K NEAREST NEIGHBOUR IMPUTATION

APPENDIX H IMPUTATION WITH BOOTSTRAP	
H.1 Degradation Increments are i.i.d.	
H.2 Residuals are Approximately i.i.d.	
H.3 Residuals are Non-i.i.d	
REFERENCES	

#### **CHAPTER 1**

#### **INTRODUCTION**

# 1.1 The Motivations of the Research

Infrastructures are susceptible to degradation throughout their service lives. Corrosion is one of the most common degradation forms. It is a major threat to the structural integrity and safe operation of infrastructures throughout the world. Due to the severe impact of corrosion defects, various reliability estimation models and corrosion management methods have been studied (Morcillo *et al.*, 2013). The most popular models are the corrosion pit depth growth models. An extensive review of the literature reveals stochastic modeling of the corrosion growth is limited (Ghosh *et al.*, 2007).

Field data (Hu *et al.*, 2016, Valor *et al.*, 2013) reveals that the corrosion growth rate decreases as corrosion progresses. It is generally accepted that the rust layer, which mainly consists of the produced reaction products, stops water and air from contacting new metal and decreases corrosion rate. No model captures the dependency between the corrosion growth rate and the corrosion layers, as well as the measurements' uncertainties.

As corrosion progresses, corrosion pit grows larger both in radial and depth directions. At the same time, the corresponding corrosion pit's volume increases, which decreases the remaining strength of the structures. While current research mainly focuses on modeling the structural failures caused by corrosion pit depth growth, there is little work on modeling the structural failures induced by corrosion pit volume growth. There are cases where shallow pits lead to failures. Therefore, it is necessary to model the pit volume growth to accurately predict a structure's failure time and its residual life.

The corrosion growth rate is significantly affected by the environmental stresses, which include temperature, relative humidity and the pH levels (Guedes Soares *et al.*, 2011). As has been shown by Hughes *et al.* (2016), there is an inherent relationship between the underlying levels of stresses and the corrosion electric current, which can be interpreted as the corrosion rate. Therefore, the effect of environmental stresses needs to be accurately incorporated into the corrosion growth model.

The corrosion degradation measurements are susceptible to errors, missing observations, noises and biases (Huang *et al.*, 2009, Seem, 2007). With the advancement of sensing technology, more sensors that monitor different degradation indicators are readily available. By employing multiple sensors, more accurate information can be obtained since they portray different perspectives of the underlying degradation process. Meanwhile, the joint behavior of different sensors has not been investigated. Moreover, due to sensor failures and human errors, missing data may occur in the data acquisition process. Therefore,

appropriate data imputation approaches and data integration approaches need to be investigated. These problems are the subject of investigation of this dissertation.

## **1.2 Problem Definition**

#### **1.2.1** Stochastic Modeling of Corrosion Growth

#### 1.2.1.1 Stochastic Modeling of Corrosion Depth Growth

After a corrosion pit is initiated at micro-particles in the metal matrix (Trzaskoma, 1990), it begins to grow both in the radial and depth directions. The corrosion growth rates in these directions are not deterministic. This is due to the uncertainties of material microstructure, physical constitution and other environmental factors. For example, the non-uniformity of constituents that compose the steel alloy may lead to an inhomogeneous metal dissolution rate (corrosion growth rate) in different locations of the alloy. The non-homogeneous microstructure of the metal around the welds may also lead to various corrosion growth rates.

The environmental stresses, which include the relative humidity, the temperature and the pH levels significantly influence the corrosion growth rate (Soares *et al.*, 2009). According to the corrosion kinetics, the stresses affect the corrosion growth rate by changing the corrosion chemical reaction current (Hughes *et al.*, 2016). However, physics-

based stress incorporation models are sparse. Meanwhile, it is widely observed that as corrosion progresses, the corrosion growth rate decreases (Murer *et al.*, 2013, Vanaei *et al.*, 2017). This is mainly caused by the formation of corrosion rust layer, which prevents reagent such as water and air from corroding the uncorroded metal. This phenomenon may also be interpreted as that there is a relationship between the corrosion growth rate and the thickness of the corrosion layer. The literature is also sparse in addressing the phenomenon.

# 1.2.1.2 Stochastic Modeling of Corrosion Volume Growth

Critical corrosion pit depth has been used as a structural failure threshold for a long time (Al-Amin *et al.*, 2016, Burstein *et al.*, 2004). The structure fails when the maximum corrosion pit depth exceeds the threshold. However, there are cases where the corrosion pit with non-critical depth but a large volume of corrosion results in pipeline burst (Netto *et al.*, 2005). Therefore, structure failures induced by corrosion pit volume growth should be carefully investigated as well.

Challenges arise when dealing with this problem. First, the morphologies of the corrosion pits are highly stochastic and irregular (Horner *et al.*, 2011, Kioumarsi *et al.*, 2016, Trzaskoma, 1990), as shown in Figure 1.1 (Horner *et al.*, 2011). We may observe that the corrosion propagates with different rates in all directions. It is inappropriate to assume that

the shape of a pit is a perfect smooth hemisphere or half ellipsoid as presented in (Kondo, 1989, Yu *et al.*, 2015) when estimating the pit volume growth. Second, the proposed corrosion pit volume growth model should be compatible with existing corrosion depth growth models. Third, there is a dependency between the corrosion depth growth rate of a location in the pit and its relative location to the center of the pit. The locations close to the center are likely to have large depth, while the locations close to the boundary of the pit are usually shallow. Moreover, the depth growth rates of the locations close to the center are similar to each other. Likewise, the growths in locations close to the boundary are similar.



Figure 1.1 Three-dimensional X-ray micro-tomographic image of a corrosion pit by

Horner et al. (2011) (notice the morphologies of the corrosion growth)

Corrosion pits may propagate to other areas of the material, leading to the initiation of new pits. Existing statistical methods on quantitatively modeling the propagation and volume growth of corrosion pits are sparse. The challenge lies in the following three aspects: first, the corrosion pits are related to the particles in the material. Second, the initiation probability of a particle is related to its size and the influence of its nearby pits. Third, nearby pits may coalesce and form larger pits, which have a larger impact on the reduction of the strength of materials. A comprehensive model that deals with the challenges mentioned above needs to be developed.

# 1.2.2 Stochastic Modeling of Degradation Branching Processes

Degradation branching is a common phenomenon in many real-life applications. The degradation of a location/unit not only increases with time but also propagates to other locations/units in the same system. While the degradation of an individual location/unit has been studied extensively, the research on degradation branching is sparse. A stochastic model that considers both degradation growth and degradation propagation is proposed in this dissertation.

Missing data have challenged researchers since the beginning of the field research. The challenge has been particularly acute for degradation monitoring, that is, research involving multiple time series on the same individuals. Missing data arise due to data transmission failure, human manipulation errors, etc. Data interpolation may replace the missing values with estimated values. Simple data interpolation methods include mean/mode substitution, nearest neighbor substitution (Graham, 2009) and others. However, they may lead to biased statistical estimates. Regression models are applied to interpolate missing values assuming the underlying true path is a linear function of time (Van Buuren, 2012). However, the regression models assume the residuals that represent the uncertainties follow a normal distribution, which is unrealistic. Better stochastic data interpolation models should be proposed to capture different types of missing data.

#### **1.2.4** Multiple Sensor Data Interpolation and Integration

When multiple sensors are used to monitor the degradation process, we need a model to assign weights to the sensors and find the integrated measurement. However, potential challenges remain. First, due to sensor failures or human errors, observations may be missing. This may lead to inaccurate parameter estimates. Second, some of the sensors may be sensitive to different sizes or types of defects. For instance, the eddy current approach presents more accurate measurements when defects are cracks and are located close to the surface. X-rays are sensitive to volumetric defects and ultrasound is sensitive to defects that are perpendicular to the ultrasound rays. Likewise, despite its accuracy, the metallographic microscope is only able to detect open defects. Third, as all sensors are monitoring the same defect, the measurements from different sensors are highly correlated. For example, there is a high probability that the ultrasonic signal and eddy current signal of a defect are both strong or weak.

#### **1.3** Organization of the Dissertation

This dissertation is organized as follows. In Chapter 2, we present a comprehensive literature review of the corrosion models, reliability metrics, missing data imputation approaches, data integration approaches and degradation branching models. In Chapter 3, we investigate the stochastic approaches to model the corrosion volume growth, as well as the influences of covariates and the propagation of corrosion. In Chapters 4 and 5, the degradation branching models are developed. In Chapter 6, missing degradation data imputation models using stochastic bridges and data integration approaches are proposed. In Chapter 7, we give conclusions and future research directions.

# **CHAPTER 2**

# LITERATURE REVIEW

#### 2.1 Literature Review on Degradation Models

#### 2.1.1 General Corrosion Concepts

Corrosion occurs in several widely differing forms. Classification is usually based on one of the three factors (Davis, 2000):

- *Nature of the Corrodent*: Corrosion can be "wet" or "dry." The former usually requires a liquid or moisture; the latter usually involves reaction with high-temperature gases.
- *Corrosion Mechanisms*: This involves either electrochemical or direct chemical reactions.
- *Corrosion Appearance*: Corrosion is either uniform, and the metal corrodes at the same rate over the entire surface, or it is localized, in which case only small areas are affected.

Classification by appearance, which is particularly useful in the failure analysis, is based on identifying forms of corrosion by visual observation with either the naked eyes or magnification. The morphology of the defects is the basis for classification. Figure 2.1 schematically illustrates some of the most common forms of corrosion (Davis, 2000). Eight forms of wet (aqueous) corrosion can be identified based on the appearance of the corroded metal. These are:

- Pitting corrosion (Designation, 2009, Frankel, 1998, Hubbell *et al.*, 1985)
- Uniform corrosion (Schweitzer, 1983, Wranglén, 1985)
- Crevice corrosion (Szklarska-Smialowska et al., 2005)
- Galvanic corrosion (Oldfield, 1988)
- Erosion corrosion (Kerr *et al.*, 1958), including fretting corrosion
- Intergranular corrosion (Lin et al., 1995) and exfoliation
- Dealloying (Vukmirovic *et al.*, 2002)
- Environmentally assisted cracking, including stress-corrosion cracking, corrosion fatigue (Suresh *et al.*, 1981)



Figure 2.1 Schematics of the common forms of corrosion (Davis, 2000)

Amongst all types of corrosion, pitting is one of the most common and damaging forms of corrosion especially in marine and offshore structures. Pitting takes the form of cavities on the surface of a metal. It starts with the local breakdown of the protective surface film. It may cause the perforation of thin sections, as well as create stress concentrations that may trigger the onset of fatigue cracking or other types of corrosion. The damage is generally assumed to nucleate on the bare surface as a pit due to localized galvanic corrosion surrounding the exposed constituent particles in the alloy (Harlow *et al.*, 1998). As the pit continues to grow, neighboring particles are exposed, and likewise, they are corroded and contribute to further pit growth. The corrosion of metals is generally due to an irreversible oxidation-reduction (redox) reaction between the metal and an oxidizing agent present in the environment:

$$metal + oxidizing \ agent \rightarrow oxidized \ metal + reducing \ agent$$
 (2.1)

For example, the corrosion of iron in the presence of hydrochloric acid is

$$Fe + 2HCl \rightarrow FeCl_2 + H_2$$
 (2.2)

In the aqueous phase, hydrochloric acid and ferrous chloride exist in ionic form. We can therefore also write Eq. (2.2) as:

$$Fe + 2H^+ + 2Cl^- \rightarrow Fe^{2+} + 2Cl^- + H_2$$
 (2.3)

In this case, the oxidizing agent is the solvated proton  $H^+$ . The products of the reaction are the solvated ferrous ion  $Fe^{2+}$  and gaseous hydrogen  $H_2$ . Under neutral and alkaline conditions, the corrosion of metals is generally caused by a reaction of the metal with oxygen. For example, when exposed to air and humidity iron forms rust *FeOOH* (Landolt, 2007).

$$Fe + O_2 + H_2O \to 4FeOOH \tag{2.4}$$

As the reaction proceeds, the produced rust accumulates and forms a layer preventing air and water vapor from contacting metal (Vanaei *et al.*, 2017). As a result, the corrosion rate decreases as reaction proceeds and more corrosion products are produced.

Because pitting corrosion is common and basic, in this dissertation we mainly discuss this type of corrosion. Corrosion may lead to thinning of walls, crack initiation (Ebara, 2007, Hu *et al.*, 2016, Huang *et al.*, 2013) and ultimately cause perforation, leakage, reduced strength and even burst of components (Cheng *et al.*, 2012, Kioumarsi *et al.*, 2016, Nakai *et al.*, 2004, Otieno *et al.*, 2008).

# 2.1.2 Physics-based Corrosion Degradation Models

Following Hoeppner (1979), by assuming a hemispherical pit of radius r and a constant bulk dissolution rate B, Kondo (1989) proposes that the pit volume increases proportionally to time t:

$$\frac{2}{3}\pi r^3 = Bt \tag{2.5}$$

and hence

$$r \propto t^{1/3} \tag{2.6}$$

Harlow *et al.* (1995), Wei (2001) and Assoul *et al.* (2008) assume that a pit grows at a constant volumetric rate dV/dt. In keeping with Kondo *et al.* (1989), they propose that

$$\frac{dV}{dt} = \frac{MI_{p_0}}{nF\rho} \exp\left[-\frac{E_a}{RT}\right]$$
(2.7)

where *M* is the molecular weight of the material,  $\eta$  is the valence; F = 96514c/mol is Faraday's constant,  $\rho$  is the density,  $E_a$  is the activation energy, R = 8.314 J/mol-K is the universal gas constant, *T* is the absolute temperature,  $I_{p_0}$  is the pitting current coefficient.

Since the assumption of hemisphere shape is far too unrealistic, a reasonable approximation (Harlow *et al.*, 1995) of the shape is one half of a prolate spheroid that has volume

$$V = \frac{2}{3}\pi ab^2 \tag{2.8}$$

where *a* and *b* are one-half the lengths of the major and minor axes. For a reaction where the activation overvoltage limits the rate, the relationship between the rate of reaction (expressed by a current density i) and the driving force for the reaction, or potential *E*, is presented by the Butler-Volmer equation (Roberge, 2000):

$$i = i_0 \exp\left[\frac{\alpha n F\left(E - E_{rev}\right)}{RT}\right]$$
(2.9)

where *F* is the Faraday constant = 96,487 C/equivalent, *n* is the charge on the ion in equivalents/mol, *R* is the gas constant = 8.314 J/mol-K,  $\alpha$  is the unitless charge transfer coefficient,  $E_{rev}$  is the reversible potential where each of the half-reactions will be in equilibrium and  $i_0$  is the reaction current in equilibrium. We will discuss the application of Eq. (2.9) on developing the stress incorporation models in Chapter 3.

Despite the extensive study of physics-based corrosion models, their limitations are obvious. First, these models are difficult to obtain. Extensive experiments must be carried out based on strong knowledge of physics, chemistry and metallurgical properties of the components. Second, physics-based degradation models tend to result in less accurate prediction as field conditions are complicated.

# 2.1.3 Deterministic Corrosion Degradation Models

Deterministic corrosion growth models are most frequently discussed among all existing models, which include power functions, logarithmic functions, linear functions and power-linear functions.

It is observed that pit-growth laws are of the form of power functions (Chookah, 2009, Goldstein, 1960, Kawai *et al.*, 1985, Noor *et al.*, 2011, Panchenko *et al.*, 2016, Panchenko *et al.*, 2014, Turnbull *et al.*, 2006a, Turnbull *et al.*, 2006b, Turnbull *et al.*, 2004):

$$h = bt^n \tag{2.10}$$

where h is the maximum depth of a pit, t is time, b and n are the parameters related to material-environment systems. The value of n is nearly constant. b changes wildly depending on the material-environmental system. Little information on b is available but would be best determined by field data.

The use of logarithmic functions to model the corrosion depth has also been investigated (de la Fuente *et al.*, 2007, De la Fuente *et al.*, 2011, Haynie *et al.*, 1971, Ma *et al.*, 2010, Morcillo *et al.*, 1993, Syed, 2013):

$$\log D = \log A + n \log t \tag{2.11}$$

where *D* is the maximum depth, *A* and *n* are constant parameters, *t* is time. The variation of *n* indicates the degree of the corrosion process, n > 1 indicates the corrosion acceleration process, whereas n < 1 indicates the corrosion deceleration process.

Linear functions, due to its simplicity, are investigated by Achterbosch *et al.* (2006) and Din *et al.* (2009):

$$y = \hat{\alpha}^0 + \hat{\alpha}^1 t \tag{2.12}$$

where y is the corrosion pit depth;  $\hat{\alpha}^0$  and  $\hat{\alpha}^1$  are parameters to be determined by field data.

The bi-linear model (ISO, 1992) is proposed for the estimation of possible corrosion losses in each category of atmosphere corrosivity. In this model, the parabolic variation of corrosion losses during the first years until the onset of the stationary phase (up to 10 years) is replaced, in first approximation, by a linear increase of corrosion loss D with an average rate of  $r_{av}$ . At the stationary stage after ten years, the corrosion process occurs at a constant rate  $r_{lin}$ . It should be noted that the annual gain in D amounts to  $r_{lin}$ . The total corrosion loss D is represented by the sum of corrosion losses over the first ten years and over a subsequent period when the stationary corrosion process has established:

$$D = tr_{av} + (t - 10)r_{lin}$$
(2.13)

The metal corrosion loss under atmospheric conditions is also found to obey a power-linear function (ISO, 1992, Knotkova *et al.*, 2010, McCuen *et al.*, 1992, Morcillo *et al.*, 2013, Panchenko, 2013, Panchenko *et al.*, 2016, Panchenko *et al.*, 2014): a power law at the initial stage assuming a parabolic increase in metal corrosion loss during the first ten years of exposure and a linear law for exposures longer than ten years.

$$D(t < 10) = r_{corr}t^{b} \tag{2.14}$$

$$D(t \ge 10) = r_{corr} \left[ 10^{b} + b(10^{b-1})(t-10) \right]$$
(2.15)

where *D* is corrosion mass loss;  $r_{corr}$  means corrosion mass losses over the first year; *b* is a coefficient characterizing the environment; *t* is time.

# 2.1.4 Stochastic Corrosion Degradation Models

Deterministic models have been used extensively due to its simplicity. However, the underlying corrosion growth process is inherently stochastic and includes temporal and spatial variability (Stewart *et al.*, 2008, Stewart *et al.*, 2007). The uncertainties (e.g., measurement error, randomness associated with the corrosion growth and the material properties) are not incorporated in the deterministic models. Meanwhile, the deterministic models only present point estimators regarding the time to failure, which is unuseful in real life. Stochastic models are better alternatives under this condition. The stochastic degradation models in the literature can be categorized as random variable-based models
(Al-Amin *et al.*, 2016, De Leon Escobedo *et al.*, 2005, Meeker *et al.*, 1998, Pandey *et al.*, 2005, Stewart *et al.*, 2008, Velázquez *et al.*, 2009, Zaman *et al.*, 1994) and stochastic process-based models (Caleyo *et al.*, 2009, Hong, 1999a, Hong, 1999b, Little *et al.*, 2004, Maes *et al.*, 2009, Morrison *et al.*, 1992, Ossai *et al.*, 2016, Provan *et al.*, 1989, Qin *et al.*, 2003, Teixeira *et al.*, 2008, Valor *et al.*, 2007, Whiteside II, 2008, Zhou *et al.*, 2012). The latter includes Markov Chains, Gamma process models, inverse Gaussian process models, Brownian motion models and geometrical Brownian motion models.

A general degradation path is proposed by Meeker and Escobar (Meeker *et al.*, 1998). The degradation of unit *i* at time  $t_j$  is modeled as  $y_{ij} = Y_{ij} + \varepsilon_{ij}$  where  $Y_{ij}$  is the actual path and  $\varepsilon_{ij}$  is the residual deviation that follows a normal distribution with mean zero. Velázquez *et al.* (2009) investigate a multivariate regression model on modeling maximum pit depth as

$$\hat{d}_{\max} = \left(k_0 + \sum_{i=1}^n k_i x_i\right) \left(t - t_0\right)^{n_0 + \sum_{j=1}^m \eta_j x_j}$$
(2.16)

where  $\hat{d}_{max}$  is the predictor of the maximum pit depth,  $x_i$  represents the *i*<sup>th</sup> predictor variable,  $t_0$  is corrosion initiation time,  $k_i$  and  $\eta_j$  are the regression coefficients for this predictor. Because the growth rates in the variable-based models are time-independent random variables, they do not capture the temporal variability of the corrosion growth process.

The homogeneous and non-homogenous Markov Chains are used to model the growth of pitting corrosion and determine the optimal inspection intervals for the pipelines. Giorgio *et al.* (2011) propose that the transition probabilities between unit states depend on both the age and state of the unit. The main limitations of the MC-based models are: the accuracy of the model is sensitive to the total number of discrete damage states (Giorgio *et al.*, 2011) and it is not straightforward to evaluate the transition probability function when inspection data are imperfect.

The use of the Gamma process to characterize the deterioration of engineering structures in the context of optimal maintenance decision or time-dependent reliability analysis has been reported extensively in the literature (Cheng *et al.*, 2012, Kallen *et al.*, 2005, Pandey *et al.*, 2005, Park *et al.*, 2005, van Noortwijk *et al.*, 1996, Yuan *et al.*, 2008, Yuan *et al.*, 2009, Zhou *et al.*, 2012). Kallen *et al.* (2005) model degradation of steel pressure vessels and optimal maintenance intervals with a Gamma process. van Noortwijk *et al.* (2007) model the crest-level decline of dikes and evaluate the time-dependent reliability of dikes subjected to sea waves. Zhou *et al.* (2012) use the Gamma process model to characterize the growth of defect depth for evaluating the time-dependent system reliability of pipelines containing multiple active corrosion defects. To improve the Gamma process model, Pandey *et al.* (2005) study the dependency between degradation and time. The dependency of the degradation in the next moment on the current degradation state is considered by Fan *et al.* (2015). Zhang (2014) studies a hierarchical Bayesian-updating Gamma process model. In addition to Gamma processes, an improved Gamma process is proposed to model nonlinear degradation growth (Fan *et al.*, 2015). Multivariate Gamma process is applied to cases where multiple degradation indicators exist (Peng *et al.*, 2016, Zhou *et al.*, 2010).

Two stochastic processes analogs to the Gamma processes, namely the inverse Gaussian process and the Brownian motion process have been reported to model degradation process (Wang, 2010, Wang *et al.*, 2010). The inverse Gaussian process consists of independent increments that follow the inverse Gaussian distributions, whereas the Brownian motion process consists of independent increments that follow Gaussian distributions.

Gebraeel *et al.* (2005) model the degradation process of products with an exponential function where the randomness from unit to unit is described with a centered Brownian motion. The randomness from unit to unit is modeled with a standard normal distribution (Gebraeel, 2006, Gebraeel *et al.*, 2005). Wang (2010) proposes a degradation-based remaining useful life prediction method where the degradation is modeled using a Brownian motion with drift. The drift parameter is adaptive to the history of condition monitoring information. Wu *et al.* (2011) propose a degradation-based reliability estimation method with a random failure threshold.

A notable drawback of using the Brownian motion process to model the corrosion growth is that it cannot rigorously characterize the monotonic nature of the corrosion growth since the Gaussian distributed increments can be either positive or negative. The inverse Gaussian process overcomes this drawback because it is positively defined (Zhang, 2014). Another process, geometric Brownian motion (Elsayed *et al.*, 2004), differs from the Brownian motion in that the former characterizes the logarithm of the degradation as Brownian motion and is used to model corrosion growth by Zhang (2014).

However, these models are limited as they mainly characterize the corrosion depth growth and do not consider other features such as corrosion pit volume, corrosion pit morphology, depth to diameter ratio and others. There are cases where even shallow pits can lead to failures, e.g., rupture of pipelines (Ma *et al.*, 2011, Netto *et al.*, 2005, Stewart *et al.*, 2008). The stochastic nature of corrosion pit volume growth is not captured either.

# 2.1.5 Degradation Propagation and Stochastic Branching Processes

In some cases, the degradation not only grows over time but also propagates over space. Degradation data of this kind usually exhibits complex correlation structures in space and time. The spatiotemporal models have been extensively studied to provide flexible and effective ways to construct covariance functions (Carlin et al., 2014, Cressie et al., 1999, Fuentes et al., 2003, Gneiting, 2002, Liu et al., 2018a, Reich et al., 2011). Kernel functions are used by Liu et al. (2018b), Wikle et al. (1999), Brown et al. (2000) and Sigrist et al. (2015) to express the amount of degradation at a certain location as a weighted sum of its neighbors. Brown et al. (2000) propose the use of convolution to approximate the propagation phenomenon in space under a constant vector field. Sigrist et al. (2015) present a spatio-temporal Gaussian process directly derived from the solution of a stochastic Partial Differential Equation (PDE) describing the convection-diffusion process. Meanwhile, a Fourier spectral method has been used for efficient computation. Calder (2007) proposes a dynamic spatial-temporal model in which the quantity of interest is expressed as the convolution of a latent process. Such a model fits into the general dynamic modeling framework for spatial-temporal data introduced in (Stroud et al., 2001). Yao et al. (2014) develop a degradation behavior propagation model based on the small world clustering. An Ant Colony Optimization algorithm is adopted to obtain the degradation propagation path with a strong pervasion ability.

The branching process is defined as follows (Rahimov, 1995): at time t = 0, it is assumed that only one particle exists. After a random time L, the particle splits into vindependent particles that experience the same transformations. Other branching process models are developed by specifying different reproduction process assumptions of the particles. In particular, when L follows an exponential distribution, the branching process is a branching Brownian Motion process (Fang *et al.*, 2012, Harris *et al.*, 2009). These stochastic branching processes are used in modeling the transport of contaminants in groundwater (Marseguerra *et al.*, 1997), the spread of epidemics (Wang, 1980), population and demography (Berestycki *et al.*, 2010, Berestycki *et al.*, 2015, Wu, 2010), neutrons in nuclear reactions (Vatutin *et al.*, 1985), growth of large organic molecules (Dorogov *et al.*, 1988), and population migration (Sawyer, 1976, Zoia *et al.*, 2014). However, their interests are obtaining the sum of the living particles (reproduction) at a certain time *t* or the maximum distance between the particles and the origin.

Stochastic branching processes are proposed to model the earthquake crack propagation (Kagan, 1982, Libicki *et al.*, 2005). The basic assumption is that an earthquake process is initiated by a set of infinitesimal shear dislocations occurring according to a specific time, space and orientation distributions. The foundation of a mathematical branching model is a branching tree as shown in Figure 2.2. A tree is a set of vertices connected by edges and contains no loops (edges that form an enclosed region). The root vertex is referred to as a parent if it is adjacent to other vertices. In the applications of the model, the vertices and edges are related to actual physical quantities. There are four relevant distributions in the model: Uniform distribution, Poisson distribution, Power-law distribution and Cauchy distribution. The Uniform distribution is commonly used to deal with homogeneous spaces

(Libicki *et al.*, 2005), e.g., the distribution of branching locations. The Poisson distribution is used to model the distribution of the number of offsprings when branching occurs (Leon-Garcia, 1994). The Power-law distribution (Kagan, 1994, Turcotte *et al.*, 2016) and Cauchy distribution (Rice, 2006) describe the rotations of the crack branches' surfaces with respect to their parent crack's surface.



Figure 2.2 A graphical representation of definitions associated with the branching tree

Figure 2.3 shows the rotations of a crack surface. Let  $O_1P$  denote the direction of the parent crack branch, which is perpendicular to the plane. After rotating around the x axis for an angle of  $\varphi_0$  and the y axis for an angle of  $\varphi_1$ , which follow power-law distribution and Cauchy distribution, its offspring finally propagates along the direction of  $O_1P'$ .



Figure 2.3 The rotations of a crack surface

Rabinovitch *et al.* (2011) assume that the secondary fractures are initiated at distances greater than a certain minimal radius r from the primary's vertex. As shown in Figure 2.4 (a),  $\beta$  is the angular location and  $\delta$  is the direction of branching from a flaw adjacent to the primary fracture. Figure 2.4 (b) shows the angle  $\beta_1$  where the stress at flaw tip reaches its maximum and the relationship between branching direction  $\delta_1$  and fracture velocity. Thus it is implied that  $\beta$  is the angle at which the flaw turns into a secondary fracture and  $\beta$  is the direction that crack propagates. Crack branching angles have been experimentally measured and reported for various materials by different researchers. The branching angles reported by various investigators in their dynamic crack propagation studies are shown in Table 2.1.



**Figure 2.4** (a) Definitions of the angular location and direction of branching from a flaw adjacent to the primary fracture (b) The branching direction as a function of fracture

velocity (Rabinovitch et al., 2011)

Source	Branching angle, $\delta$ in degrees
Bullen et al. (1970)	60
Freudenthal (1968)	22-42
Christie et al. (1952)	25
Clark et al. (1966)	34
Congleton et al. (1973)	30-40

 Table 2.1 Experimentally observed branching angles in fracture mechanics specimens

Kirchner et al. (1981)	25
Kalthoff (1971)	28
Kobayashi et al. (1974)	24-34
Ramulu et al. (1984)	28-34
Ravi-Chandar et al. (1984)	20-90
Sakai et al. (1991)	16-70

Though there are similarities between the above branching process and the process being investigated in this dissertation, we are interested in the total degradation of the branches (the total distance of the particles from the origin). Moreover, the degradation branching process under study differs from existing ones in that a branch never terminates but continues to degrade and branches over time. We briefly describe our branching process as follows and use crack propagation as an example.

Crack branching and propagation is a common cause of failures in many materials. Failures due to crack growth and branching have been observed in alumina-silicon carbide composites loaded in fatigue (Han *et al.*, 1989). An initial crack is shown to spawn many sub-cracks until the overall crack network induces failure. The crack branching behavior (how expansive the crack network) is observed to be dependent upon the magnitude of the applied load and the temperature. Physically, there are two proposed mechanisms for crack branching. The first is due to crack deflection, wherein a crack impinges on an interface, such as a grain boundary, and deflects into the interface causing the formation of two cracks along the interface (Ming-Yuan *et al.*, 1989, Parmigiani *et al.*, 2006). The second is due to

crack jumping from a primary interface onto a secondary interface, leading to the nucleation of a new crack which is separate from the initial crack (Dempsey *et al.*, 1985, Ramulu *et al.*, 1985, Sills *et al.*, 2015). The propensity for each process is dictated by the strength, toughness, and elastic characteristics of the materials and interfaces comprising the solid.

In addition to the failure of materials, the crack growth and branching degradation is one of the major causes of structure failures (Guo et al., 2018) such as the failures of concretes and fuel cell membranes (Singh et al., 2017). When the crack length or number of cracks formed reach specified values, failures occur as demonstrated in the following example. Polymer electrolyte membrane (PEM) fuel cells are widely used in transport applications such as electric vehicles to supply electrical energy by converting hydrogen and oxygen (air) into water. In the PEM fuel cells, the polymeric membranes that separate the electrodes allow certain protons to pass through them. It is of vital importance for the membranes to be efficient and reliable. However, the fluctuation of operating conditions poses significant durability challenges for the PEM fuel cells. Figure 2.5 (a) shows an infrared image of the failed membranes, where the bright sections correspond to the high local temperature indicating the through-membrane leakage caused by the cracks. Figure 2.5 (b) shows the cross-sectional views of the branched membrane cracks, which cause crossover leakage. The crack is initiated on the left-hand side and branches during its

propagation to the right. The cracks grow and branch with time. The degradation of the fuel cells, which is represented by leakage amount, is related to the total length of all the crack branches.



**Figure 2.5** (a) Infrared image of the failed membranes of Polymer Electrolyte Membrane (PEM) Fuel Cell. The bright regions represent leakage caused by cracks (b) Cross-sectional views of branched PEM cracks (Singh *et al.*, 2017)

Likewise, Thompson (2007) shows that the total cracks' degradation is the main criterion for fuel cell membranes as shown in Figure 2.6.



Figure 2.6 Crack propagation and branching in a fuel cell membrane (Thompson, 2007)

The branching process can also be exemplified by the case when a system is configured such that the degradation of a unit may cause the adjacent units to begin to deteriorate when the degradation level of the preceding unit reaches a threshold. In our work, we simply use the initiation of cracks in locations, rather than units.

## 2.2 Literature Review on Missing Degradation Data Interpolation

Degradation data are crucial to accelerated degradation testing (ADT) and prognostic health management (PHM). However, in the actual situations, missing values of degradation data are common due to the failure of data transmission or manipulation errors (Fan *et al.*, 2014).

Sometimes a group of data with missing values is abandoned. Examples of this type of techniques include listwise deletion and pairwise deletion (Graham *et al.*, 1993, Wothke,

2000). The listwise deletion, which is also known as complete-case analysis or case deletion, is to discard units whose information is incomplete (Schafer *et al.*, 2002). In contrast, pairwise deletion uses different sets of sample units for different parameters. For more discussions on the properties of case deletion and further references, see Chapter 3 of Little *et al.* (2014). It is a waste of data resources and reduces statistical power. The loss of valuable information may lead to inaccurate model parameters. Schafer *et al.* (2002) and King *et al.* (2001) both report that high rates of listwise deletion lead to serious parameter bias and inefficiency. Brown (1994), Graham *et al.* (1996) and Wothke (2000) report the shortcomings of listwise deletion by simulation. It is possible to reduce biases from case deletion by the judicious application of weights (Liang *et al.*, 1986, Robins *et al.*, 1995). For a review of weighting in the context of sample surveys, see Little *et al.* (2014).

Rather than deleting the incomplete data or ignoring the missing data, the expectationmaximization (EM) algorithm is utilized to "fill in the missing data" with the best guess at what it might be under the current estimate of the unknown parameters and re-estimate the parameters. Repeat the expectation and maximization procedures multiple times until a criterion is satisfied. Dempster *et al.* (1977) show that, rather than filling in the missing data values per se, we are also filling in the complete-data sufficient statistics. These statistics depend on the model specification. Overviews of the EM algorithm can be found in (Little *et al.*, 2014, McLachlan *et al.*, 2007, Schafer, 1997, Schneider, 2001). It is feasible to replace the missing data with plausible values and proceed with the desired analysis. This is called imputation or interpolation. Simple data interpolation techniques include mean/mode substitution (Graham, 2009), hot deck imputation (Ernst, 1980, Marker *et al.*, 2002), cold-deck imputation and composite methods (Schieber, 2005). Hot deck replaces each missing value by a random draw from the observed values. By contrast, cold deck selects donors from another dataset. Composite methods combine ideas from different methods above (David *et al.*, 1986, Schieber, 2005). For example, the hot-deck and regression imputation are combined by calculating the predicted means from a regression model and a residual randomly chosen from the empirical residuals to the predicted value when forming values for imputation.

The imputation methods above are about point estimation of the missing data and model parameters. Some approaches incorporate additional uncertainties. Efron (1994) uses bootstrap to generate imputed samples and inference parameters. Regression-based models are widely used to incorporate the uncertainties in the missing data imputation (Allison, 1987, Buck, 1960, Fan *et al.*, 2014, Van Buuren, 2012, Yates, 1933). Popular regression-based data imputation models are summarized as follows (Van Buuren, 2012)

• Predict.  $\dot{y} = \hat{\beta}_0 + X_{mis}\hat{\beta}_1$ , where  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are the least-squares estimates calculated from the observed data.

- Predict + noise.  $\dot{y} = \hat{\beta}_0 + X_{mis}\hat{\beta}_1 + \dot{\varepsilon}$ , where  $\dot{\varepsilon}$  is randomly drawn from the normal distribution as  $\dot{\varepsilon} \sim N(0, \hat{\sigma}^2)$
- Bayesian multiple imputation.  $\dot{y} = \dot{\beta}_0 + X_{mis}\dot{\beta}_1 + \dot{\varepsilon}$ , where  $\dot{\varepsilon} \sim N(0, \hat{\sigma}^2)$ .  $\dot{\beta}_0$ ,  $\dot{\beta}_1$ and  $\dot{\varepsilon}$  are random draws from their posterior distribution. It is also named "predict + noise + parameters' uncertainty".
- Bootstrap multiple imputation.  $\dot{y} = \hat{\beta}_0 + X_{mis}\hat{\beta}_1 + \dot{\varepsilon}$ , where  $\dot{\varepsilon} \sim N(0, \hat{\sigma}^2)$ .  $\hat{\beta}_0$ ,  $\hat{\beta}_1$ and  $\dot{\varepsilon}$  are the least-squares estimates calculated from a bootstrap sample taken from the observed data. This is an alternative way to implement "predict + noise + parameters uncertainty".

Multiple imputation (Laird, 1988) repeats a simple univariate imputation multiple times, analyzes each imputed dataset and pools the results into one estimate.

The standard Brownian bridge is proposed to impute missing data in finance by DiCesare (2006). Brownian bridge predicts missing data between two "tie ends," assuming the underlying process is a Brownian motion that starts and ends with a fixed value.

For data with multiple features amongst which the inputs are partially missing, approaches such as Support Vector Machine (SVM), A Framework for Imputing Missing Values Using Co-Appearance, Correlation and Similarity Analysis (FIMUS), *k*-Nearest Neighbors

Imputation (kNNI) have been proposed. FIMUS (Rahman et al., 2014) fills in the missing values based on the co-appearances of the values, correlations between the attributes and similarity of the values belonging to an attribute. Pelckmans et al. (2005) propose a convex optimization approach to learn a classification rule from observational data when missing data occur based on SVM. The main idea is to incorporate the uncertainty due to the missing data into an appropriate risk function. An imputing approach based on the Support Vector Regression (SVR) is proposed by Wang et al. (2006). The approach utilizes an orthogonal coding input scheme, which makes use of multi-missing values in one row of a certain gene expression profile and imputes the missing value into a much higher dimensional space to obtain better performance. kNNI finds the k-nearest neighbor records of the observation that has missing attributes and imputes the missing attributes with the average values (Batista et al., 2003). However, it has also been pointed out that choosing a suitable value for k may be a challenging task (Wu *et al.*, 2008). Considering the local similarity of data, many methods based on clustering, for example, k-means or fuzzy cmeans (FCM), have been proposed (Aydilek et al., 2013, Li et al., 2004, Li et al., 2013). The underlying principles are as follows: first random select k complete data objects as k centroids, then iteratively modify the partition to reduce the sum of distances for each object from the centroid of the cluster to which the object belongs until a user-specified threshold  $\varepsilon$  is obtained, the last process is to fill in all the missing attributes by taking the mean of some nearest objects. Decision Tree-based Missing Value Imputation

Technique (DMI) makes use of an entropy-based decision tree algorithm and expectation maximization-based imputation technique (Rahman *et al.*, 2011).

# 2.3 Literature Review on Data Integration

Data integration approaches are generally classified into three categories: data-level integration, feature-level integration and decision-level integration (Klein, 1999, Steinberg *et al.*, 1999). Data-level integration combines raw measurements of individual sensors and forms a unified indicator (Gros *et al.*, 2000, Hua *et al.*, 2013, Liu *et al.*, 2015, Liu *et al.*, 2013, Liu *et al.*, 2016, Yan *et al.*, 2016). Feature-level integration extracts and integrates features such as the frequency spectrum of abnormal signals (Amolins *et al.*, 2007, Balakrishnan *et al.*, 2012, Pajares *et al.*, 2004). Decision-level integration integrates conclusions drawn by individual sensors to reach a comprehensive one. In the case of remaining life prediction, decision-level integration combines life prediction results obtained with different sensors and then concludes final judgment on the condition of the components (units).

Data integration can also be classified into the following categories based on the specific techniques used: Bayesian inference integration (Mitchell, 2007), evidential reasoning integration (Shafer, 1976, Zhao *et al.*, 2014), fuzzy theory integration (Liu *et al.*, 2008, Zadeh, 1999), Dempster-Shafer theory integration (Dempster, 1968, Murphy, 1998, Yager

*et al.*, 2008), wavelet integration (Balakrishnan *et al.*, 2012, Pajares *et al.*, 2004, Qiu *et al.*, 2005), Artificial Neural Networks integration (Ling *et al.*, 2008, Liu *et al.*, 2008), clusterbased data integration (Alyannezhadi *et al.*, 2016, Shi *et al.*, 2012) and other techniques based on possibility theory (Benferhat *et al.*, 2006, Destercke *et al.*, 2009).

Even though extensive research has been conducted on multi-sensor data integration for fault diagnosis and prognosis, sparse work has addressed data integration of multi-sensor degradation information or unified degradation path. Wei *et al.* (2011) weigh different sensor signals by minimizing the uncertainty of unified measurements, but the data from different sensors are considered as independent instead of correlated with each other. The first passage time distribution of even 2-dimensional Brownian motion with drift is quite complex (Dominé *et al.*, 1993). It is more difficult to obtain explicit reliability functions for more dimensions of more advanced degradation processes.

#### **CHAPTER 3**

# **CORROSION GROWTH MODELING**

#### 3.1 Introduction

In this chapter, we propose two stochastic models to capture the corrosion depth growth and corrosion volume growth respectively. The improved inverse Gaussian (IIG) process is developed as a stochastic degradation model to describe the corrosion depth growth. The model captures dependency between corrosion increments and corrosion depth length and is more robust than the inverse Gaussian (IG) model. Assuming that the corrosion pit growth in the depth direction and radial direction both follow IIG processes, the corrosion pit volume growth model is developed. Moreover, a physics-based stochastic model is proposed to describe the propagation of corrosion pits. The influence of stresses of temperature, relative humidity and pH level are investigated based on chemical reaction mechanisms. The models are validated with actual corrosion data.

# 3.2 Corrosion Depth Growth Modeling

### 3.2.1 Inverse Gaussian Process

Suppose that a degradation process  $\{d(t), t \ge 0\}$  follows an IG process with scale parameter  $\lambda$  and shape (mean) function  $\Lambda(t)$  as described by Ye *et al.* (2014). Let  $I_g(\cdot)$  denote the probability density function (PDF) of IG distribution. It has the following properties:

• d(t) has independent increments:  $d(t_2) - d(t_1)$  and  $d(t_4) - d(t_3)$  are independent of each other for  $\forall t_4 > t_3 > t_2 > t_1$ 

• The degradation increments follow IG distributions:  

$$(d(t_2)-d(t_1)) \sim Ig(\Lambda(t_2)-\Lambda(t_1), \lambda(\Lambda(t_2)-\Lambda(t_1))^2)$$
 for  $\forall t_2 > t_1$ 

Let d(t) denote the degradation state at time t, where  $t \ge 0$ . d(0) is the starting degradation state. When the shape function  $\Lambda(t) = \mu t$ , where  $\mu$  is a constant parameter, the IG process appropriately describes a fatigue crack growth degradation process (Peng *et al.*, 2014). The PDF and cumulative distribution function (CDF) of d(t)-d(0) are represented as follows (Guo *et al.*, 2018)

$$f_{D(t)-D(0)}\left(\left(d(t)-d(0)\right)|\Lambda(t), \ \lambda\Lambda^{2}(t)\right) \\ = \sqrt{\frac{\lambda\Lambda^{2}(t)}{2\pi \left(d(t)-d(0)\right)^{3}}} \exp\left[-\frac{\lambda\left(\left(d(t)-d(0)\right)-\Lambda(t)\right)^{2}}{2\left(d(t)-d(0)\right)}\right]$$
(3.1)
$$= \sqrt{\frac{\lambda\mu^{2}t^{2}}{2\pi \left(d(t)-d(0)\right)^{3}}} \exp\left[-\frac{\lambda\left(\left(d(t)-d(0)\right)-\mu t\right)^{2}}{2\left(d(t)-d(0)\right)}\right], \left(d(t)-d(0)\right) > 0$$

$$F_{D(t)-D(0)}\left(\left(d(t)-d(0)\right)|\Lambda(t),\lambda\Lambda^{2}(t)\right) = \Phi\left[\sqrt{\frac{\lambda\mu^{2}t^{2}}{(d(t)-d(0))}}\left(\frac{(d(t)-d(0))}{\mu t}-1\right)\right] +$$
(3.2)  

$$\exp(2\lambda\mu t)\Phi\left[-\sqrt{\frac{\lambda\mu^{2}t^{2}}{(d(t)-d(0))}}\left(\frac{(d(t)-d(0))}{\mu t}+1\right)\right]$$

The expectation and variance are given by

$$E\left(\left(d\left(t\right)-d\left(0\right)\right)\right) = \Lambda\left(t\right) = \mu t \text{ and } Var\left(\left(d\left(t\right)-d\left(0\right)\right)\right) = \frac{\Lambda\left(t\right)}{\lambda} = \frac{\mu t}{\lambda}$$

### 3.2.2 Improved Inverse Gaussian Process

In this section, we develop an IIG process model based on the IG process. Suppose the degradation process  $\{d(t), t \ge 0\}$  is observed at every discrete unit of time. Assume that at time  $t_j$ , the degradation state is  $d(t_j)$ . The degradation increment  $\Delta d(t_j) = d(t_{j+1}) - d(t_j)$  denotes the degradation during  $(t_j, t_{j+1})$ . Rather than modeling this degradation increment with respect to time  $t_j$ , we use the starting degradation  $d(t_j)$  as the reference. In other words, the degradation in a unit time  $\Delta d(t_j)$  follows an IG distribution in accordance with the starting degradation  $d(t_j)$ . Assume that the shape function takes a linear form as  $\Lambda(d(t_j)) = \mu_0 + \mu_1 d(t_j)$ , where  $\mu_0$  and  $\mu_1$ . are constants, the PDF of  $\Delta d(t_j)$  is

$$f_{\Delta D(t_{j})}\left(\Delta d\left(t_{j}\right) \middle| \Lambda\left(d\left(t_{j}\right)\right), \ \lambda \Lambda^{2}\left(d\left(t_{j}\right)\right)\right)$$

$$= \sqrt{\frac{\lambda \Lambda^{2}\left(d\left(t_{j}\right)\right)}{2\pi\Delta d^{3}\left(t_{j}\right)}} \exp\left[-\frac{\lambda\left(\Delta d\left(t_{j}\right) - \Lambda\left(d\left(t_{j}\right)\right)\right)^{2}}{2\Delta d\left(t_{j}\right)}\right], \ \Delta d\left(t_{j}\right) > 0$$
(3.3)

The expectation and variance are

$$E\left(\Delta d\left(t_{j}\right)\right) = \Lambda\left(d\left(t_{j}\right)\right) = \mu_{0} + \mu_{1}d\left(t_{j}\right) \text{ and } Var\left(\Delta d\left(t_{j}\right)\right) = \frac{\Lambda\left(d\left(t_{j}\right)\right)}{\lambda} = \frac{\mu_{0} + \mu_{1}d\left(t_{j}\right)}{\lambda}$$

## 3.2.3 Remaining Life Prediction and Failure Probability Estimation

Suppose the threshold of the degradation process is  $c_d$ . The distribution of remaining life for the IIG process is obtained through iterative computation. Assume we have degradation measurements until time  $t_{\tau}$  as  $\{d(0), ..., d(t_{\tau})\}$ , the remaining life is obtained as follows

- 1) Starting from j=1.
- 2) Generate a random number  $\Delta d(t_{\tau+j-1})^*$  following the distribution  $Ig(\Lambda(d(t_{\tau+j-1})), \lambda\Lambda^2(d(t_{\tau+j-1})))$ . Degradation state now becomes  $d(t_{\tau+j}) = d(t_{\tau+j-1}) + \Delta d(t_{\tau+j-1})^*$

- 3) Compare  $d(t_{\tau+j})$  with the threshold  $c_d$ . If  $d(t_{\tau+j}) < c_d$ , set j = j+1, go back to step (2). If  $d(t_{\tau+j}) \ge c_d$ , the predicted failure time is  $t_{\tau+j}$  and the predicted remaining life is  $t_{\tau+j} - t_{\tau}$ .
- 4) Repeat steps (1)-(3) for N times, the mean remaining life is obtained by taking the average of the N predicted remaining lives. Let  $T_k$  denote the k percentile of the predicted remaining life. The 90% confidence interval of the predicted remaining life is  $(T_{0.05}, T_{0.95})$ .

The probability that the unit fails in the next time increment given that it survived at time  $t_{\tau}$  with the starting degradation state  $d(t_{\tau})$  is as follows

$$P(T < t_{\tau+1})$$

$$= P(\Delta d(t_{\tau}) + d(t_{\tau}) > c_{d})$$

$$= P(\Delta d(t_{\tau}) > c_{d} - d(t_{\tau})) \qquad (3.4)$$

$$= 1 - \Phi\left(\sqrt{\frac{\lambda}{c_{d} - d(t_{\tau})}} \left(\Lambda(d(t_{\tau})) - c_{d} + d(t_{\tau})\right)\right)$$

$$+ \exp\left(2\lambda\Lambda(d(t_{\tau}))\right) \Phi\left(-\sqrt{\frac{\lambda}{c - d(t_{\tau})}} \left(\Lambda(d(t_{\tau})) + c_{d} - d(t_{\tau})\right)\right)$$



3.2.4.1 Model Validation with Crack Growth Data

Figure 3.1 (a) Degradation versus time (b) degradation versus starting crack length (c) degradation increments versus time (d) degradation increments versus starting crack length.

In this section, we use Bogdanoff and Kozin's data (Bogdanoff *et al.*, 1985) to validate the proposed model. This dataset is characterized as linear degradation paths with equally spaced degradation observations. Twenty-one units are placed under testing for 0.12 million cycles. The crack length for each unit is measured every 0.01 million cycles. The

starting crack length is set as 0.9 inches for all the units to shorten the testing time. The unit is considered to fail when the crack length reaches 1.6 inches. Figure 3.1 (a) is the plot of the 21 degradation paths versus time. Figure 3.1 (b) shows the degradation paths versus crack length. From these two figures, we observe that the degradation rate is linearly related to the degradation states. The variance is also smaller in Figure 3.1 (b). Figure 3.1 (c) and 3.1 (d) are the plots of crack increments for the 21 units versus time and starting crack length respectively. Again, compared with that in Figure 3.1 (c), the lines in Figure 3.1 (d) converge better. This means that starting crack length is a suitable independent variable in this example.

The parameter estimation approach is provided in Appendix A. The crack length data are provided in Appendix B. To fit the IIG model, we reorganize the original data into pairs  $\left[d\left(t_{j}\right), \Delta d\left(t_{j}\right)\right], t_{j} \ge 0$ , where  $d\left(t_{j}\right)$  is the starting crack length at time  $t_{j}$  and  $\Delta d\left(t_{j}\right)$  is the degradation increment in the next 0.01 million cycles. We use the crack growth data of the first twelve units because they represent samples with similar properties. This can be inferred from the estimation results of parameters with the IIG model for the two groups of degradation paths respectively. For the first twelve units,  $\left\{\hat{\mu}_{0}, \hat{\mu}_{1}, \hat{\lambda}\right\} = \{-0.1125, 0.1585, 400.5748\}$ , and the remaining nine units  $\left\{\hat{\mu}_{0}, \hat{\mu}_{1}, \hat{\lambda}\right\} = \{-0.0714, 0.1039, 404.7764\}$ . From Figure 3.2 (a), the 90% probability interval of the predicted degradation path with parameters estimated from the first twelve units deviates from the other nine units.



Figure 3.2 (a) The predicted 90% probability interval and the other nine units (b) The

failure probability curves for IG and IIG

The estimation of parameters  $\{\mu_0, \mu_1, \lambda\}$  for the IIG model of the first twelve paths are

$$\hat{\mu}_0 = -0.1125, \quad \hat{\mu}_1 = 0.1585, \quad \hat{\lambda} = 400.5748$$

Meanwhile, we use the observed data to fit the IG process. The parameters are obtained

using the R codes as shown in Appendix A. The results are as follows:

$$\hat{\mu} = 0.06985, \ \hat{\lambda} = 0.27075$$

Then, by applying the Monte Carlo simulation, we obtain the failure probability plots for the IIG and the IG model, as shown in Figure 3.2 (b). From the figure, we observe that the IIG model fits the observed data better than the IG model. The mean squared error of the failure probability prediction by the two methods are  $MSE_{IIG} = 0.0011$ ,  $MSE_{IG} = 0.0125$ . In this case, the IIG model has better accuracy than the IG model.

# 3.2.4.2 IIG Model Validation with the Corrosion Growth Data

In this section, we apply the IIG model to the corrosion growth data by Soares *et al.* (2009). The data sample size is increased by generating corrosion data with the deterministic model in the paper by Soares *et al.* (2009) and randomizing them with the Brownian motion process with drift, where the mean is the same as that of the corrosion data. The plot of the corrosion growth data is shown in Figure 3.3 (a).



Figure 3.3 (a) The corrosion depth growth data (b) The expectation and 95% confidence

interval with the IIG model

The estimated parameters are obtained as follows

$$\hat{\mu}_0 = 0.09699$$
,  $\hat{\mu}_1 = -0.04769$ ,  $\hat{\lambda} = 209.2$ 

The expectation and the 95% confidence interval of the corrosion depth growth are shown in Figure 3.3 (b). Meanwhile, we use the observed data to fit the IG process. The estimations of the corresponding parameters are:

$$\hat{\mu} = 0.06445, \quad \hat{\lambda} = 91.54410$$

Then, by applying the Monte Carlo simulation, we obtain the failure probability curves for the IIG and the IG model as shown in Figure 3.4. From the figure, we observe that the IIG model fits the observed data better than the IG model. The mean squared errors of the failure probability predictions obtained by the two methods are  $MSE_{IG} = 0.07982$  and  $MSE_{IIG} = 0.01917$ . For this corrosion dataset, the IIG model has better accuracy than the IG model.



Figure 3.4 Failure probability curve for IG and IIG

By comparing the estimated parameters obtained with the IIG model in the two cases, we find that the dependency between degradation increments and degradation states are both captured, which is reflected in the sign of  $\hat{\mu}_1$ . For the crack growth dataset,  $\hat{\mu}_1 = 0.1585$  shows that the crack increments are positively related to initial crack length, while for the corrosion growth dataset,  $\hat{\mu}_1 = -0.04769$  indicates that corrosion growth increments are negatively related to the initial corrosion depth.

# 3.2.5 IIG Model Robustness Evaluation

Model uncertainty is pervasive in reliability estimation. A key question is how robust the model is while the assumptions are not fully satisfied. In this section, a comparison between

the IG process and the IIG process is performed. The results show that the IIG process is more robust than the IG process.

### 3.2.5.1 IIG data

We generate 20 degradation paths based on the IIG model. Suppose that the degradation paths are censored at time t = 7 and time t = 11. In these two cases, the IG and IIG models are utilized to fit the data and obtain the estimation of the model parameters. The two models are then used to predict the failure probability of the units after the censored time. The results are shown in Figure 3.5. It is evident that the IG model does not effectively model the nonlinear growth of the degradation path generated by the IIG model.



Figure 3.5 Use the regular IG model to fit IIG data when (a) data are censored at t = 7 (c) data are censored at t = 11, use the IIG model to fit IIG data when (b) data are

censored at t = 7 (d) data are censored at t = 11.

### 3.2.5.2 IG data

In contrast with the scenario in section 3.2.5.1, 20 degradation paths are generated based on the IG model. Suppose the degradation paths are censored at times t = 8 and t = 12. We fit the data with both the IG and IIG models and obtain the corresponding models' parameters. The predicted failure times of censored units using both models are obtained and shown in Figures 3.6. It is evident that the IIG model effectively predicts the failure time of the censored degradation path generated by the IG model. From the analysis of the two scenarios, we conclude that the IIG model is more robust than the IG model. The IIG process can be applied even when the degradation path does not have nonlinear growth. This is because the IIG process is based on the starting degradation state.



Figure 3.6 Use the IIG model to fit IG data when (a) data are censored at t = 12 (b) data are censored at t = 8, use the IG model to fit IG data when (c) data are censored at t = 12

(d) data are censored at t = 8.

#### 3.3 Corrosion Volume Growth Modeling

In the corrosion reaction process, the produced corrosion layer prevents the atmosphere from contacting new metal, inhibits the transport of reactant and decreases growth rate (Ghahari *et al.*, 2011, Li *et al.*, 2009, Vanaei *et al.*, 2017). Accordingly, the corrosion growth rate decreases as corrosion proceeds. We let the corrosion growth rates in the radial and the depth directions be dependent on the accumulated corrosion states in the model. As shown in Figure 3.7 (a), the corrosion pit is assumed to be consisted of several sectors, with the growth of the radius of each sector following an IIG process. We also derived an alternative corrosion volume growth model based on improved Gamma process, interested readers may refer to Appendix C.



Figure 3.7 Schematic diagram of the corrosion pit growth

Meanwhile, as shown in the plan view of the corrosion pit in Figure 3.7 (b), we assume it is divided into many pixels (squares) and that the pixel depth varies from one to another.

The depth growth of each pixel over time follows an IIG process. The dark pixels represent corroded pixels while the gray ones represent newly corroded pixels in the next time increment. As corrosion progresses, the size of the pit grows both radially and in the depth direction. The volume growth of the corrosion pit in the next time increment consists of two parts: volume growth of the corroded pixels and volume growth of the pixels corroded in the next time increment, which are called new corroded pixels.

# 3.3.1 Volume Growth of the Corroded Pixels

We model the depth growth/sector radius growth with an IIG process, where the shape parameter is a function of the corrosion level (pixel depth or sector radius), such that the phenomenon that corrosion growth rate decreases as corrosion increases is characterized. Let  $d_i(t)$  denote the corrosion depth of the corroded pixel *i* at time *t*. The depth of pixel *i* at time *t* is  $d_i(t+1) = \Delta D_i(t) + d_i(t)$ , where  $\Delta D_i(t)$  is the corrosion depth growth of pixel *i* in the time interval (t, t+1). We use  $Ig(\cdot)$  to denote the PDF of IG distribution. The PDF of  $\Delta D_i(t)$  is

$$f_{\Delta D_i(t)}\left(\Delta d_i(t)\right) = Ig\left(\mu_0 + \mu_1 d_i(t), \lambda\left(\mu_0 + \mu_1 d_i(t)\right)^2\right)$$
(3.5)

where parameters  $\mu_0 > 0$ ,  $\mu_1 < 0$ ,  $\lambda > 0$  and they are the same for all pixels. The area of a pixel is  $s_0$ . Suppose the area of existing corroded pixels at time t is s(t), the total

number of corroded pixels at time t is  $n(t) = \frac{s(t)}{s_0}$ . Let  $\Delta V_i(t)$  denote the volume growth of pixel i in the time interval (t, t+1). The overall volume increment of the pit in time interval (t, t+1) by all the corroded pixels is

$$\sum_{i=1}^{n(t)} \Delta V_i(t) = s_0 \sum_{i=1}^{n(t)} \Delta D_i(t)$$
(3.6)

 $\sum_{i=1}^{n(t)} \Delta D_i(t) \text{ is a new IG variable with mean } \sum_{i=1}^{n(t)} (\mu_0 + \mu_1 d_i(t)) \text{ and scale parameter}$  $\lambda \left( \sum_{i=1}^{n(t)} (\mu_0 + \mu_1 d_i(t)) \right)^2 \text{. By properties of IG distribution, } s_0 \sum_{i=1}^{n(t)} \Delta D_i(t) \text{ is an IG variable}$ 

with mean  $\sum_{i=1}^{n(t)} \left( \mu_0 + \mu_1 d_i(t) \right) = s_0 \left( n(t) \mu_0 + \mu_1 \frac{v(t)}{s_0} \right) \quad \text{and} \quad \text{scale parameter}$ 

$$\lambda s_0 \left( n(t) \mu_0 + \mu_1 \frac{v(t)}{s_0} \right)^2$$
, where  $v(t)$  is the volume of the pit at time t.

#### 3.3.2 Volume Growth of the Newly Corroded Pixels



Figure 3.8 Corrosion area growth
The corroded area becomes larger as corrosion pit grows. As shown in Figure 3.8 (a), suppose the pit's area is divided into  $n_s$  sectors. The angle of each sector is  $\xi = \frac{2\pi}{n_s}$ . Suppose at time t the radius of the  $j^{th}$  sector is  $r_j(t)$ . Let the radius growth follow an IIG process with parameters  $\eta_0 > 0$ ,  $\eta_1 < 0$ ,  $\omega > 0$  as

$$\Delta R_j(t) \sim Ig\left(\eta_0 + \eta_1 r_j(t), \ \omega \left(\eta_0 + \eta_1 r_j(t)\right)^2\right)$$
(3.7)

The radii of the sectors grow independently. As shown in Figure 3.8 (b), by simple geometry, the length of the arc is  $2r_j(t)\sin\frac{\xi}{2}$ . Since the sectors are in the same corrosion pit, the parameters are assumed to be the same. All the newly corroded pixels begin to grow in the depth direction by the time they are corroded. When the radius increment is small (growth time increment is small) and the angle is small, the area growth can be approximated with a rectangle as shown in Figure 3.8 (b). The area growth of sector j in (t, t+1) is  $\Delta S_j(t)=2r_j(t)\sin\frac{\xi}{2}\cdot\Delta R_j(t)$ , which follows an IG distribution as shown in Eq. (3.8).

$$\Delta S_{j}(t) \sim Ig\left(2r_{j}(t)\sin\frac{\xi}{2}\cdot\left(\eta_{0}+\eta_{1}r_{j}(t)\right), \ 2r_{j}(t)\sin\frac{\xi}{2}\cdot\omega\left(\eta_{0}+\eta_{1}r_{j}(t)\right)^{2}\right) \quad (3.8)$$

Equivalently, the number of pixels in sector j corroded in (t, t+1) is

$$\Delta N_{j}(t) = \frac{\Delta S_{j}(t)}{s_{0}} \sim Ig \begin{pmatrix} \frac{2r_{j}(t)}{s_{0}} \sin \frac{\xi}{2} \cdot (\eta_{0} + \eta_{1}r_{j}(t)), \\ \frac{2r_{j}(t)}{s_{0}} \sin \frac{\xi}{2} \cdot \omega (\eta_{0} + \eta_{1}r_{j}(t))^{2} \end{pmatrix} (3.9)$$

Correspondingly, the increase of the number of corroded pixels in the pit is  $\Delta N(t) = \sum_{j=1}^{n_t} \Delta N_j(t) \quad \text{. These pixels grow with an initial depth of} \\ \left\{ d_i(t) = 0, \ n(t+1) \le i \le n(t) + \Delta n(t) \right\} \quad \text{Similar to corroded pixels, the corresponding} \\ \text{volume growth of the sum of these pixels is an IG variable with mean} \\ \sum_{i=n(t)+1}^{n(t)+\Delta n(t)} \left( \mu_0 + \mu_1 d_i(t) \right) = \mu_0 \Delta n(t) \text{ and scale parameter } \omega \left( \Delta n(t) \mu_0 \right)^2 \text{.}$ 

#### 3.3.3 Total Incremental Volume

The total incremental volume in the next time increment is the sum of volume increments of corroded pixels and new corroded pixels. Let  $\Gamma(v(t)) = (n(t) + \Delta n(t))\mu_0 + \mu_1 \frac{v(t)}{s_0}$ .

According to the properties of IG distribution, we have

$$\Delta V(t) \sim Ig(s_0 \Gamma(v(t)), \lambda s_0 \Gamma^2(v(t)))$$
(3.10)

we find that the overall volume growth in the next time increment follows IG distribution, where the mean and shape parameters dependent on the number of corroded pixels by time t, the number of pixels corroded in (t, t+1), corrosion volume at time t and pixel area  $s_0$ .

## 3.3.4 Incremental Volume Under Stresses

The corrosion volume growth is significantly affected by the atmospheric conditions, which include temperature, relative humidity (Rh), pH levels and others. According to Hughes *et al.* (2016), the corrosion volume growth rate (metal dissolution rate) is related to corrosion current density. Because corrosion is an electrochemical process, corrosion current arises when there is charge transfer between the anode (metal) and cathode. The corrosion reaction current is approximately described as

$$i_{c} = i_{0} \exp\left[-\frac{\alpha n F\left(E - E_{rev}\right)}{RT}\right]$$
(3.11)

where *n* is the charge on the ion in equivalents/mol, *F* is Faraday constant = 96,487C/equivalent, *R* is the gas constant = 8.314 J/mol-K,  $E_{rev}$  is the reversible potential, *E* is the applied potential,  $\alpha$  is the charge transfer coefficient,  $i_0$  is the exchange current which is a current of a given single electrode at equilibrium when the electrode material experiences no loss or gain. It is a function of soluble species concentration. According to Eyring (1935), the rate of reaction is proportional to the concentration of reagent  $H^+$  and water vapor in the air. Because the concentration  $c(H^+)=10^{-pH}$ , we modify Eq. (3.11) as

$$i_{c} = \beta_{0} e^{\beta_{1}Rh} e^{\beta_{2}pH} \exp\left[\frac{\beta_{3}}{T}\right]$$

$$= \beta_{0} \exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right)$$
(3.12)

where  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  are constants. Eq. (3.10) describes the base corrosion (volume) rate, which corresponds to base corrosion current density (exchange current rate). We incorporate the effect of temperature, *pH* level, and Rh by modifying Eq. (3.10) as

$$\Delta V(t) \sim Ig \begin{pmatrix} s_0 \Gamma(v(t)) \exp\left(\beta_1 Rh + \beta_2 pH + \frac{\beta_3}{T}\right), \\ \lambda s_0 \left(\Gamma(v(t)) \exp\left(\beta_1 Rh + \beta_2 pH + \frac{\beta_3}{T}\right)\right)^2 \end{pmatrix} (3.13)$$

Similarly, the increase in depth under stresses is described as

$$\Delta D_{i}(t) \sim Ig \begin{pmatrix} \left(\mu_{0} + \mu_{1}d_{i}(t)\right)\exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right), \\ \lambda\left(\left(\mu_{0} + \mu_{1}d_{i}(t)\right)\exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right)\right)^{2} \end{pmatrix}$$
(3.14)

#### 3.3.5 Reliability Estimation and Remaining Life Prediction

#### 3.3.5.1 Reliability Estimation

Because failure occurs when either the depth of the corrosion pit reaches the depth threshold or the volume of the corrosion pit reaches the volume threshold, we consider it as a competing risk model. We use  $IG(\cdot)$  to denote the CDF of IG distribution. In terms of the corrosion volume failure mode, the reliability, which is the probability that the component (structure) fails in time t+1 given that it has not failed at time t, is:

$$R_{\nu}(t) = P_{\nu}(T_{f} > (t+1)) = P(\Delta V(t) + \nu(t) < c_{\nu})$$

$$= P(\Delta V(t) < c_{\nu} - \nu(t)) \qquad (3.15)$$

$$= IG \left( c_{\nu} - \nu(t) \middle| s_{0} \Gamma(\nu(t)) \exp\left(\beta_{1}Rh - \beta_{2}pH - \frac{\beta_{3}}{T}\right), \\ \lambda s_{0} \left(\Gamma(\nu(t)) \exp\left(\beta_{1}Rh - \beta_{2}pH - \frac{\beta_{3}}{T}\right)\right)^{2} \right)$$

where  $T_f$  is the failure time,  $c_v$  is the threshold for corrosion volume growth failure mode.  $c_v$  can be obtained according to the procedures in Appendix D. Let  $d^*(t)$ represent the maximum depth of a corrosion pit. For the depth failure mode, reliability is given by:

$$R_{d}(t) = P_{d}(T_{f} > (t+1)) = P(\Delta D^{*}(t) + d^{*}(t) < c_{d})$$

$$= P(\Delta D^{*}(t) < c_{d} - d^{*}(t))$$

$$= IG \begin{pmatrix} \mu_{0} + \mu_{1}d_{i}(t) \exp(\beta_{1}Rh - \beta_{2}pH - \frac{\beta_{3}}{T}), \\ \lambda \left( (\mu_{0} + \mu_{1}d_{i}(t)) \exp(\beta_{1}Rh - \beta_{2}pH - \frac{\beta_{3}}{T}) \right)^{2} \end{pmatrix}$$
(3.16)

where  $c_d$  is the threshold for corrosion depth growth failure mode. Assuming the two failure modes are independent, the reliability at time t is

$$R(t) = R_d(t)R_v(t)$$
(3.17)

#### 3.3.5.2 Remaining Life Prediction

The distribution of the remaining life is obtained by using an iterative computational method as described next. Assume we have measurements of a corrosion pit growth (3-D

images) up to time t. The corrosion volume at that time is v(t) and the number of corroded pixels is n(t). The remaining life is predicted as follows:

- (1) Set k = t.
- (2) Generate a random number  $\Delta n(k)$  using Eq. (3.9), update  $\{r_j(k)\}$  with  $\{r_j(k+1)\}$ .
- (3) Generate a random number  $\Delta v(k)$  using Eq. (3.10).
- (4) Obtain  $v(k+1) = v(k) + \Delta v(k)$ . If  $v(k+1) < c_v$ , set k = k+1, go to step (2).

Otherwise, the failure time is k+1 and the remaining life T = k+1-t.

(5) Repeat steps (2) - (4)  $N_p$  times,

Let  $T_m$  denote the *m* percentile remaining life predicted. The obtained 90% confidence interval of predicted remaining life is  $(T_{0.05}, T_{0.95})$ .

#### 3.3.6 Case Study

Because the proposed model needs datasets that contain 3-D data of corrosion pit, which can hardly be found in existing literature, we use corrosion depth growth data to demonstrate the applicability of the model. We use two datasets: corrosion depth growth data of steel by Caleyo *et al.* (2009) and corrosion depth growth data of ships by Soares *et al.* (2009). In the first dataset, the measurements are obtained under only one operating condition. In the second dataset, the corrosion data are obtained under six different operating conditions. The parameters are estimated with MLE. The details of the procedures are provided in Appendix E.

#### 3.3.6.1 Caleyo's Dataset

To illustrate the use of the model, we fit the corrosion depth growth data by Caleyo *et al.* (2009) with the depth growth model as developed in Eq. (3.5) and estimate its parameters. The estimated parameters are  $\{\hat{\mu}_0, \hat{\mu}_1, \hat{\lambda}\} = \{20.78, -0.33, 0.59\}$ . The actual data and predicted mean depth growth path are shown in Figure 3.9. Assuming the corrosion growth in the depth and the radial directions are controlled by the same set of parameters, i.e.,  $\{\mu_0, \mu_1, \lambda\} = \{\eta_0, \eta_1, \omega\}$ , we simulate the corrosion pit 3-D growth data with the estimated parameters as shown in Figure 3.10.



Figure 3.9 Corrosion data and predicted mean corrosion growth path



**Figure 3.10** Selected simulated corrosion pit 3-D images from t = 3 to t = 9

Using the simulated corrosion growth data by time 10 we obtain the volume model parameters as  $\{\hat{\eta}_0, \hat{\eta}_1, \hat{\omega}, \hat{\mu}_0, \hat{\mu}_1, \hat{\lambda}\} = \{20.86, -0.30, 0.62, 21.12, -0.31, 0.58\}$ . Suppose we are considering the metal loss of a pipe with  $\sigma_{flow} = 0.35 Mpa$ ,  $t_p = 63 \mu m$ ,  $D = 100 \ \mu m$ ,  $\underline{L} = 100 \ \mu m$  and the operating pressure of the pipe is  $P_0 = 0.6 \times 0.35 = 0.21 Mpa$ ,  $v_0 = 630,000 \ cubic \ \mu m$  and  $c_v = 504,340 \ cubic \ \mu m$ . The critical depth threshold is  $c_d = 63 \ \mu m$ . With estimated parameters, the expectation and 90% probability intervals of volume growth and maximum depth growth curves are shown in Figure 3.11. We observe that the predicted failure times are significantly different, which shows the necessity of considering the volume loss failure mechanism. The reliability of the two individual failure modes and joint failure modes are plotted in Figure 3.12.



Figure 3.11 Predicted volume growth and maximum depth growth

We compare the proposed model with the physics-based model by Kondo *et al.* (1989) and the power-law mass loss function by Panchenko *et al.* (2016) using Caleyo's dataset. As shown in Figure 3.13, the proposed model matches well with the physics-based model and the power-law empirical model. We find that in the late stage of the corrosion growth process, the physics-based model overpredicts the volume growth compared with the proposed model, which predicts that the corrosion volume growth rate decreases. This can be explained as that the physics-based model does not consider the fact that the recently formed corrosion layers prevent new metal from contacting the external environment. However, these layers may lead to the decrease of the corrosion growth rate.



Figure 3.12 Reliability when considering different failure modes



Figure 3.13 Corrosion loss prediction comparison using different models

# 3.3.6.2 Soares's Dataset

The model is also validated with corrosion depth growth data from six ships operating on different routes on the Pacific Ocean by Soares *et al.* (2009). The data sample size is increased by generating corrosion data with the deterministic model by Soares *et al.* (2009) and randomized with the Brownian motion process with drift, where the mean of the Brownian motion process is the same as that of the corrosion data. We censor the data at time 10 and estimate the parameters of the model. The predicted corrosion depth growth after time 10 and time to failure are shown in Figure 3.14 (a). The sum of squared errors (SSE) of the predicted mean path and the percent error (PE) of failure times are calculated

and shown in Table 3.1. Let  $t_0$  denote the censoring time. SSE and PE are defined as follows

$$SSE = \sum_{j=t_0+1}^{T_f} \left( d^*(j) - \hat{d}^*(j) \right)^2$$
$$PE = \frac{T_f - T_r}{T}$$

where  $d^*(j)$  is the maximum depth of the pit at time j.  $\hat{d}^*(j)$  is the predicted mean maximum depth of the pit at j.  $T_f$  is the observed failure time and  $T_r$  is the true failure time. The predicted corrosion volume growth paths are shown in Figure 3.14 (b). The reliability function considering both the corrosion depth growth and volume growth is shown in Figure 3.15 when maximum depth is 1.2 while the volume threshold is 2.0.



Figure 3.14 Corrosion depth and volume prediction when data are censored at t = 10

Ship index	1	2	3	4	5	6
T (Kelvin)	283.85	289.75	291.55	293.35	293.55	296.05
Rh(%)	85.5	82.8	83.3	79.8	82.3	81.6
Ph	8.16	8.15	8.15	8.14	8.18	8.18
PE of $T$	1.685	1.163	-1.625	3.870	-0.132	-1.644
SSE	0.00067	0.00130	0.00133	0.00406	0.00056	0.00063

Table 3.1 The time to failure prediction accuracy and sum of squared errors



Figure 3.15 Reliability when considering both depth growth and volume growth

# 3.4 Corrosion Propagation Modeling

In the corrosion degradation process, the corrosion pits not only grow larger but also propagate to other locations on the material over time, leading to the initiation and growth of new pits. In this section, we develop a physics-based stochastic model that describes the propagation of corrosion pits in both time and space. As the initiation of pits is related to material composition and the existing pits' locations, we consider three aspects: characterization of particles in a material, modeling of pits initiation and modeling of multiple pits growth, propagation and overlap/coalescence.

## 3.4.1 Characterization of the Particles

It is widely known that corrosion pits initiate from the particles in the alloy (Cawley *et al.*, 1996). According to Liao (1998), Cullin *et al.* (2011) and Cawley *et al.* (1996), the particles in the alloy tend to aggregate and form clusters as shown in Figure 3.16 (a). It is observed that the pits induced from larger particle clusters are larger than those induced by smaller-size particles, as shown in Figure 3.16 (b). The area/size distribution of the particles in the LT (Rolling Transverse), LS (Rolling-Short Transverse) planes of a material can be characterized by a bi-mode Weibull distribution (Harlow, 2012) as shown in Figure 3.17, where the CDF is the product of the CDFs representing the two modes as shown in Eq. (3.18)

$$F_{s}(x) = F_{1s}(x)F_{2s}(x)$$
(3.18)

The indexes  $1_s$  and  $2_s$  indicate the modes associated with the lower and upper tails of the data respectively. Using a three-parameter Weibull distribution, the CDF of the particle sizes is

$$F_{s}(x) = 1 - \exp\left[-\left(\frac{x}{\beta_{1}}\right)^{\alpha_{1}}\right] - \exp\left[-\left(\frac{x}{\beta_{2}}\right)^{\alpha_{2}}\right] + \exp\left[-\left(\frac{x}{\beta_{1}}\right)^{\alpha_{1}} - \left(\frac{x}{\beta_{2}}\right)^{\alpha_{2}}\right] \quad (3.19)$$

where  $\alpha_1$  and  $\beta_1$  are the parameters for the first Weibull distribution.  $\alpha_2$  and  $\beta_2$  are the parameters for the second Weibull distribution.



Figure 3.16 Particles and pits (Liao et al., 1998)

Cawley *et al.* (1996) state that the particle centers are uniformly distributed on the surface. We incorporate both particles' sizes and centers in modeling the pits propagation in our model.



Figure 3.17 Area data for particles on the LS and LT surfaces of common 7075-T6 aluminum alloy with estimated CDFs (Harlow, 2012)

# 3.4.2 Corrosion Pits Initiation

A particle's ability to initiate a pit depends on its size (Suter *et al.*, 2001). Meanwhile, the particles close to the corroded pits are more susceptible to corrosion because of the chemical reaction of the pits' corrosion. Let  $PP(P_i, t)$  denote particle  $P_i$ 's pitting potential, which is its ability to initiate a new pit at time t. We obtain  $PP(P_i, t)$  as follows:

$$PP(P_i, t) = \left(g_1 S(P_i) + g_2 \sum_{j=1}^{n_p(t)} L_j(P_i, t)\right)$$
(3.20)

where  $g_1$  and  $g_2$  are constants.  $P_i$  denotes the index of a particle,  $S(P_i) = 2\pi r_{P_i}^2$  is the effective surface area of particle  $P_i$ , where  $r_{P_i}$  is the radius of particle  $P_i$ . The total number of pits at time t is denoted as  $n_p(t)$ .  $L_j(P_i, t)$  represents the corrosion effects

of the chemical reaction produced by the  $j^{th}$  pit on particle  $P_i$ 's ability to initiate a new pit. We define  $L_i(P_i, t)$  as follows

$$L_{j}(P_{i}, t) = v_{j}(t) \exp\left[-\frac{(x_{P_{i}} - x_{j})^{2} + (y_{P_{i}} - y_{j})^{2}}{2\sigma_{L}^{2}}\right]$$
(3.21)

where  $v_j(t)$  is the volume of the  $j^{th}$  pit at time t,  $x_j$  and  $y_j$  are the coordinates of the  $j^{th}$  pit's center (which are also the coordinates of the center of the particle that has initiated the  $j^{th}$  pit),  $x_{P_i}$  and  $y_{P_i}$  are the coordinates of particle  $P_i$  and  $\sigma_L$  is a constant. As time t increases, whenever  $PP(P_i, t)$  reaches a threshold  $c_{pp}$ , particle  $P_i$  initiates a pit. Eq. (3.20) indicates that every corrosion pit influences the pitting ability of the uncorroded particles. The amount of influence depends on the distances of the particle to the corresponding pit.

Without loss of generality, consider a specimen with a size of  $a_s \times b_s$ . To model the pits initiation process, we propose the following procedures:

- 1) Randomly generate  $n_{pc}$  pairs of coordinates for the  $n_{pc}$  particles, where the x and y coordinates follow uniform distributions  $U(0, a_s)$  and  $U(0, b_s)$ .
- 2) For each of the particle, randomly generate a number z which follows the distribution with CDF  $F_s(z)$  as its size. Set the starting time as t = 0.

3) Calculate the pitting potentials of all the particles according to Eq. (3.20). If ∃i such that PP(Pi, t) > cpp, particle Pi is corroded at time t+1. Let t=t+1, repeat step (3) until all the particles are corroded.

# 3.4.3 Corrosion Growth of Multiple Pits



Figure 3.18 The overlap of the pits

After new pits are initiated, they grow both in the depth direction and the radial directions as described in section 3.3. In the multiple-pits growth process, pits may overlap and grow jointly. Figure 3.18 shows a schematic diagram of the growth and overlap process of four pits, where their centers are  $O_1$ ,  $O_2$ ,  $O_3$  and  $O_4$ . At time  $t_1$ , the pits are away from each other. Starting from  $t_2$ , pits begin to overlap with each other until time  $t_4$ , when the four pits merge into a larger-sized pit that continues to grow until the overall volume reaches the failure threshold.



**Figure 3.19** The coalescence of the two pits at  $t_2$ 

The overlap of the pits can be considered as the overlap of a series of sectors in each pit. Figure 3.19 (a) is an enlargement of the overlap section between pits 1 and 2 in Figure 3.18 (b). We define a rule to decide if the sectors in one pit overlap with their counterparts in another pit and stop growing. As shown in Figure 3.19 (a), the sectors of pit 1 on one side of the overlapped area are  $S_{1j_1}$ ,  $S_{1j_1+1}$ ,  $S_{1j_1+2}$ , the sectors of pit 2 on the corresponding side are  $S_{2j_2}$ ,  $S_{2j_2+1}$ ,  $S_{2j_2+2}$ . Let  $\alpha_{1j_1}$  denote the angle between  $O_1O_2$  and the median of sector  $S_{1j_1}$ .  $r_{ij}(t)$  is the radius of the  $j^{th}$  sector of pit *i* at time *t*, which follows an IIG process. Using simple geometry as shown in Figure 3.19 (b), we observe that sector  $S_{1j_1}$  stops growing at time *t* if there is at least a sector  $S_{2j_1}$ ,  $j_2 \le j \le j_2 + \frac{n_s}{4}$  such that

$$r_{1j_{1}}(t)\cos(\alpha_{1j_{1}}) + r_{2j}(t)\cos(\alpha_{2j}) > O_{1}O_{2}$$
(3.22)

We define an indicator function as shown in Eq. (3.23) to indicate whether a sector from pit 1 stops growing at time t.

$$I_{1j_{1}}(t) = \operatorname{sgn}\left\{r_{1j_{1}}(t)\cos(\alpha_{1j_{1}}) + r_{2j}(t)\cos(\alpha_{2j}) - O_{1}O_{2}\right\}$$
(3.23)

where

$$\operatorname{sgn}(x) = \begin{cases} -1 \text{ if } x < 0\\ 1 \text{ if } x > 0 \end{cases}$$

If  $I_{1j_i}(t) = -1$ , the corresponding sector overlaps with other sectors and stops growing. Otherwise, it continues to grow. Let  $t_{ij}$  denote the overlap time of sector  $S_{ij}$ , we define  $I_{ij}(t) = -1$  for  $t \ge t_{ij}$ . The volume growth of all the pits is composed of two parts: the volume growth of corroded pixels and newly corroded pixels as discussed in the following section.

## 3.4.3.1 The Depth and Volume Growth of Corroded Pixels

Assume at time t there are  $n_p(t)$  pits. Let  $d_{ik}(t)$  denote the corrosion depth of the corroded pixel k in the  $i^{th}$  pit at time t. The depth of pixel k in the  $i^{th}$  pit at time t+1 is  $D_{ik}(t+1) = \Delta D_{ik}(t) + d_{ik}(t)$ , where  $\Delta D_{ik}(t)$  is the corrosion depth growth of pixel k in the  $i^{th}$  pit in the time interval (t, t+1). The PDF of  $\Delta D_{ik}(t)$  is

$$f_{\Delta D_{ik}(t)}(\Delta d_{ik}(t)) = Ig(\mu_{i0}(t) + \mu_{i1}(t)d_{ik}(t), \lambda(\mu_{i0}(t) + \mu_{i1}(t)d_{ik}(t))^{2}) \quad (3.24)$$

where  $\mu_{i0}(t)$ ,  $\mu_{i1}(t)$  are defined in Eqs. (3.25) and (3.26) as

$$\mu_{i0}(t) = \mu_0 S_i(t) = \mu_0 \sum_{j=1}^{np_i(t)} 2\pi r_{P_j}^2$$
(3.25)

$$\mu_{i1}(t) = \mu_1 S_i(t) = \mu_1 \sum_{j=1}^{n_{P_i}(t)} 2\pi r_{P_j}^2$$
(3.26)

where  $S_i(t)$  is the sum of the surface areas of the particles pit *i* at time *t*,  $\mu_0$ ,  $\mu_1$ and  $\lambda$  are constants and  $np_i(t)$  is the total number of pixels in pit *i* at time *t*. Suppose that the area of all the existing corroded pixels in the  $i^{th}$  pit at time *t* is  $s_i(t)$ , the total number of corroded pixels at time *t* is  $n_i(t) = \frac{s_i(t)}{s_0}$ , where  $s_0$  is the area of a pixel. Let  $\Delta V_{ik}(t)$  denote the volume increase of pixel *k* in the  $i^{th}$  pit in the time interval (t,t+1). The overall volume increment of the pit in the time interval (t,t+1)

$$\sum_{i=1}^{n_{p}(t)} \sum_{k=1}^{n_{i}(t)} \Delta V_{ik}\left(t\right) = s_{0} \sum_{i=1}^{n_{p}(t)} \sum_{k=1}^{n_{i}(t)} \Delta D_{ik}\left(t\right)$$
(3.27)

where  $n_p(t)$  is the total number of pits at time *t*. As  $\Delta D_{ik}(t)$  is an IG variable with shape parameter  $\mu_{i0}(t) + \mu_{i1}(t)d_{ik}(t)$  and scale parameter  $\lambda (\mu_{i0}(t) + \mu_{i1}(t)d_{ik}(t))^2$ , according to properties of IG distribution we have

$$\sum_{i=1}^{n_{p}(t)}\sum_{k=1}^{n_{i}(t)}\Delta V_{ik}(t) \sim Ig\begin{pmatrix} s_{0}\left(\sum_{i=0}^{n_{p}(t)}\mu_{i0}(t)n_{i}(t)+\frac{\sum_{i=0}^{n_{p}(t)}\mu_{i1}(t)v_{i}(t)}{s_{0}}\right),\\ \lambda s_{0}\left(\sum_{i=0}^{n_{p}(t)}\mu_{i0}(t)n_{i}(t)+\frac{\sum_{i=0}^{n_{p}(t)}\mu_{i1}(t)v_{i}(t)}{s_{0}}\right)^{2}\right)$$
(3.28)

where  $v_i(t)$  is the volume of pit *i* at time *t*.

# 3.4.3.2 Volume Growth of Newly Corroded Pixels

The radius of each sector is  $r_{ij}(t)$ , where  $1 \le i \le n_p(t)$ ,  $1 \le j \le n_s$ . The radial growth of

the 
$$j^{th}$$
 sector in pit *i* is  $\Delta R_{ij}(t) \left( \frac{I_{ij}(t) + 1}{2} \right)$ , where  

$$\Delta R_{ij}(t) \sim Ga \left( \eta_{i0}(t) + \eta_{i1}(t) r_{ij}(t), \ \omega \left( \eta_{i0}(t) + \eta_{i1}(t) r_{ij}(t) \right)^2 \right)$$
(3.29)

where  $\eta_{i0}(t)$ ,  $\eta_{i1}(t)$  are defined in Eqs. (3.30) and (3.31).

$$\eta_{i0}(t) = \eta_0 S_i(t) = \eta_0 \sum_{j=1}^{np_i(t)} 2\pi r_{P_j}^2$$
(3.30)

$$\eta_{i1}(t) = \eta_1 S_i(t) = \eta_1 \sum_{j=1}^{n_{P_i}(t)} 2\pi r_{P_j}^2$$
(3.31)

where  $\eta_0$ ,  $\eta_1$  and  $\omega$  are constants,  $S_i(t)$  is the sum of the surface areas of the particles pit *i* at time *t*,  $np_i(t)$  is the total number of pixels in pit *i* at time *t*. The

area growth of the  $j^{th}$  sector in the  $i^{th}$  pit is  $\Delta S_{ij}(t) \left( \frac{I_{ij}(t) + 1}{2} \right)$ , where

$$\Delta S_{ij}(t) \sim Ig \begin{pmatrix} 2r_{ij}(t)\sin\frac{\xi}{2}(\eta_{i0}(t) + \eta_{i1}(t)r_{ij}(t)), \\ 2r_{ij}(t)\sin\frac{\xi}{2}\cdot\omega(\eta_{i0}(t) + \eta_{i1}(t)r_{ij}(t))^2 \end{pmatrix}$$
(3.32)

where  $\xi$  is the contained angle of each sector and  $\Delta S_{ij}(t)$  is the area growth of sector j in pit *i* in (t,t+1) if it continues to grow. Equivalently, the number of pixels in

sector  $\dot{j}$  corroded in (t, t+1) is  $\Delta N_{ij}(t) \left(\frac{I_{ij}(t)+1}{2}\right)$ , where

$$\Delta N_{ij}(t) = \frac{\Delta S_{ij}(t)}{s_0} \sim Ig \begin{pmatrix} 2\frac{r_{ij}(t)}{s_0} \sin\frac{\xi}{2} (\eta_{i0}(t) + \eta_{i1}(t)r_{ij}(t)), \\ 2\frac{r_{ij}(t)}{s_0} \sin\frac{\xi}{2} \cdot \omega (\eta_{i0}(t) + \eta_{i1}(t)r_{ij}(t))^2 \end{pmatrix} (3.33)$$

The increase of the number of new corroded pixels in pit i is

$$\Delta N_i(t) = \sum_{j=1}^{n_s} \Delta N_{ij}(t) \left(\frac{I_{ij}(t) + 1}{2}\right)$$
(3.34)

The volume growth by all these new corroded pixels is

$$\sum_{i=1}^{n_{p}(t)} \sum_{k=\Delta n_{i}(t)+1}^{\Delta n_{i}(t)+n_{i}(t)} \Delta V_{ik}(t) = s_{0} \sum_{i=1}^{n_{p}(t)} \sum_{k=\Delta n_{i}(t)+1}^{\Delta n_{i}(t)+n_{i}(t)} \Delta D_{ik}(t) \sim \left( s_{0} \left( \sum_{i=0}^{n_{p}(t)} \mu_{i0}(t) \Delta n_{i}(t) \right), \ \lambda s_{0} \left( \sum_{i=0}^{n_{p}(t)} \mu_{i0}(t) \Delta n_{i}(t) \right)^{2} \right)$$
(3.35)

#### 3.4.3.3 Total Incremental Volume

The total incremental volume in the next time instant is the sum of the volume increments of corroded pixels and new corroded pixels. The overall volume increment of all the pits in (t, t+1) is

$$\Delta V(t) = \sum_{i=1}^{n_p(t)} \sum_{k=1}^{n_i(t)} \Delta V_{ik}(t) + \sum_{i=1}^{n_p(t)} \sum_{k=1+n_i(t)}^{\Delta n_i(t)+n_i(t)} \Delta V_{ik}(t)$$
(3.36)

Using properties of IG distribution, we have

$$\Delta V(t) \sim Ig \left( s_{0} \left( \sum_{i=0}^{n_{p}(t)} \mu_{i0}(t) (n_{i}(t) + \Delta n_{i}(t)) + \frac{\sum_{i=0}^{n_{p}(t)} \mu_{i1}(t) v_{i}(t)}{s_{0}} \right), \\ s_{0} \lambda \left( \sum_{i=0}^{n_{p}(t)} \mu_{i0}(t) (n_{i}(t) + \Delta n_{i}(t)) + \frac{\sum_{i=0}^{n_{p}(t)} \mu_{i1}(t) v_{i}(t)}{s_{0}} \right)^{2} \right)$$
(3.37)

where  $n_p(t)$  is the total number of pits at time t,  $n_i(t)$  is the total number of pixels in pit i at time t and  $\Delta n_i(t)$  is the number increase of pixels in pit i in (t, t+1), which is predicted by Eq. (3.34). The parameters  $\mu_{i0}(t)$ ,  $\mu_{i1}(t)$  are obtained by Eqs. (3.25) and (3.26) and  $\lambda$  is a constant. In terms of the corrosion volume failure mode, the reliability is:

$$R_{\nu}(t) = P(T_{f} > (t+1)) = P(\Delta V(t) + \nu(t) < c_{\nu})$$

$$= P(\Delta V(t) < c_{\nu} - \nu(t))$$

$$= IG\left(c_{\nu} - \nu(t)\right) \left(s_{0}\left(\sum_{i=0}^{n_{p}(t)} \mu_{i0}(t)(n_{i}(t) + \Delta n_{i}(t)) + \frac{\sum_{i=0}^{n_{p}(t)} \mu_{i1}(t)\nu_{i}(t)}{s_{0}}\right), \\ s_{0}\lambda\left(\sum_{i=0}^{n_{p}(t)} \mu_{i0}(t)(n_{i}(t) + \Delta n_{i}(t)) + \frac{\sum_{i=0}^{n_{p}(t)} \mu_{i1}(t)\nu_{i}(t)}{s_{0}}\right)^{2}\right) (3.38)$$

where  $T_f$  is the failure time,  $c_v$  is the volume threshold and  $IG(\cdot)$  is the CDF of IG distribution.

# 3.4.4 Multiple Pits Initiation and Growth: Case Study

We use a case study to illustrate the pits' initiation and growth process over time. The total number of particles is assumed to be 10. We follow Harlow (2012) and use the parameters of the CDF of particle sizes as  $\alpha_1 = 0.34$ ,  $\beta_1 = 850$ ,  $\alpha_2 = 2$ ,  $\beta_2 = 2$ . The particles are generated as shown in Figure 3.20 with red circles. Figure 3.21 shows the selected figures of overlap of the pits over time, where the particles are also shown in red circles. The total volume growth of the pits is shown in Figure 3.22 (a). The reliability of the unit is shown in Figure 3.22 (b).



Figure 3.20 The particles in a material



Figure 3.21 Selected figures of the pits over time



**Figure 3.22** (a) Total volume growth of the pits (b) Reliability of the unit with multiple

pits for threshold

# 3.5 Summary

We propose an IIG process model to capture the dependency between the corrosion state and the corrosion growth rate. The corresponding failure probability and remaining life are accurately predicted based on the IIG model. The proposed models provide more accurate and robust results than the IG model. Based on the IIG process, we develop the first stochastic model that describes the corrosion pit volume growth over time. It captures the phenomenon where a critical amount of volume loss leads to the failure of a component. An iterative remaining life prediction approach is proposed. The consistency of the proposed model with the empirical mass loss model and physics-based model is verified with real data. The propagation of corrosion is studied. The physical factors, including the spatial and size distributions of the particles and the influence of the corroded pits are incorporated into the pits' initiation model. The pits overlap process is also studied and a simulation model is provided to validate the analytical model.

#### **CHAPTER 4**

#### **DEGRADATION BRANCHING MODELING**

#### 4.1 Introduction

Degradation branching is a common phenomenon in many real-life applications. The degradation of a location not only increases with time but also propagates to other locations in the same system. While the degradation of an individual location has been studied extensively, research on degradation branching is sparse. In this chapter, we develop a general stochastic degradation branching model that characterizes both the degradation growth and degradation propagation. The probabilistic properties of the general degradation branching processes are analyzed. Reliability metrics such as the mean time to failure, mean residual life, failure probability and others are also investigated. In particular, closed-form expressions for the expectation and variance of the degradation and selected reliability metrics are obtained when the time to branch follows an exponential distribution. The model is validated using actual crack growth data and can be applied to corrosion propagation.

The degradation branching process (DBP) is described as follows: (1) The degradation starts in one location and propagates to other locations, (2) After a random time another degradation branch is initiated, (3) The degradation of each branch follows a stochastic

process, (4) Failure occurs when the total degradation of all the branches reaches a failure threshold. Consider the crack growth as an example, the crack is initiated due to the mechanical stresses and continues to grow. After a random time, it initiates another branch (it occurs due to internal stresses or crossing of grain boundaries). Failure occurs when the total crack length of all the branches exceeds the overall "failure" threshold (Wang *et al.*, 2020b).

# 4.2 The Degradation Branching Process Model

#### 4.2.1 Properties of the DBP

First, we define the general DBP as shown in Figure 4.1. Let the stochastic process  $X_0(t)$  denote the degradation of the first location which starts to degrade at time t = 0. The branching occurs after a random time interval  $\tau_1$  and the main degradation path  $X_0(t)$  continues to grow after the new degradation branch is initiated. Note that  $\tau_i$  may have different interpretations in different DBPs depending on the properties of the real-life applications. We define two types of DBPs: Type I DBP and Type II DBP. In Type I DBP,  $\tau_i$  is a variable corresponding to the time that  $X_{i-1}(t)$  reaches its threshold  $c_i$  (the first passage time) and incurs branching while in Type II DBP  $\tau_i$  is a random variable determined by other physical processes independent of the degradation of the branches. Figure 4.1 illustrates the Type I DBP. The degradation of the first branch at time t is

denoted as  $X_1(t-\tau_1)$ , where  $t-\tau_1$  is the length of time of its degradation. Meanwhile, the first degradation branch continues to grow and initiates another branch after a random time  $\tau_2$ , where  $\tau_2$  is a random variable corresponding to the time that  $X_1(t-\tau_1)$ reaches its threshold  $c_1$ . The degradation of the second branch at time t is  $X_2(t-\tau_1-\tau_2)$ , where  $t-\tau_1-\tau_2$  is the length of time of its degradation. In general, the threshold  $c_i$  for each branch to initiate a new branch may be different. The branching continues until the cumulative degradation of all the branches exceeds the overall failure threshold c. In comparison, in Type II DBP the initiations of new branches are independent of the degradation of the original branches. Let  $X_0(t)$  denote the length of the initial degradation and  $X_1(t-\tau_1)$  denote the length of the first branch (say crack length as an example) at time t. The notation  $\tau_1$  denotes the time that the Stress Intensity Factor (SIF) (Sundaram *et al.*, 2018) reaches the threshold  $s_0$  and leads to the initiation of the first crack branch. The notation  $X_2(t-\tau_1-\tau_2)$  denotes the crack length of the second branch at time t and  $\tau_2$  is the time that SIF reaches the threshold  $s_1$  and initiates the second branch. The overall threshold c corresponds to the maximum allowance of total crack length (Sundaram et al., 2018).



Figure 4.1 Schematic diagram of the DBP

Suppose  $\tau_i$  follows a distribution with a probability density function (PDF)  $f(\tau_i)$  and a cumulative distribution function (CDF)  $F(\tau_i)$ . Let  $X_i^*(t)$  denote the degradation of the  $i^{th}$  branch at time t and  $X_0(t) = X_0^*(t)$ . Note that  $X_i^*(t)$ ,  $i \ge 1$  are different from  $X_i\left(t - \sum_{j=1}^i \tau_j\right)$ ,  $i \ge 1$  in that the latter are the conditional degradation given  $\tau_j$ ,  $1 \le j \le i$ are known. Let  $h_i\left(x \left| t - \sum_{j=1}^i \tau_j \right.\right)$  denote the conditional PDF of  $X_i\left(t - \sum_{j=1}^i \tau_j\right)$ , where  $\{\tau_i, 1 \le i \le n\}$  are the known initiation times of the first n branches. Let  $h\left(x \left| t, \{\tau_i, 1 \le i \le n\}\right)$  denote the conditional PDF of these branches' total degradation including the first degraded location, which is  $\sum_{i=1}^n X_i\left(t - \sum_{j=1}^i \tau_j\right) + X_0(t)$ . For example, in the case of crack growth and branching where the degeneration of the branches follows

Brownian Motion with a drift vector  $\boldsymbol{\mu} = (\mu_0, ..., \mu_n)'$  and diffusion vector

$$\boldsymbol{\sigma} = \left(\sigma_{0}, \dots, \sigma_{n}\right)^{\prime} \text{ and } \left\{\tau_{i}, 1 \leq i \leq n\right\} \text{ are known, the conditional total crack length is}$$

$$\sum_{i=1}^{n} X_{i} \left(t - \sum_{j=1}^{i} \tau_{j}\right) + X_{0}\left(t\right) \sim N\left(\mu_{0}t + \sum_{i=1}^{n} \left(t - \sum_{j=1}^{i} \tau_{j}\right)\mu_{i}, \sigma_{0}^{2}t + \sum_{i=1}^{n} \left(t - \sum_{j=1}^{i} \tau_{j}\right)\sigma_{i}^{2}\right) \text{ . As}$$

shown in the proof in Appendix F, the PDF of the total degradation at time t is given by Eq. (4.1),

$$h^{*}(x|t) = \sum_{n=0}^{n_{b}} \int_{0}^{t} \cdots \int_{0}^{t-1} \prod_{i=1}^{\tau_{i}} f(\tau_{i}) \left(1 - F\left(t - \sum_{i=1}^{n} \tau_{i}\right)\right) h(x|t, \{\tau_{i}, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_{i} \quad (4.1)$$

where  $n_b$  is the limit of the number of branches, which can be either finite or infinite depending on the real-life applications. When the system consists of a finite number of units and the degradation of each unit (branch) is represented by a unique indicator,  $n_b$  is finite and equals the total number of units in the system. The degradation branches correspond to the degradation of the units. Otherwise, when there is only one unit and the degradation starts from one location of the unit and propagates and branches to other locations, then  $n_b$  is infinite ( $n_b = \infty$ ) as described in the case of crack growth and branching process. Here the degradation branches correspond to the length of the crack branches.

When  $n_b = \infty$  and the distribution of time to branch  $f(\tau_i)$  is known, the expected number of branches at time t can be determined as follows. Let  $f^*(s)$  be the Laplace

transform of f(t). Let M(t) denote the expected number of branches at time t. Let

 $m(t) = \frac{dM(t)}{dt}$ . As shown in the proof in Appendix F, m(t) is obtained as

$$m(t) = \mathcal{L}^{-1}\left(\frac{f^*(s)}{1 - f^*(s)}\right)$$

The expected number of branches at time t is

$$E(N|t) = M(t)$$

$$= \int_{0}^{t} m(\tau) d\tau$$
(4.2)

The expectation of the total amount of degradation of all the branches including the first degraded location at time t is

$$E\left(X^{*}(t)\right) = \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} \left(g\left(n|\tau\right) \left(\sum_{i=0}^{n} E\left(dX_{i}^{*}(\tau)\right)\right)\right)\right)$$
(4.3)

where  $E(dX_i^*(\tau))$  is the expectation of the degradation increment of the  $i^{th}$  branch in  $(\tau, \tau + d\tau)$  and  $g(n|\tau)$  is the probability that there are n branches at time t. Note that  $dX_i^*(\tau)$  is not the differential form of  $X_i^*(\tau)$  as many stochastic processes are not differentiable, instead, it represents the amount of degradation in the small time interval of  $(\tau, \tau + d\tau)$ . When the expected degradation growth rates of all the branches are  $\mu$  as in the cases of Brownian Motion, Gamma processes and IG processes where the mean function is  $\mu t$ , then Eq. (4.3) reduces to

$$E(X^*(t)) = \mu\left(\int_{\tau=0}^{t} E(N|\tau)d\tau + t\right)$$
(4.4)

where  $E(N|\tau)$  is the expected number of branches at time  $\tau$ . Proofs of Eqs. (4.3) and (4.4) are given in Appendix F. When the failure threshold of the total degradation is c, the mean time to failure (MTTF) can be obtained by solving  $E(X^*(t)) = c$ . In the example of crack growth and branching, the MTTF corresponds to the expected time when the total length of all the cracks reaches the threshold and leads to failure.

Suppose at time t the total degradation is x(t), the total number of branches is n(t)and the last branch is initiated at  $t_1$ . We obtain the mean residual life (MRL) using Eq. (4.5) and its proof is given in Appendix F.

$$MRL = \left\{ t_{r} \int_{\tau=t}^{t_{r}} \sum_{i=0}^{n(t)} E(dX_{i}^{*}(\tau)) + \int_{\tau=t_{r}}^{t_{r}-t_{l}} \int_{j=1}^{t_{r}-t_{l}} \int_{j=1}^{t_{r}-t_$$

The variance of the total degradation at time t is obtained in Eq. (4.6), the proof is in Appendix F. Consider the example of crack growth and branching, Eq. (4.6) obtains the uncertainty of the total length of the crack branches at time t which originates from the uncertainty of the total number of crack branches and the uncertainty of the length of the crack branches.

$$Var(X^{*}(t)) = \sum_{n=1}^{n_{b}} \left( g_{n} \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right)^{2} \right) - \left( \sum_{n=1}^{n_{b}} \left( g_{n} \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right) \right)^{2} + v_{0} + \sum_{n=1}^{n_{b}} \left( g_{n} \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} \left( v_{i}^{n} + \left( m_{i}^{n} \right)^{2} \right) \right) \right) - \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right)^{2} + 2\sum_{\substack{i_{1} \geq 1 \\ i_{1} < i_{2}}}^{n} \left( \dot{F}_{i_{2}} \left( \frac{g_{n-i_{2}}}{g_{n}} m_{i_{1}}^{n} m_{i_{2}}^{n} \right) \right) \right) \right) \right)$$

$$(4.6)$$

where 
$$\dot{F}_i(x) = \int_0^t \cdots \int_0^{i-1} \prod_{j=1}^i f(\tau_j) x \prod_{j=1}^i d\tau_j$$
,  $g_{n-i} = g\left(n-i\left|t-\sum_{j=1}^i \tau_j\right|\right)$ ,  $g_n = g\left(n\left|t\right)$ ,  
 $m_0 = Var\left(X_0(t)\right)$ ,  $m_i^n = E\left(X_i\left(t-\sum_{j=1}^i \tau_j\left|n\right|\right)\right)$  and  $v_i^n = Var\left(X_i\left(t-\sum_{j=1}^i \tau_j\left|n\right|\right)\right)$ .

The reliability and the failure time distribution of the system for a given threshold c are obtained in Eqs. (4.7) and (4.8):

$$R(t) = \sum_{n=0}^{n_b} \int_{0}^{t} \cdots \int_{0}^{t-1} \prod_{i=1}^{n} f(\tau_i) \left( 1 - F\left(t - \sum_{i=1}^{n} \tau_i\right) \right) H(c|t, \{\tau_i, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_i \quad (4.7)$$

$$\overline{f}(t) = \frac{\partial}{\partial t} \left( 1 - \sum_{n=0}^{n_b} \int_0^t \cdots \int_0^{t-\sum_{i=1}^{n-1} \tau_i} \prod_{i=1}^n f(\tau_i) \left( 1 - F\left(t - \sum_{i=1}^n \tau_i\right) \right) H(c|t, \{\tau_i, 1 \le i \le n\}) \prod_{i=1}^n d\tau_i \right)$$
(4.8)
where  $H(x|t, \{\tau_i, 1 \le i \le n\})$  denotes the conditional CDF of the total degradation of all the branches (including the first degraded location) at time t, given  $\{\tau_i, 1 \le i \le n\}$  are known. The proof of Eq. (4.7) is given in Appendix F. In the case of crack growth and branching, Eq. (4.7) obtains the probability that the total length of all the cracks is less than the failure threshold c given the parameters of  $f(\tau_i)$ and the parameters of the crack growth of the branches (such as the  $\mu$  and  $\sigma$  in the case of Brownian Motions) are known. Eq. (4.8) obtains the PDF of the failure time given the failure threshold is c.

The mean and variance of the degradation of the  $i^{th}$  branch at time t are obtained in Eqs. (4.9) and (4.10) respectively. The proofs of Eqs. (4.9) and (4.10) are given in Appendix F.

$$E(X_{i}^{*}(t)) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j}) E\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right)\right) \prod_{j=1}^{i} d\tau_{j}$$
(4.9)

$$Var(X_{i}^{*}(t)) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1} \tau_{j}} \prod_{j=1}^{i} f(\tau_{j}) E\left(X_{i}^{2}\left(t-\sum_{j=1}^{i} \tau_{j}\right)\right) \prod_{j=1}^{i} d\tau_{j}$$
$$-\left(\int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1} \tau_{j}} \prod_{j=1}^{i} f(\tau_{j}) E\left(X_{i}\left(t-\sum_{j=1}^{i} \tau_{j}\right)\right) \prod_{j=1}^{i} d\tau_{j}\right)^{2}$$
(4.10)

We consider a special case of Type II DBP where the time to branch follows an exponential distribution. The approach can be applied to other distributions that can be approximated by the exponential distribution. Consider a general degradation process whose mean function and variance function of any branch in the time interval (t, t+dt) is  $\mu dt$  and  $\sigma^2 dt$  respectively (e.g. Brownian Motion), when the time to branch follows an exponential distribution with rate  $\lambda$ , the PDF, expectation and variance of the total degradation are obtained in Eqs (4.11), (4.12) and (4.13) respectively. Moreover, the PDF, the expectation and the variance of the total degradation are uniquely determined by  $\lambda$ ,  $\mu$  and  $\sigma$ . The proofs are given in Appendix F.

$$h^{*}(x|t) = e^{-\lambda t} \sum_{n=0}^{n_{b}} \lambda^{n} \int_{0}^{t} \cdots \int_{0}^{t-1} h(x|t, \{\tau_{i}, 1 \le i \le n\}) d\tau_{n} \cdots d\tau_{1}$$
(4.11)

where

$$h(x|t, \{\tau_{i}, 1 \le i \le n\}) = \left(\sqrt{2\pi \left((n+1)t - \sum_{i=1}^{n} (n-i+1)\tau_{i}\right)\sigma^{2}}\right)^{-1} e^{-\frac{\left(x - \mu \left((n+1)t - \sum_{i=1}^{n} (n-i+1)\tau_{i}\right)\right)^{2}}{2\left((n+1)t - \sum_{i=1}^{n} (n-i+1)\tau_{i}\right)\sigma^{2}}}.$$

$$E(X^{*}(t)) = \mu \left(\frac{1}{2}\lambda t^{2} + t\right)$$

$$(4.12)$$

$$Var(X^{*}(t)) = \frac{1}{3}\mu^{2}\lambda t^{3} + \frac{1}{2}\sigma^{2}\lambda t^{2} + \sigma^{2}t$$
(4.13)

The expected number of branches at time t is obtained by Eq. (4.14), which indicates that the expected number of branches initiated is uniquely determined by the rate of branching  $\lambda$  and time t.

$$E(N|t) = \int_{0}^{t} \lambda d\tau = \lambda t$$
(4.14)

When the overall failure threshold is c, the MTTF is

$$MTTF = \frac{-1 + \sqrt{1 + \frac{2\lambda c}{\mu}}}{\lambda}$$
(4.15)

The reliability at time t is

$$R(t) = e^{-\lambda t} \sum_{n=0}^{n_b} \lambda^n \int_0^t \cdots \int_0^{t-\sum_{i=1}^{n-1} \tau_i} H(c|t, \{\tau_i, 1 \le i \le n\}) d\tau_n \cdots d\tau_1$$
(4.16)

where 
$$H(x|t, \{\tau_i, 1 \le i \le n\}) = \frac{1}{2} \left[ 1 + erf\left(\frac{x - \mu\left((n+1)t - \sum_{i=1}^n (n-i+1)\tau_i\right)}{\sqrt{2}\left((n+1)t - \sum_{i=1}^n (n-i+1)\tau_i\right)\sigma^2}\right) \right]$$
 denotes

the conditional CDF of the total degradation of all the branches (including the first degraded location) at time t, given  $\{\tau_i, 1 \le i \le n\}$  are known. Note that  $erf(\cdot)$  is the error function. The PDF of time to failure is

$$\overline{h}(t) = \frac{\partial}{\partial t} \left( 1 - e^{-\lambda t} \sum_{n=0}^{n_b} \lambda^n \int_0^t \cdots \int_0^{t-1} H(c|t, \{\tau_i, 1 \le i \le n\}) d\tau_n \cdots d\tau_1 \right)$$
(4.17)

Suppose the total degradation at time t is x(t), the total number of branches is n(t)and the last branch is initiated at  $t_l \le t$ . The MRL is given in Eq. (4.18) and its proof is shown in Appendix F.

$$MRL = \inf\left\{t_{r} \left| \mu(n(t)+1)(t_{r}-t) + \frac{\mu}{2g(0|t-t_{l})} \sum_{n=1}^{n_{b}-n(t)} \left(g(n|t_{r}-t_{l})n\frac{(t-t_{l})^{n+1}}{(t_{r}-t_{l})^{n}}\right) = c - x(t)\right\}$$
(4.18)

Using Eqs. (4.9) and (4.10) we obtain the mean and variance of the degradation of branch *i* as given in Eqs. (4.19) and (4.20) respectively.

$$E\left(X_{i}^{*}\left(t\right)\right) = \mu\left(\left(t - \frac{i}{\lambda}\right) + \left(\sum_{j=0}^{i-1} \frac{\left(i - j\right)}{j!} t^{j} \lambda^{j-1}\right) e^{-\lambda t}\right)$$
(4.19)

$$Var(X_{i}^{*}(t)) = \frac{\mu^{2}}{\lambda^{2}} \left(\lambda^{2}t^{2} - 2i\lambda t + (i+1)i - e^{-\lambda t} \sum_{j=0}^{i-1} t^{j} \left(\frac{(i-j+1)(i-j)}{j!}\lambda^{j}\right)\right) + \sigma^{2} \left(t - \frac{n}{\lambda} + \left(\sum_{j=0}^{n-1} \frac{(n-j)}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right) - \left(\mu \left(t - \frac{i}{\lambda} + \left(\sum_{j=0}^{i-1} \frac{(i-j)}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right)\right)^{2}$$
(4.20)

Next, we discuss the scenario where Type I DBP can be approximated by the Branching Brownian Motion processes. In Type I DBP the initiations of the new branches are triggered when the degradation levels of the branches reach their thresholds  $\{c_i\}$ . The parameters of the distribution of time to branch are obtained from the corresponding branches' degradation process using the models described in section 4.2.1 and the metrics of the Type I DBP can be obtained accordingly. For a branch in a Brownian Motion process with drift  $\mu_i$  and diffusion  $\sigma_i$ , it is known that the time to cross the threshold  $c_i$ follows an IG distribution  $\tau_i \sim IG\left(\frac{c_i}{\mu_i}, \frac{c_i^2}{\sigma_i^2}\right)$ . Accordingly,  $f(\tau_i) = \sqrt{\frac{c_i^2}{2\pi\tau_i^3\sigma_i^2}} \exp\left(-\frac{\mu_i^2}{2\sigma_i^2}\left(\frac{1}{\tau_i} - \frac{c_i}{\tau_i^2\mu_i}\right)^2\right)$ 

The metrics of the threshold-triggered Branching Brownian Motion process can be obtained using Eqs. (4.1)-(4.10). However, in some cases, we may approximate these processes with a DBP that has branching times following an exponential distribution as discussed in the first part of this section, where closed-form or simplified results are obtained. To do so, the first requirement is that the expectations of the two distributions should be equal, i.e.  $\frac{1}{\lambda} = \frac{c_i}{\mu_i}$ . The second requirement is that the variances of the IG distribution and exponential distribution are close, i.e., the following condition should be satisfied:

$$\left(\frac{c_i}{\mu_i}\right)^3 \left(\frac{c_i^2}{\sigma_i^2}\right)^{-1} \approx \frac{1}{\lambda^2} = \frac{c_i^2}{\mu_i^2}$$

which is equivalent as

$$\sigma_i^2 \approx \mu_i c_i \tag{4.21}$$

In this case,  $IG\left(\frac{c_i}{\mu_i}, \frac{c_i^2}{\sigma_i^2}\right) \approx Exp\left(\frac{\mu_i}{c_i}\right)$ . When  $\mu_i = \mu$  and  $c_i = d$ ,  $n_b = \infty$  Eqs. (4.12),

(4.13) and (4.15) reduce to

$$E(X^*(t)) = \mu t \left(\frac{\mu}{2d}t + 1\right)$$
(4.22)

$$Var(X^{*}(t)) = \frac{1}{3d} \mu^{3} t^{3} + \frac{1}{2d} \sigma^{2} \mu t^{2} + \sigma^{2} t$$
(4.23)

$$MTTF = \frac{d}{\mu} \left( -1 + \sqrt{1 + \frac{2c}{d}} \right) \tag{4.24}$$

where *c* is the overall failure threshold. Figure 4.2 shows an example when  $\mu = 1$ ,  $\sigma = 6$  and d = 36, where the CDFs of the IG distribution and exponential distribution are close.



Figure 4.2 The CDFs of IG distribution and exponential distribution

### 4.2.3 Case Study

As crack growth leads to failures when its total length reaches the threshold as described in (Singh *et al.*, 2017), we use actual crack growth and branching data by Dondeti *et al.* 

(2019) to validate the proposed model. Three identical soda-lime glass specimens are subjected to impact loading in the test under the same condition and monitored with three different methods of Photoelasticity Experiment (PE), Digital Image Correlation (DIC) and Digital Gradient Sensing (DGS). To ensure that the observations obtained with different methods are consistent, calibration techniques such as nonlinear-squares regression, genetic algorithms, subset splitting and others have been used for calibration (Dondeti et al., 2019). Figures 4.3 and 4.4 show examples of the selected contours of DGS images and PE fringes photos of the crack growth and branching process of soda-lime glass specimens. In both examples the cracks are initiated on the left and continue to grow and propagate to the right. The crack in Figure 4.3 is initiated at  $t = 0\mu s$  and branches at  $t = 21\mu s$  while the crack in Figure 4.4 is initiated at  $t = 0\mu s$  and branches at  $t = 18\mu s$ . Note that the 'stitch marks' in Figure 4.4 along the crack path are attributed to contact stresses due to the Rayleigh waves. In this case study we investigate the DBP where only one branch is initiated at every instant of branching.



Figure 4.3 Selected figures of the contours of DGS images of crack growth and

branching process in a soda-lime glass specimen (Dondeti et al., 2019)



Figure 4.4 Selected photos of PE fringes of crack growth and branching process in a

soda-lime glass specimen (Dondeti et al., 2019).



Figure 4.5 The crack growth data of the three specimens used in the case study

Figure 4.5 shows the crack growth data of the three specimens, where the original cracks are plotted in lines with downward-pointing triangles and the branched cracks are plotted in dashed-dot lines with squares. We observe that the crack growth rates of the original cracks and branched cracks are close, so we assume the original and branched cracks' growth follow Brownian Motions with the same parameters. We censor the degradation data at time  $t=25 \ \mu s$  and obtain the crack growth parameters { $\mu=1.602$ ,  $\sigma=0.571$ } using Maximum Likelihood Estimation (MLE).

The distribution of time to branch  $f(\tau)$  is determined by the time that the Stress Intensity Factor (SIF) reaches the branching threshold (Dondeti et al., 2019, Sundaram et al., 2018). As shown in Figure 4.6, it is observed that the SIF starts from an initial lower state (0.8)and increases until it reaches the threshold (1.45) and triggers branching. The SIF then returns to the lower level after branching and the process is restarted. We model the SIF growth as a Brownian Motion process and the first passage time follows an IG distribution. IG Using MLE. we obtain the parameters of the distribution as  $\{\mu_{IG} = 27.778, \lambda_{IG} = 151.169\}$ . The PDF of the IG distribution is shown in Figure 4.7. Besides, the branching times (16, 18, 21) of the three specimens are indicated with crosses. We find that the branching times observed from the experiments match that estimated with SIF growth. This shows that the time to branch can be accurately characterized by the time that SIF reaches its threshold. When the total number of specimens is large, the parameters can be obtained using the crack branching data instead of using SIF data, as the SIF data is more difficult and expensive to obtain.



Figure 4.6 The apparent SIFs of cracks over time



Figure 4.7 The estimated PDF of time to branch based on SIF growth and observed times

of branching of specimens

The mean and 99% confidence interval of total crack length are shown in Figure 4.8. In addition, the total degradation of the cracks of the three specimens are plotted in black. We observe that the model accurately predicts the growth of the total crack length.

We estimate the reliability using Eq. (4.7) for a threshold of 60 *mm* as shown in Figure 4.9. In real-life applications where glass-made structures are used in manufacturing containers, the total crack length is related to the leakage amount. Failure occurs whenever leakage reaches an unacceptable threshold.



Figure 4.8 The actual and predicted total crack length

We estimate the reliability using Eq. (4.7) for a threshold of 60 mm as shown in Figure

4.9.



Figure 4.9 The reliability of the unit/system subject to crack growth and branching

Note that in this study, the reliability metrics are calculated numerically based on Monte Carlo (MC) simulation. The calculation is performed on the MATLAB platform using a MacBook Pro with a 16GB 2400 MHz DDR4 and 2.2 GHz Intel Core i7 processor. The calculation results are shown in Table 4.1 in the increasing order of integral corresponds to n in Eq. (4.1). The contribution of degradation of higher-order branches ( $n \ge 4$ ) is insignificant (<1/1000), so we may stop the calculation of integrals higher than order 4 and

use the total degradation of the main degradation path and the first 3 branches to approximate the total degradation, which saves a significant amount of time.

Order of	0	1	2	3	4	5	6
integral							
Time (s)	0.0093	0.0146	0.0229	0.0825	0.4141	6.3659	98.5081
Contribution	0.2240	0.7073	0.0805	0.0002	< 0.0001	< 0.0001	< 0.0001
to the total							
degradation							

 Table 4.1 The information of numerical calculation of integrals

#### 4.3 Summary

We develop a general stochastic model that characterizes the DBP. The statistical properties of the process, such as the PDF, the mean and variance of the degradation and the related reliability metrics such as the failure time distribution, the reliability and the MTTF are presented. Specifically, closed-form expressions of the expectation and variance of the total degradation and MTTF are obtained when the times to branch follow an exponential distribution. The threshold triggered DBP is also investigated. The models are validated with actual crack growth and branching data. The predictions from the models match the actual data accurately.

#### **CHAPTER 5**

### **GENERAL DIRECTIONAL DEGRADATION BRANCHING PROCESSES**

#### 5.1 Introduction

In many degradation branching applications, the number and directions of the branches significantly affect the total degradation of the system. In Chapter 4, we investigate the degradation branching process and assume that only one branch occurs at the time of branching and in a random direction. However, relaxing this assumption results in the more general case where a random number of branches is initiated at each branching instant and their directions may be bounded. In this chapter, we investigate this general directional degradation branching process and develop a stochastic degradation model that captures both the number and directions of the branches at any branching instant. Reliability metrics such as the mean time to failure, the mean residual life, the failure probability and others are also investigated.

The general directional degradation branching process (GDDBP) is described as follows: (1) The degradation starts in one location and propagates to other locations, (2) After a random time a random number of degradation branches (offsprings) are initiated, (3) Each offspring branch grows in a random direction from its parent branch, (4) The degradation of each branch follows a stochastic process and its degradation rate is related to its branching angle, (5) Failure occurs when the total degradation of all the branches reaches a failure threshold.

## 5.2 Nomenclature

n <sub>tr</sub>	The number of trees.
k	Index of a tree.
t	Time.
$N_{b,k}$	The total number of branchings occur in tree $k$ .
$N_{g,k}$	The total number of generations of branches in tree $k$ .
i	The index of generations of the branches.
<i>i'</i>	$0 \le i' \le i$ , the generation indexes of a branch's ancestors.
$n_{i,k}$	The total number of branchings where branches in
	generation $i$ in tree $k$ are initiated.
$ ilde{n}_{i,k}$	The upper bound of $n_{i,k}$ .
$j_{i'}$	The index of an offspring's $i'^{th}$ -generation ancestor's
	index.
$ ilde{j}_{i'}$	The upper bound of variable $j_{i'}$ .
$\underbrace{j_i, \ldots, j_i}_i, k$	The index of a branch, which consists of its tree index $k$ ,
	its ancestors' indexes in the $1^{st}$ to the $i-1^{th}$ generations

 $(j_1 \text{ to } j_{i-1})$  and its index  $j_i$  among the offsprings in this generation. Specially, when i = 0, the branch index is k, which indicates that this branch is an original branch. The branch whose index is  $B_{j_i, \dots, j_l, k}$ . Note that the  $\{j_{i'}, \text{ for } 0 \le i' \le i\}$  are constants in this expression. The time when  $B_{j_i, \dots, j_l, k}$  branches.

The degradation of branch  $B_{j_1, \dots, j_1, k}$ , where  $\sum_{\substack{i'=1\\i'=1}}^{i-1} \tau_{j_i, \dots, j_1, k}$ is the time it takes to initiate  $B_{j_1, \dots, j_1, k}$  from t = 0. The angle between branch  $B_{j_{i-1}, \dots, j_1, k}$  and its offspring  $B_{j_i, \dots, j_1, k}$ .

The number of offsprings of branch  $B_{\underbrace{j_i, \ldots, j_1, k}}$ . The branching threshold of branch  $B_{\underbrace{j_i, \ldots, j_1, k}}$ .

The drift and diffusion of Brownian motion process of

$$X_{\underbrace{j_i,\ldots,j_1}_i,k}\left(t-\sum_{i'=1}^{i-1}\tau_{\underbrace{j_i',\ldots,j_1}_i,k}\right).$$

The failure threshold.

The PMF of  $O_{\underbrace{j_i, \ldots, j_i}_i, k}$ .

The PDF of  $\tau_{\underbrace{j_i,\ldots,j_l,k}_i}$ .

c  $s\left(O_{\underline{j_i,\ldots,j_l,k}}\right)$   $f\left(\tau_{\underline{j_i,\ldots,j_l,k}}\right)$ 

 $B_{\underbrace{j_i,\ldots,j_1}},k$ 

 $au_{\underline{j_i,\ldots,j_l},k}$ 

 $\alpha_{\underbrace{j_i,\ldots,j_1}}$ 

 $O_{\underbrace{j_i,\ldots,j_1},k}$ 

 $C_{\underbrace{j_i,\ldots,j_l}_i,k}$ 

 $\mu_{\underbrace{j_i,\ldots,j_1}_i,k},\sigma_{\underbrace{j_i,\ldots,j_1}_i,k}$ 

 $X_{\underline{j_i,\ldots,j_l},k}\left(t-\sum_{i'=1}^{i-1}\tau_{\underline{j_i,\ldots,j_l},k}\right)$ 

$F\left( au_{j_i,\ldots,j_i,k} ight)$	The CDF of $\tau_{\underbrace{j_i,\ldots,j_l,k}_i}$ .
$\phi\left(\alpha_{\underbrace{j_i,\ldots,j_l,k}_{i}}\right)$	The PDF of angle $\alpha_{\underbrace{j_i,\ldots,j_l,k}_i}$ .
$\widetilde{Q}_k$	The numbers of offsprings initiated in the branchings in
	tree $k$ .
$\mathbf{O}_k$	The regions of integral of variable $Q_k$ .
$\mathcal{I}_k$	The times when branchings occur in tree $k$ .
$\mathbf{t}_{\widetilde{k}}^{k}$	The regions of integral of variable $\tau_k$ .
$lpha_{k}$	The angles of branches in tree $k$ .
$\mathfrak{a}_{k}$	The regions of integral of variable $\alpha_k$ .
$\mathbf{f}(\underline{\tau}_k)$	The probability that the first $N_{b,k}$ branchings occur by
	time t are $\underline{\tau}_k$ .
$\mathbf{f}'\bigl(\boldsymbol{\tau}_{\!\scriptscriptstyle k}\bigr)$	The probability that no other branchings occur by time $t$
	except those occur at $\tau_k$ .
$\mathbf{s}(\mathcal{Q}_k)$	The probability that the numbers of offsprings initiated in
	the first $N_{b,k}$ branchings are $Q_k$ .
$\boldsymbol{\phi}(\boldsymbol{\mathfrak{A}}_{k})$	The probability that the angles of branches are $\alpha_k$ .

- $\mathbf{h}(x|t, \underline{z}_k, Q_k, \underline{\alpha}_k)$  The PDF of the conditional total degradation of all the branches and the original branch in tree k at time t, given  $Q_k$ ,  $\underline{z}_k$ ,  $\underline{\alpha}_k$  are known.
- $\mathbf{g}(x, \underline{\tau}_k, \underline{Q}_k, \underline{q}_k | t)$  The probability that degradation of tree k is x, the total number of branchings occur by time t is  $N_{b,k}$  and the times that the branchings occur are  $\underline{\tau}_k$ , the angles of branches are  $\underline{q}_k$ , and the numbers of offsprings initiated in the branchings are  $\underline{Q}_k$ .
- $\Psi_{\mathbf{T}_{k},\mathbf{Q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k},Q_{k},\underline{q}_{k}|t\right)\right)$  The probability that the degradation of tree k is x, the total number of branchings occur by time t is  $N_{b,k}$  and the numbers of offsprings initiated in the branchings are  $Q_{k}$
- $\Omega_{\mathbf{Q}_{k},N_{b,k},N_{g,k}}\left(\Psi_{\mathbf{I}_{k},\mathbf{Q}_{k}}\left(\mathbf{g}\right)\right)$  The marginal distribution of degradation amount  $\tilde{X}_{k}\left(t\right)$  when there are  $N_{g,k}$  generations of branches in tree k and the total number of branchings is  $N_{b,k}$ , where  $\mathbf{g} = \mathbf{g}\left(x, \mathbf{z}_{k}, \mathbf{Q}_{k}, \mathbf{Q}_{k} | t\right)$

 $I_{b,k}$  The set of indexes of the branches that branch in tree k by time t

$I_{co,k}$	The set of indexes of the offspring branches in tree $k$ by
	time t
$I_{u,k}$	The set of indexes of the branches that do not branch by
	t. Note that $= I_{o,k} - I_{b,k}$ .
$ ilde{X}_{k}\left(t ight)$	The degradation of tree $k$ .
$\tilde{X}\left(t\right) = \sum_{k=1}^{n_{tr}} \tilde{X}_{k}\left(t\right)$	The total degradation of all the trees.
h(x t)	The PDF of the total degradation at time $t$ .
R(t)	The reliability.
$\overline{f}(t)$	The PDF of the time to failure.

#### 5.3 **General Directional Degradation Branching Process**

#### 5.3.1 Definitions and Properties of the GDDBP

5.3.1.1 The branching times, generations and degradation of the GDDBP

First, we define the general directional degradation branching process (GDDBP). Figure 5.1 shows an example of the branching pattern of GDDBP. At time t=0, a random number of  $n_{tr}$  original branches are initiated and propagate in different directions, each of which grows into a branching tree as degradation proceeds. Figure 5.1 shows the case when  $n_{tr} = 5$ . Note that only the branches initiated at t = 0 are called the original branches, those initiated at t > 0 are simply called branches. Figure 5.2 shows a schematic

diagram of the tree grown from original branch k=1 shown in Figure 5.1 and Figure 5.3 shows the degradation of the original branch and its offspring branches. The offspring branches in the three generations are shown in three different colors in Figure 5.3. Let  $B_k$ denote original branch k and the stochastic process  $X_k(t)$  denote its degradation. The branching of  $B_k$  occurs after a random time  $\tau_k$  and initiates  $O_k = 1$  new branch  $(B_{1,k})$ , whose degradation is  $X_{1,k}(t-\tau_k)$  and  $t-\tau_k$  is its actual degradation time. Note that the subscript 1 is the index of the offspring of its ancestor  $B_k$ . The new branch is in the 1<sup>st</sup> generation of branches as it is initiated from one of the original branches  $B_k$ . Meanwhile,  $X_k(t)$  continues to grow after the new degradation branch(es) is(are) initiated. The offspring(s) continue(s) to grow and may continue to initiate new branches thereafter over time. For example, two ( $O_{1,k} = 2$ ) 2<sup>nd</sup>-generation branches  $B_{1,1,k}$  and  $B_{2,1,k}$  are initiated from branch  $B_{1,k}$  at  $t = \tau_k + \tau_{1,k}$  and two ( $O_{1,1,k} = 2$ ) 3<sup>rd</sup>-generation branches  $B_{1,1,1,k}$  and  $B_{2,1,1,k}$  are initiated from branch  $B_{1,1,k}$  at  $t = \tau_k + \tau_{1,k} + \tau_{1,1,k}$ . The degradation amount of branches are  $X_{1,1,k}\left(t-\tau_k-\tau_{1,k}\right)$  ,  $X_{2,1,k}\left(t-\tau_k-\tau_{1,k}\right)$  , four these  $X_{\scriptscriptstyle 1,1,1,k}\left(t-\tau_{\scriptscriptstyle k}-\tau_{\scriptscriptstyle 1,k}-\tau_{\scriptscriptstyle 1,1,k}\right) \ \text{ and } \ X_{\scriptscriptstyle 2,1,1,k}\left(t-\tau_{\scriptscriptstyle k}-\tau_{\scriptscriptstyle 1,k}-\tau_{\scriptscriptstyle 1,1,k}\right) \ \text{respectively}.$ 



**Figure 5.1** The branching pattern of a GDDBP that has  $n_{tr} = 5$  trees



**Figure 5.2** The tree grown from original branch k = 1 in Figure 5.1



Figure 5.3 Schematic diagram of the degradation of the original branch and branches in

tree k = 1 shown in Figure 5.1

In general,  $B_{\underline{j_i, \dots, j_l}, k}$  is the  $j_i^{th}$  offspring in the  $i^{th}$  generation whose ancestors in generations 1 to i-1 are  $j_1, \dots, j_{i'}, \dots, j_{i-1}$  respectively. Note that  $j_{i'}$  denotes the index of the branch's ancestor in the  $i'^{th}$  generation. The degradation of branch  $B_{\underline{j_i, \dots, j_l}, k}$ 

is 
$$X_{\underbrace{j_i, \dots, j_i}_i, k}\left(t - \sum_{i'=0}^{i-1} \tau_{\underbrace{j_i', \dots, j_i}_i, k}\right)$$
, where  $\sum_{i'=0}^{i-1} \tau_{\underbrace{j_i', \dots, j_i}_i, k}$  denotes the time it takes  $B_{\underbrace{j_{i-1}, \dots, j_i}_{i-1}, k}$  to

initiate branch  $B_{\underbrace{j_i, \ldots, j_l, k}_i}$  starting from t = 0. In addition, note that when arranging the subscripts the  $j_1, j_2, \ldots, j_i$  are reversed as  $j_i, \ldots, j_2, j_1$  as the indexes of the latest generations of offsprings are always put in the first places.

#### 5.3.1.2 The branching mechanisms of the GDDBP

Note that the branching time  $\tau_{\underline{j_i}, \dots, \underline{j_l}, k}$  may have different interpretations in different GDDBPs depending on the characteristics of the real-life applications. We define two types of GDDBPs: Type I GDDBP and Type II GDDBP. In Type I GDDBP,  $\tau_{\underline{j_i}, \dots, \underline{j_l}, k}$  is a

variable corresponding to the time that  $X_{\underline{j_1, \dots, j_l, k}}\left(t - \sum_{i=1}^{i-1} \tau_{\underline{j_r, \dots, j_l, k}}\right)$  reaches its threshold  $c_{\underline{j_l, \dots, j_l, k}}$  (the first passage time) and initiates branching at this time while in Type II GDDBP  $\tau_{\underline{j_l, \dots, j_l, k}}$  is a random variable determined by other physical processes independent of the degradation of the branches. In the soda-lime glass crack branching example,  $\tau_{\underline{j_l, \dots, j_l, k}}$  denotes the time that the Stress Intensity Factor (SIF) (Sundaram *et al.*, 2018) reaches the threshold  $c_{\underline{j_l, \dots, j_l, k}}$  and leads to the initiation of the offspring crack branches. The overall threshold c corresponds to the maximum allowance of the total crack length of all the trees (Sundaram *et al.*, 2018).

Figure 5.3 illustrates the Type I GDDBP. At time  $\tau_k$  branch  $B_{1,k}$  is initiated from  $B_k$ when  $X_k(t)$  reaches its branching threshold  $c_k$ . Then after time  $\tau_{1,k}$  other branches  $\{B_{1,1,k}, B_{2,1,k}\}$  are initiated from branch  $B_{1,k}$  when  $X_{1,k}(t-\tau_k)$  reaches its branching threshold  $c_{1,k}$ . Similarly, after time  $\tau_{1,1,k}$  further branches  $\{B_{1,1,1,k}, B_{2,1,1,k}\}$  are initiated from branch  $B_{1,1,k}$  when  $X_{1,1,k}(t-\tau_k-\tau_{1,k})$  reaches its branching threshold  $c_{1,1,k}$ . In comparison, no branches are initiated from  $B_{2,1,k}$ ,  $B_{1,1,1,k}$  and  $B_{2,1,1,k}$  since

$$X_{2,1,k}(t-\tau_k-\tau_{1,k})$$
,  $X_{1,1,1,k}(t-\tau_k-\tau_{1,k}-\tau_{1,1,k})$  and  $X_{2,1,1,k}(t-\tau_k-\tau_{1,k}-\tau_{1,1,k})$  never  
reach their branching thresholds  $c_{2,1,k}$ ,  $c_{1,1,1,k}$  and  $c_{2,1,1,k}$  by time  $t$ . In general, the  
branching process continues until the cumulative degradation of all the branches reaches  
the overall failure threshold  $c$ .

#### 5.3.1.3 The branching angles and degradation rates of the GDDBP

Rabinovitch *et al.* (2011) show that the growth rates of offspring crack branches are related to their angles with respect to their parent branch. It is assumed that the secondary fractures are initiated at distances greater than a certain minimal radius r from the primary's vertex. As shown in Figure 5.4 (a),  $\beta$  is the angular location and  $\delta$  is the direction of branching from a flaw adjacent to the primary fracture. Figure 5.4 (b) shows the angle  $\beta_1$  where the stress at flaw tip reaches its maximum and the relationship between branching direction  $\delta_1$  and fracture velocity  $V/C_L$ , where V is the instantaneous crack propagation velocity and  $C_L$  is the longitudinal wave speed. Thus it is implied that  $\beta$  is the angle at which the flaw turns into a secondary fracture and  $\beta$  is the direction that crack propagates. It can be concluded that larger angles lead to larger crack growth rates. Let  $\alpha_{\underline{j_1,\ldots,j_k,k}}$  denote the angle of  $B_{\underline{j_1,\ldots,j_k,k}}$ , which is the angle between offspring  $B_{\underline{j_1,\ldots,j_k,k}}$  and its parent branch  $B_{\underline{j_1,\ldots,j_k,k}}$ , the mean crack growth rate of  $B_{\underline{j_1,\ldots,j_k,k}}$  is obtained using Eq. (5.1) as

$$\mu_{\underbrace{j_i,\ldots,j_l,k}_{i}} = \mu_{\underbrace{j_{i-1},\ldots,j_l,k}_{i-1}} \left( \eta_0 + \eta_1 \alpha_{\underbrace{j_i,\ldots,j_l,k}_{i}} \right), \text{ for } i \ge 1$$
(5.1)

where  $\eta_0$  and  $\eta_1$  are constants. If the degradation of branch  $B_{\underbrace{j_i, \dots, j_l, k}}$  is modeled with a Brownian motion process  $\mu_{\underbrace{j_i, \dots, j_l, k}}$  corresponds to its drift and its diffusion is obtained with Eq. (5.2):

$$\sigma_{\underbrace{j_i,\ldots,j_l,k}_{i}} = \sigma_{\underbrace{j_{i-1},\ldots,j_l,k}_{i-1}} \left( \eta_0 + \eta_1 \alpha_{\underbrace{j_i,\ldots,j_l,k}_{i}} \right)^{1/2}, \text{ for } i \ge 1$$
(5.2)

Cauchy distribution is used to model crack angles in (Rice, 2006). The PDF of the branching angles modeled by Cauchy distribution is given in Eq. (5.3).

$$\phi\left(\alpha_{\underbrace{j_{i},\ldots,j_{l},k}_{i}}\right) = \frac{\gamma}{\pi\left(\alpha_{\underbrace{j_{i},\ldots,j_{l},k}_{i}}^{2} + \gamma^{2}\right)}$$
(5.3)

where  $\gamma$  is a constant.



**Figure 5.4** (a) Definitions of the angular location and direction of branching from a flaw adjacent to the primary fracture (b) The branching direction as a function of fracture

velocity (Rabinovitch et al., 2011)

## 5.3.2 Probabilistic Properties of the GDDBP

5.3.2.1 The conditional degradation of a tree

In tree k the numbers of offsprings initiated in the branchings are denoted as vector

$$Q_k = \left\{ O_{\underbrace{j_{i-1}, \dots, j_1, k}_{i-1}, k}, \text{ for } \underbrace{j_{i-1}, \dots, j_1}_{i-1}, k \in \underline{I}_{b,k} \right\}, \text{ where } \underline{I}_{b,k} \text{ is the set of indexes of the}$$

branches that branch in tree k by time t, the times when branchings occur are denoted

as vector 
$$\underline{\tau}_k = \left\{ \tau_{\underbrace{j_{i-1}, \dots, j_l, k}}, \text{ for } \underbrace{j_{i-1}, \dots, j_l}_{i-1}, k \in \underline{I}_{b,k} \right\}, \text{ the angles of the branches are denoted}$$

as vector  $\alpha_k = \left\{ \alpha_{\underbrace{j_i, \dots, j_i, k}_{i}, k}, \text{ for } \underbrace{j_i, \dots, j_i}_{i}, k \in I_{o,k} \right\}, \text{ where } I_{o,k} \text{ is the set of indexes of the}$ 

offspring branches in tree k by time t. In the example of Figures 5.2 and 5.3,  $I_{b,k} = \{(k), (1,k), (1,1,k)\}$ ,  $I_{o,k} = \{(1,k), (1,1,k), (2,1,k), (1,1,1,k), (2,1,1,k)\}$ ,

$$Q_k = (O_k = 1, O_{1,k} = 1, O_{1,1,k} = 2)'$$
,  $\tau_{k} = (\tau_k, \tau_{1,k}, \tau_{1,1,k})'$  and

Let 
$$f\left(\tau_{\underbrace{j_{i},\ldots,j_{l},k}{i}}\right)$$
 and  $F\left(\tau_{\underbrace{j_{i},\ldots,j_{l},k}{i}}\right)$  denote respectively the probability density function  
(PDF) and cumulative distribution function (CDF) of  $\tau_{\underbrace{j_{i},\ldots,j_{l},k}{i}}$ . Given  $\underline{\tau}_{k}$ ,  $\underline{Q}_{k}$  and  $\underline{q}_{k}$   
are known, the conditional total degradation of tree  $k$  at time  $t$  is  
 $\sum_{\underbrace{j_{i},\ldots,j_{l},k\in I_{o,k}}{i}} X_{\underbrace{j_{i},\ldots,j_{l},k}{i}}\left(t - \sum_{i'=0}^{i-1} \tau_{\underbrace{j_{i'},\ldots,j_{l},k}{i}}\right)$  and its PDF is denoted by  $\mathbf{h}(x|t,\underline{\tau}_{k},Q_{k},\underline{q}_{k})$ . For

example, in the crack growth and branching shown in Figure 5.3, if the degradation of the branches follow Brownian Motions with drift vector and diffusion vector

$$\boldsymbol{\mu} = \left(\mu_{k}, \ \mu_{1,k}, \ \mu_{2,1,k}, \ \mu_{2,1,k}, \ \mu_{2,1,1,k}\right)', \quad \boldsymbol{\sigma} = \left(\sigma_{k}, \ \sigma_{1,k}, \ \sigma_{2,1,k}, \ \sigma_{2,1,k}, \ \sigma_{2,1,1,k}\right)'$$
and  $\boldsymbol{\tau}_{k}, \ \boldsymbol{Q}_{k}, \ \boldsymbol{Q}_{k}, \ \boldsymbol{Q}_{k}$  are known, the conditional total crack length in tree  $k$  is shown in Eq. (5.4).

$$\sum_{\substack{j_{1}, \dots, j_{1}, k \in I_{o,k} \\ i}} X_{\underline{j_{1}, \dots, j_{1}, k}} \left( t - \sum_{i'=0}^{i-1} \tau_{\underline{j_{i'}, \dots, j_{1}, k}} \right)$$

$$\sim N \begin{pmatrix} \sum_{\substack{j_{1}, \dots, j_{1}, k \in I_{o,k} \\ i} \\ k \in I_{o,k} \\ k \neq I_{o,k} \\ k$$

# 5.3.2.2 The PDF, mean and variance of degradation of the GDDBP

Let  $\tilde{X}_{k}(t)$  denote the degradation of tree k, The PDF of  $\tilde{X}_{k}(t)$  is given by Eq. (5.5)

$$\tilde{h}_{k}\left(x|t\right) = \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \mathbf{\Omega}_{\mathbf{Q}_{k},N_{b,k},N_{g,k}}\left(\mathbf{\Psi}_{\mathbf{\tau}_{k},\mathbf{q}_{k}}\left(\mathbf{f}\left(\boldsymbol{\tau}_{k}\right)\mathbf{f}'\left(\boldsymbol{\tau}_{k}\right)\mathbf{s}\left(\boldsymbol{Q}_{k}\right)\mathbf{\phi}\left(\boldsymbol{\alpha}_{k}\right)\mathbf{h}\left(x|t,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\alpha}_{k}\right)\right)\right)$$
$$= \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \mathbf{\Omega}_{\mathbf{Q}_{k},N_{b,k},N_{g,k}}\left(\mathbf{\Psi}_{\mathbf{\tau}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\alpha}_{k}\right|t\right)\right)\right)$$
(5.5)

where  $N_{b,k}$  is the total number of branchings that occur when growing tree k,  $N_{g,k}$  is the total number of generations of all the branches in tree k,  $\mathbf{f}(\underline{\tau}_k)$  is the probability that the first  $N_{b,k}$  branchings occur at  $\underline{\tau}_k$  as shown in Eq. (5.6),  $\mathbf{f}'(\underline{\tau}_k)$  is the probability that no other branchings occur by time t except those occur at  $\underline{\tau}_k$  as shown in Eq. (5.7),  $\mathbf{s}(Q_k)$  is the probability that the numbers of offsprings initiated in the first  $N_{b,k}$ branchings are  $Q_k$  as shown in Eq. (5.8),  $\phi(\underline{\alpha}_k)$  is the probability that the angles of the branches are  $\underline{\alpha}_k$  as shown in Eq. (5.9).

$$\mathbf{f}\left(\boldsymbol{\tau}_{k}\right) = \prod_{\underbrace{j_{i-1}, \dots, j_{1}, k \in \boldsymbol{I}_{b,k}}_{i-1}} f\left(\boldsymbol{\tau}_{\underbrace{j_{i-1}, \dots, j_{1}, k}_{i-1}}\right)$$
(5.6)

$$\mathbf{f}'(\underline{\tau}_k) = \prod_{\underline{j_{i-1}, \dots, j_l, k \in \underline{I}_{u,k}} \left( 1 - F\left(t - \sum_{i'=0}^{i-1} \tau_{\underline{j_i, \dots, j_l, k}}\right) \right)$$
(5.7)

$$\mathbf{s}(\mathcal{Q}_{k}) = \prod_{\underbrace{j_{i-1}, \dots, j_{1}, k \in \underline{I}_{b,k}}_{i-1}} s\left(\mathcal{O}_{\underbrace{j_{i-1}, \dots, j_{1}, k}_{i-1}}\right)$$
(5.8)

$$\boldsymbol{\phi}(\boldsymbol{\alpha}_{k}) = \prod_{\substack{j_{i}, \dots, j_{i}, k \in \boldsymbol{I}_{o,k} \\ i}} \boldsymbol{\phi}\left(\boldsymbol{\alpha}_{\underbrace{j_{i}, \dots, j_{i}, k}_{i}}\right)$$
(5.9)

In Eq. (5.8)  $s\left(O_{\underline{j_{i-1},\ldots,j_l,k}}\right)$  is the Probability Mass Function (PMF) of  $O_{\underline{j_{i-1},\ldots,j_l,k}}$ ,  $\phi\left(\alpha_{\underline{j_i,\ldots,j_l,k}}\right)$  is the PDF of  $\alpha_{\underline{j_i,\ldots,j_l,k}}$ ,  $I_{u,k} = I_{o,k} - I_{b,k}$ . Let  $\mathbf{g}\left(x, \underline{\tau}_k, Q_k, \underline{\alpha}_k | t\right) = \mathbf{f}\left(\underline{\tau}_k\right)\mathbf{f}'(\underline{\tau}_k)\mathbf{s}(Q_k)\mathbf{\phi}(\underline{\alpha}_k)\mathbf{h}(x|t, \underline{\tau}_k, Q_k, \underline{\alpha}_k)$ , it is the probability that the total degradation of tree k is x, the total number of branchings occur by time t is  $N_{b,k}$ and the times that the branchings occur are  $\underline{\tau}_k$ , the directional angles of the branches are

$$\Psi_{\mathfrak{r}_{k},\mathfrak{g}_{k}}\left(\mathbf{g}\left(x,\mathfrak{T}_{k},\mathcal{Q}_{k},\mathfrak{Q}_{k},\mathfrak{Q}_{k}\left|t\right.\right)\right)=\int_{0}^{t}\dots\int_{0}^{t-\sum_{i=0}^{t-2}\tau_{j_{i}',\dots,j_{1},k}}\dots\int_{\alpha_{i}}^{\alpha_{u}}\dots\int_{\alpha_{i}}^{\alpha_{u}}\mathbf{g}\left(x,\mathfrak{T}_{k},\mathcal{Q}_{k},\mathfrak{Q}_{k}\right|t\right)$$

$$\sum_{\substack{j_{i-1},\dots,j_{k}\\i=1}}^{\infty}o_{j_{i-1},\dots,j_{k},k}\cdots\int_{\alpha_{i}}^{\alpha_{u}}\mathbf{g}\left(x,\mathfrak{T}_{k},\mathcal{Q}_{k},\mathfrak{Q}_{k}\right)\right)$$

$$\times\prod_{\substack{j_{i},\dots,j_{i},k\in I_{o,k}\\i}}d\alpha_{j_{i},\dots,j_{i},k}\prod_{j_{i-1},\dots,j_{i},k\in I_{b,k}}d\tau_{j_{i-1},\dots,j_{i},k}$$

$$(5.10)$$

The notation  $\Omega_{\mathbf{Q}_k, N_{b,k}, N_{g,k}} \left( \Psi_{\mathbf{I}_k, \mathbf{q}_k} \left( \mathbf{g} \left( x, \mathbf{z}_k, \mathbf{Q}_k, \mathbf{q}_k | t \right) \right) \right)$  denotes the marginal distribution of  $\tilde{X}_k(t)$  when there are  $N_{g,k}$  generations of branches and the total number of branchings is  $N_{b,k}$  in tree k as shown in Eq. (5.11). It is obtained by summing  $\Psi_{\mathbf{I}_k, \mathbf{q}_k} \left( \mathbf{g} \left( x, \mathbf{z}_k, \mathbf{Q}_k, \mathbf{q}_k | t \right) \right)$  when  $\mathbf{Q}_k$  takes different sets of known values in the regions of integral  $\mathbf{Q}_k$ .

where 
$$\sum_{\substack{j_1, \dots, j_l, k \in I_{b,k} \\ i}} O_{\underline{j_i, \dots, j_l, k}} > 0 \text{ for } \forall 1 \le i \le N_{g,k}, Q_k \text{ is the vector of nonzero } O_{\underline{j_{i-1}, \dots, j_l, k} \atop i \le j_{j-1}}$$

for  $1 \le j_{i'} \le \tilde{j}_{i'}$ ,  $1 \le i' \le i-1$ ,  $1 \le i \le N_{g,k}$ , where  $\tilde{j}_{i'}$  is the upper bound of variable  $j_{i'}$  defined as follows:

$$\tilde{j}_{i'} = \begin{cases} O_{1,k}, \ i' = 1\\ \sum_{j_{i'}=1}^{\tilde{j}_{i'-1}} O_{\underline{j_{i'}, \dots, j_2}, 1, k}, \ 1 < i' \le i \end{cases}$$

The notation  $n_{i,k}$  denotes the number of branchings that initiate branches in generation *i* and  $\tilde{n}_{i,k}$  denotes its upper bound, which is obtained as follows:

$$\tilde{n}_{i',k} = \begin{cases} \min\left\{\hat{j}_{i'}, N_{b,k}\right\}, i' = 1\\ \min\left\{\hat{j}_{i'}, N_{b,k} - \sum_{i'=1}^{i'-1} n_{i',k}\right\}, i' > 1 \end{cases}$$

where

$$\widehat{j}_{i'} = \begin{cases} 1, \ i' = 1 \\ O_{1,k}, \ i' = 2 \\ o_{\underline{j_{i'-2}, \dots, j_{l'}k}} \\ \sum_{j_{i'-1}=1}^{i'-2} O_{\underline{j_{i'-1}, \dots, j_{1}, k}}, \ 2 < i' \le i \end{cases}$$

In Eq. (5.11), all of the possible combinations of  $Q_k$  are enumerated and the corresponding vectors  $\underline{\tau}_k$  and  $\underline{\alpha}_k$  are defined. The total degradation is  $\tilde{X}(t) = \sum_{k=1}^{n_{tr}} \tilde{X}_k(t)$ , where  $n_{tr}$  is a random variable with a PMF of  $w(n_{tr})$ . The PDF of the

total degradation of all the trees is obtained as shown in Eq. (5.12).

$$h(x|t) = \sum_{n_{tr=1}}^{u_{tr}} w(n_{tr}) \prod_{k=1}^{n_{tr}} *h_k(x|t)$$
(5.12)

where  $h_k(x|t)$ ,  $1 \le k \le n_{tr}$  is the PDF of  $\tilde{X}_k(t)$  and  $\prod_{k=1}^{n_{tr}} *h_k(x|t)$  denotes the

convolution of  $h_k(x|t)$  for  $1 \le k \le n_{tr}$ . The expectation of the total degradation is

$$E\left(\tilde{X}\left(t\right)\right) = \sum_{n_{tr}=1}^{u_{tr}} w(n_{tr}) \sum_{k=1}^{n_{tr}} \int_{-\infty}^{\infty} \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \boldsymbol{\Omega}_{\boldsymbol{Q}_{k},N_{g,k}} \left(\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{Q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathcal{Q}}_{k},\boldsymbol{\mathfrak{Q}}_{k}\left|t\right)x\right)\right) dx$$
(5.13)

According to the law of total variance, the variance of the total degradation is

$$Var\left(\tilde{X}(t)\right) = \sum_{n_{tr}=1}^{u_{tr}} w(n_{tr}) \operatorname{var}\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}(t) \middle| n_{tr}\right) + \sum_{n_{tr}=1}^{u_{tr}} w(n_{tr}) \left(E\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}(t) \middle| n_{tr}\right)\right)^{2} - \left(\sum_{n_{tr}=1}^{u_{tr}} w(n_{tr}) \left(E\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}(t) \middle| n_{tr}\right)\right)\right)^{2}$$
(5.14)

where  $\operatorname{var}\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}(t) \middle| n_{tr}\right)$  is the variance of the total degradation of the  $n_{tr}$  trees as shown in Eq. (5.15) and  $E\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}(t) \middle| n_{tr}\right)$  is the mean of the total degradation of the  $n_{tr}$ 

trees as shown in Eq. (5.16).

$$\operatorname{var}\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}\left(t\right) \middle| n_{tr}\right) = \sum_{k=1}^{n_{tr}} \int_{-\infty}^{\infty} \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \boldsymbol{\Omega}_{\boldsymbol{Q}_{k},N_{b,k},N_{g,k}} \left(\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{t}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{t}}_{k},\boldsymbol{\mathcal{Q}}_{k},\boldsymbol{\mathfrak{q}}_{k}\left|t\right)x^{2}\right)\right) dx - \sum_{k=1}^{n_{tr}} \left(\int_{-\infty}^{\infty} \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \boldsymbol{\Omega}_{\boldsymbol{Q}_{k},N_{b,k},N_{g,k}}\left(\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{t}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{t}}_{k},\boldsymbol{\mathcal{Q}}_{k},\boldsymbol{\mathfrak{q}}_{k}\left|t\right)x\right)\right) dx\right)^{2}$$
(5.15)

$$E\left(\sum_{k=1}^{n_{tr}} \tilde{X}_{k}\left(t\right) \middle| n_{tr}\right) = \sum_{k=1}^{n_{tr}} \left(\int_{-\infty}^{\infty} \sum_{N_{b,k}=0}^{\infty} \sum_{N_{g,k}=0}^{N_{b,k}} \mathbf{\Omega}_{\mathbf{Q}_{k},N_{b,k},N_{g,k}}\left(\mathbf{\Psi}_{\mathbf{\tau}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\middle|t\right)x\right)\right) dx\right)$$
(5.16)

#### 5.3.2.3 The reliability metrics of degradation of the GDDBP

Starting from t = 0, when the failure threshold is c, the reliability is

$$R(t) = \int_{x=-\infty}^{c} h(x|t)$$
(5.17)

The PDF of the failure time is

$$\overline{f}(t) = \frac{\partial}{\partial t} \left( 1 - \int_{x = -\infty}^{c} h(x|t) \right)$$
(5.18)

The residual life is investigated as follows. Figure 5.2 shows tree k at time t, when the total degradation of tree k is  $x_k(t)$ , the branching times are  $\underline{\tau}_k$ , the numbers of offsprings are  $Q_k$ , the branching angles are  $\underline{\alpha}_k$ , the set of indexes of the branches that branched by time t is  $\underline{I}_{b,k}$  and the set of the indexes of the branches that have not branched by time t is  $\underline{I}_{u,k}$ . Figure 5.5 shows tree k at time  $t_r > t$ , where  $\underline{I}_{b,k} = \{(k), (1,k), (1,1,k)\}, \underline{I}_{u,k} = \{(2,1,k), (1,1,1,k), (2,1,1,k)\}, \underline{\tau}_k = (\tau_k, \tau_{1,k}, \tau_{1,1,k})', Q_k = (O_k = 1, O_{1,k} = 1, O_{1,1,k} = 2)'$  and  $\underline{\alpha}_k = (\alpha_{1,k}, \alpha_{1,1,k}, \alpha_{2,1,k}, \alpha_{2,1,1,k})'$ . Note that the branches whose indexes are in  $\underline{I}_{u,k} = \{(2,1,k), (1,1,1,k), (2,1,1,k)\}$  have not branched by t and may grow into trees after t.



Figure 5.5 The growth of tree k after time t, the bold solid lines represent branches

that have not branched by t

Eq. (5.19) shows the total degradation growth of all the  $n_{tr}$  trees in  $(t, t_r)$ 

$$\sum_{k=1}^{n_{tr}} \tilde{X}_{k}\left(t_{r} \middle| t, \underline{\tau}_{k}, \underline{\mathcal{Q}}_{k}, \underline{\alpha}_{k}, \underline{\mathcal{I}}_{u,k}\right) = \sum_{k=1}^{n_{tr}} \sum_{\underline{j_{i}, \dots, j_{1}, k \in \underline{\mathcal{I}}_{b,k}}} X_{\underline{j_{i}, \dots, j_{1}, k}\left(t_{r} - t\right) + \sum_{k=1}^{n_{tr}} \sum_{\underline{j_{i}, \dots, j_{1}, k \in \underline{\mathcal{I}}_{u,k}}} \tilde{X}_{\underline{j_{i}, \dots, j_{1}, k}\left(t_{r} \middle| t, \underline{\tau}_{k}, \underline{\mathcal{Q}}_{k}, \underline{\alpha}_{k}\right)$$

$$(5.19)$$

where  $\sum_{\substack{j_1, \dots, j_l, k \in I_{b,k} \\ i}} X_{j_1, \dots, j_l, k}(t_r - t)$  is the sum of the degradation growth of the branches

that have branched before t in  $(t, t_r)$ . Let  $\tilde{X}_{\underline{j_1, \dots, j_l, k}}(t_r | t, \underline{z}_k, \underline{Q}_k, \underline{q}_k)$  denote the degradation of the sub-tree grown from branch  $B_{\underline{j_1, \dots, j_l, k}}$  in  $(t, t_r)$ . For simplicity, let  $B_{k'}$  denote  $B_{\underline{j_1, \dots, j_l, k}}$ , it is the "original branch" of its ancestors after t. Similarly, the offspring of  $B_{k'}$  whose ancestors in generations 1 to i-1 after  $B_{k'}$  are  $j'_1, \dots, j'_{i'}, \dots, j'_{i-1}$  is denoted as  $B_{\underline{j_1, \dots, j_l, k'}}$ . In Figure 5.5,  $B_{2,1,1,k}$  is defined as  $B_{k'}$  and its next-generation offsprings are denoted as  $B_{1,k'}$  and  $B_{2,k'}$ . The numbers of offsprings initiated in the branchings after  $B_{k'}$  are denoted as vector  $Q'_k$ , the times when branchings

occur are denoted as vector  $\underline{\tau}'_{k}$ , the angles of the branches are denoted as vector  $\underline{\alpha}'_{k}$ , the set of indexes of the branches that branch by time  $t_{r}$  is denoted by  $\underline{I}'_{b,k}$ , the set of indexes of the offspring branches by time  $t_{r}$  is denoted by  $\underline{I}'_{o,k}$ . Let  $t_{l} = \sum_{i=0}^{i-1} \tau_{\underline{j}_{i}, \ldots, \underline{j}_{l}, k}$ , it denotes the time that branch  $B_{\underline{j}_{i}, \ldots, \underline{j}_{l}, k}$  is initiated, where  $\underline{j}_{i}, \ldots, \underline{j}_{l}, k \in \underline{I}_{b,k}$ . Let  $h_{b,t}(t_{r})$  denote the PDF of  $\sum_{k=1}^{n_{r}} \sum_{\underline{j}_{i}, \ldots, \underline{j}_{l}, k \in \underline{I}_{b,k}} X_{\underline{j}_{i}, \ldots, \underline{j}_{l}, k}(t_{r}-t)$ . The PDF of  $\tilde{X}_{\underline{j}_{i}, \ldots, \underline{j}_{l}, k}(t_{r}|t, \underline{\tau}_{k}, Q_{k}, \underline{\alpha}_{k})$  is

obtained with Eq. (5.20).

$$h_{\underbrace{j_{i},\ldots,j_{l},k}_{i}}\left(t_{r}\left|t\right.\right) = \sum_{N_{b,k}^{\prime}=0}^{\infty} \sum_{N_{g,k}^{\prime}=0}^{N_{b,k}^{\prime}} \mathbf{\Omega}_{\mathbf{Q}_{k},N_{b,k}^{\prime},N_{g,k}^{\prime}}\left(\mathbf{\Psi}_{\mathbf{\mathfrak{I}}_{k}^{\prime},\mathbf{\mathfrak{G}}_{k}^{\prime}}\left(\mathbf{g}\left(x,\mathbf{\mathfrak{I}}_{k}^{\prime},\mathcal{Q}_{k}^{\prime},\mathbf{\mathfrak{G}}^{\prime}\left|t_{k}\right.\right)\right)\right)$$
(5.20)

where  $\underbrace{j_i, \ldots, j_1}_{i}, k \in \underline{I}_{u,k}$  and  $\Psi'_{\underline{\tau}'_k, \underline{q}'_k} \left( \mathbf{g} \left( x, \underline{\tau}'_k, \underline{Q}'_k, \underline{\alpha}' | t_k \right) \right)$  is defined in Eq. (5.21).

$$\Psi_{\underline{\mathbf{x}}_{k},\underline{\mathbf{y}}_{k}'}^{\prime}\left(\mathbf{g}\left(x,\underline{\mathbf{x}}_{k}',\underline{\mathbf{Q}}_{k}',\underline{\mathbf{q}}'|t_{k}\right)\right)=\int_{t-t_{l}}^{t_{r}-t_{l}}\dots\int_{0}^{t_{r}-t_{l}-\sum_{i=0}^{i-2}\tau_{j_{i}'}\dots,j_{i},k}\dots\int_{\alpha_{l}}^{\alpha_{u}}\dots\int_{\alpha_{l}}^{\alpha_{u}}\dots\int_{\alpha_{l}}^{\alpha_{u}}\dots$$

$$\times\frac{\mathbf{g}\left(x,\underline{\mathbf{x}}_{k}',\underline{\mathbf{Q}}_{k}',\underline{\mathbf{q}}'|t_{k}\right)}{\left(1-F\left(t-t_{l}\right)\right)}\prod_{j_{i}',\dots,j_{l}',k\in I_{o,k}'}d\alpha_{j_{i}',\dots,j_{i}',k'}\prod_{i=1}\prod_{i=1}^{i-1}d\tau_{j_{i-1}',\dots,j_{l}',k'}d\tau_{j_{i-1}',\dots,j_{l}',k'}$$
(5.21)

The PDF of the total degradation of the  $n_{tr}$  trees in  $(t, t_r)$  is obtained with Eq. (5.22)

$$h(x|t_{r},t) = h_{b,t}(t_{r}) * \prod_{k=1}^{n_{tr}} * \prod_{\substack{j_{1}, \dots, j_{1}, k \in I_{u,k}}} * h_{\underline{j_{1}, \dots, j_{1}, k}}(t_{r}|t)$$
(5.22)
where \* is the convolution operator,  $\prod_{\substack{j_i, \dots, j_l, k \in I_{u,k} \\ i}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the convolution of } h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ for all } \underbrace{j_i, \dots, j_l}_{i}, k \in I_{u,k}, \quad \prod_{k=1}^{n_r} *\prod_{\substack{j_1, \dots, j_l, k \in I_{u,k} \\ i}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l, k}\left(t_r | t\right) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l, k \in I_{u,k}} *h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes the } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l} (h_i | t_i) \text{ denotes } h_{j_i, \dots, j_l} = \sum_{i=1}^{n_r} \prod_{j_i, \dots, j_l}$ 

convolution of all  $\prod_{\substack{j_1, \dots, j_1, k \in I_{u,k} \\ i}} *h_{j_1, \dots, j_1, k}(t_r | t)$  for  $1 \le k \le n_{tr}$ . When the failure threshold is

c, the reliability of the system at  $t_r > t$  is

$$R(t_{r}) = \int_{x=-\infty}^{c-\sum_{k=1}^{n_{tr}} x_{k}(t)} h(x|t_{r},t)$$
(5.23)

The mean residual life is obtained as follows. The expectation of degradation growth of tree k in  $(t, t_r)$  is the sum of the expectation of the growth of the branches that have branched and those that have not branched as shown in Eq. (5.24).

$$E\left(\tilde{X}_{k}\left(t_{r}\big|t,\underline{\tau}_{k},\underline{Q}_{k},\underline{\alpha}_{k},\underline{I}_{u,k}\right)\right)$$

$$=\sum_{\underline{j_{i},\ldots,j_{i}},k\in I_{b,k}}E\left(X_{\underline{j_{i},\ldots,j_{i}},k}\left(t_{r}-t\right)\right)+\sum_{\underline{j_{i},\ldots,j_{i}},k\in I_{u,k}}\left(E\left(\tilde{X}_{\underline{j_{i},\ldots,j_{i}},k}\left(t_{r}\big|t,\underline{\tau}_{k},\underline{Q}_{k},\underline{\alpha}_{k}\right)\right)\right)$$
(5.24)

where  $E\left(\tilde{X}_{\underbrace{j_i, \ldots, j_l, k}}\left(t_r | t, \underline{\tau}_k, \underline{Q}_k, \underline{\alpha}_k\right)\right)$  denotes the expectation of degradation of the subtree grown from branch  $B_{\underbrace{j_i, \ldots, j_l, k}_{i}}$ , where  $\underbrace{j_i, \ldots, j_l}_{i}, k \in \underline{I}_{u,k}$  as shown in Eq. (5.25) as

follows.

$$E\left(\tilde{X}_{\underbrace{j_{i},\ldots,j_{i},k}}\left(t_{r}\left|t,\underline{\tau}_{k},Q_{k},\underline{\alpha}_{k}\right.\right)\right)$$

$$=\int_{-\infty}^{\infty}x\sum_{N_{b,k}^{\prime}=0}^{\infty}\sum_{N_{g,k}^{\prime}=0}^{N_{b,k}^{\prime}}\boldsymbol{\Omega}_{\underline{Q}_{k},N_{b,k}^{\prime},N_{g,k}^{\prime}}\left(\boldsymbol{\Psi}_{\underline{\tau}_{k}^{\prime},\underline{q}_{k}^{\prime}}\left(\boldsymbol{g}\left(x,\underline{\tau}_{k}^{\prime},Q_{k}^{\prime},\underline{\alpha}^{\prime}\left|t_{k}\right.\right)\right)\right)dx$$
(5.25)

The MRL is

$$MRL = \inf\left\{t_r \left| E\left(\tilde{X}_k\left(t_r \middle| t, \tau_k, Q_k, Q_k, Q_k, I_{u,k}\right)\right) = c - \sum_{k=1}^{n_{tr}} x_k\left(t\right)\right\}$$
(5.26)

#### 5.4 Case Study

In this section, we use a simulated dataset to show the applicability of the proposed model.

#### 5.4.1 Assumptions and Parameters Used in the Case Study

In the case study, we let  $O_{\underline{j_i, \dots, j_i}}$  follow a Bernoulli distribution,  $\tau_{\underline{j_i, \dots, j_i}}$  follow a uniform distribution,  $\alpha_{\underline{j_i, \dots, j_i}}$  follow a uniform distribution and  $X_{\underline{j_i, \dots, j_i, k}}\left(t - \sum_{i'=0}^{i-1} \tau_{\underline{j_r, \dots, j_i, k}}\right)$  follow Brownian motions. According to Figure 5.4, the

relationship between crack propagation rates and branching angle  $\delta$  is modeled with a linear model as shown in Figure 5.6, where the intercept is 0.1753 and the slope is 0.0066. Therefore, the coefficients  $\eta_0$  and  $\eta_1$  in Eqs. (5.1) and (5.2) are 0.1753 and 0.0066 respectively.



Figure 5.6 The linear fit of the crack propagation rate—branching angle data

The parameters  $\alpha_l$  and  $\alpha_u$  are obtained from the paper by Kobayashi *et al.* (1974) as follows:

$$\phi: \alpha_{\underline{j_i, \dots, j_i}} \sim unif(\alpha_i = 24^\circ, \ \alpha_u = 34^\circ)$$

The other parameters are shown as follows:

$$t = 10$$
  
s:  $s\left(O_{j_{i}, \dots, j_{i}}\right) = \begin{cases} 0.5, \text{ if } O_{j_{i}, \dots, j_{i}} = 1\\ 0.5, \text{ if } O_{j_{i}, \dots, j_{i}} = 2 \end{cases}$   
f:  $\tau_{j_{i}, \dots, j_{i}} \sim unif(4.5, 5.5)$   
w:  $w(n_{tr}) = \begin{cases} 0.5, \text{ if } n_{tr} = 4\\ 0.5, \text{ if } n_{tr} = 5 \end{cases}$ 

$$X_{\underbrace{j_i,\ldots,j_1},k}\left(t-\sum_{i'=0}^{i-1}\tau_{\underbrace{j_i',\ldots,j_1},k}\right) \sim N\left(\mu_{\underbrace{j_i,\ldots,j_1},k}\left(t-\sum_{i'=0}^{i-1}\tau_{\underbrace{j_{i'},\ldots,j_1},k}\right), \ \sigma_{\underbrace{j_i,\ldots,j_1},k}^2\left(t-\sum_{i'=0}^{i-1}\tau_{\underbrace{j_{i'},\ldots,j_1},k}\right)\right)$$

## 5.4.2 Total Degradation Estimation

First, the different branching patterns are enumerated as shown in Table 5.1.

Case ID	$N_{g,k}$	$N_{b,k}$	<i>n</i> <sub>1,<i>k</i></sub>	<i>n</i> <sub>2,k</sub>	<i>n</i> <sub>3,k</sub>
1	1	1	1	None	None
2	2	2	1	1	None
3		3	1	2	None

Table 5.1 The different cases of branching patterns in the case study

The PDF of is obtained by using Eq. (5.6) as follows

$$\begin{split} \tilde{h}_{k}\left(x\left|t\right) &= \mathbf{\Omega}_{\mathbf{Q}_{k},1,1}\left(\mathbf{\Psi}_{\mathbf{I}_{k},\mathbf{Q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\alpha}_{k}\left|t\right)\right)\right) + \mathbf{\Omega}_{\mathbf{Q}_{k},2,2}\left(\mathbf{\Psi}_{\mathbf{I}_{k},\mathbf{Q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\alpha}_{k}\left|t\right)\right)\right) \\ &+ \mathbf{\Omega}_{\mathbf{Q}_{k},3,2}\left(\mathbf{\Psi}_{\mathbf{I}_{k},\mathbf{Q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\alpha}_{k}\left|t\right)\right)\right)\right) \end{split}$$

To calculate  $\Omega_{\mathbf{Q}_{k},N_{b,k},N_{g,k}}\left(\Psi_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\mathbf{z}_{k},\mathbf{Q}_{k},\mathbf{q}_{k}\left|t\right)\right)\right)$  when  $N_{b,k}$  and  $N_{g,k}$  take different values, specific sub-categories of each case are enumerated as shown in Figures 5.7, 5.8 and 5.9 respectively. We discuss them in detail as follows:

There are two sub-categories when  $N_{g,k} = 1$  and  $N_{b,k} = 1$  as shown in Figure 5.7. The calculation of  $\Omega_{\mathbf{Q}_{k},1,1}\left(\Psi_{\mathfrak{r}_{k},\mathfrak{q}_{k}}\left(\mathbf{g}\left(x,\mathfrak{r}_{k},\mathcal{Q}_{k},\mathfrak{q}_{k}\left|t\right)\right)\right)$  is shown in Table 5.2.



Figure 5.7 The subcategories of Case 1 in Table 5.1

148		
	Sub-category 1	Sub-category 2
$\mathbf{f}(\boldsymbol{ au}_k)$	$f( au_k)$	$f( au_k)$
$\mathbf{f}'(\mathbf{z}_k)$	$\left(1-F\left(t-\tau_{k}\right)\right)$	$\left(1-F\left(t-\tau_{k}\right)\right)^{2}$
$\mathbf{s}(\mathcal{Q}_k)$	$s(O_k = 1)$	$s(O_k = 2)$
$\boldsymbol{\phi}(\boldsymbol{\mathfrak{Q}}_k)$	$\phi(lpha_{\scriptscriptstyle 1,k})$	$\phiig(lpha_{_{1,k}}ig)\phiig(lpha_{_{2,k}}ig)$
$\mathbf{h}ig(x t,  au_k,  ilde Q_k,  ilde Q_kig)$	$N\left(\begin{matrix}\mu_{k}t+\mu_{1,k}\left(t-\tau_{k}\right),\\\sigma_{k}^{2}t+\sigma_{1,k}^{2}\left(t-\tau_{k}\right)\end{matrix}\right)$	$N\begin{pmatrix} \mu_{k}t + \mu_{1,k}(t - \tau_{k}) \\ + \mu_{2,k}(t - \tau_{k}), \\ \sigma_{k}^{2}t + \sigma_{1,k}^{2}(t - \tau_{k}) \\ + \sigma_{2,k}^{2}(t - \tau_{k}) \end{pmatrix}$
$\mathbf{g}\left(x,\underline{\tau}_{k},\underline{O}_{k},\underline{\alpha}_{k}\left t\right.\right)$	Eq. (5.27)	Eq. (5.28)
$\Psi_{\mathbf{\mathfrak{T}}_{k},\mathbf{\mathfrak{Q}}_{k}}\left(\mathbf{g}\left(x,\mathbf{\mathfrak{T}}_{k},\mathbf{\mathcal{Q}}_{k},\mathbf{\mathfrak{Q}}_{k}\left t\right.\right)\right)$	Eq. (5.29)	Eq. (5.30)

Table 5.2	The cal	culation	steps	in	Case	1
			1			

$\boldsymbol{\Omega}_{\mathbf{Q}_{k},N_{b,k}=l,N_{g,k}=l}\left(\boldsymbol{\Psi}_{\mathbf{\tilde{t}}_{k},\mathbf{\tilde{q}}_{k}}\left(\mathbf{g}\right)\right)$	Eq. (5.31)
---	------------

The  $\mathbf{g}(x, \underline{\tau}_k, Q_k, \underline{\alpha}_k | t)$  and  $\Psi_{\underline{\tau}_k, \underline{q}_k}(\mathbf{g}(x, \underline{\tau}_k, Q_k, \underline{\alpha}_k | t))$  for the two sub-categories are obtained using Eqs. (5.27), (5.28) and Eqs. (5.29), (5.30)

$$\mathbf{g}(x, \underline{\tau}_{k,1}, \underline{Q}_{k,1}, \underline{\alpha}_{k,1} | t) = f(\tau_k) (1 - F(t - \tau_k)) s(O_k = 1) \phi(\alpha_{1,k}) N\left(x \begin{vmatrix} \mu_k t + \mu_{1,k}(t - \tau_k), \\ \sigma_k^2 t + \sigma_{1,k}^2(t - \tau_k) \end{vmatrix}\right)$$
(5.27)

$$\mathbf{g}(x, \tau_{k,2}, Q_{k,2}, \alpha_{k,2} | t) = f(\tau_k) (1 - F(t - \tau_k))^2 s(O_k = 2) \phi(\alpha_{1,k}) \phi(\alpha_{2,k}) N \begin{pmatrix} \mu_k t + \mu_{1,k} (t - \tau_k) \\ + \mu_{2,k} (t - \tau_k), \\ \sigma_k^2 t + \sigma_{1,k}^2 (t - \tau_k) \\ + \sigma_{2,k}^2 (t - \tau_k) \end{pmatrix}$$
(5.28)

$$\Psi_{\mathfrak{T}_{k},\mathfrak{Q}_{k}}\left(\mathbf{g}\left(x,\mathfrak{Z}_{k,1},\mathfrak{Q}_{k,1},\mathfrak{Q}_{k,1}\right|t\right)\right) = \int_{0}^{t} \int_{\alpha_{l}}^{\alpha_{u}} \mathbf{g}\left(x,\mathfrak{Z}_{k,1},\mathfrak{Q}_{k,1},\mathfrak{Q}_{k,1}\right|t\right) d\tau_{k} d\alpha_{1,k}$$
(5.29)

$$\Psi_{\mathbf{\mathfrak{T}}_{k},\mathbf{\mathfrak{g}}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k,2},\boldsymbol{\mathcal{Q}}_{k,2},\boldsymbol{\boldsymbol{\alpha}}_{k,2}\left|t\right.\right)\right) = \int_{0}^{t} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \mathbf{g}\left(x,\boldsymbol{\tau}_{k,2},\boldsymbol{\mathcal{Q}}_{k,2},\boldsymbol{\boldsymbol{\alpha}}_{k,2}\left|t\right.\right) d\tau_{k} d\alpha_{1,k} d\alpha_{2,k} \quad (5.30)$$

$$\boldsymbol{\Omega}_{\boldsymbol{Q}_{k},1,1}\left(\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{g}}_{k}}\left(\boldsymbol{g}\left(\boldsymbol{x},\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{\mathfrak{Q}}_{k},\boldsymbol{\mathfrak{g}}_{k}\left|\boldsymbol{t}\right\right)\right)\right) \\ = \boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{g}}_{k}}\left(\boldsymbol{g}\left(\boldsymbol{x},\boldsymbol{\tau}_{k,1},\boldsymbol{Q}_{k,1},\boldsymbol{\mathfrak{Q}}_{k,1}\right|\boldsymbol{t}\right)\right) + \boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{g}}_{k}}\left(\boldsymbol{g}\left(\boldsymbol{x},\boldsymbol{\tau}_{k,2},\boldsymbol{Q}_{k,2},\boldsymbol{\mathfrak{Q}}_{k,2}\right|\boldsymbol{t}\right)\right) \tag{5.31}$$

Case 2:

There are 4 sub-categories when  $N_{g,k} = 2$  and  $N_{b,k} = 2$  (sub-categories 3' and 4' are duplicates) as shown in Figure 5.8. The calculation of  $\Omega_{Q_k,2,2}\left(\Psi_{\mathfrak{r}_k,\mathfrak{q}_k}\left(\mathbf{g}(x,\mathfrak{r}_k,\mathcal{Q}_k,\mathfrak{q}_k|t)\right)\right)$  is shown in Table 5.3.



Figure 5.8 The subcategories of Case 2 in Table 5.2

	Sub-category 1	Sub-category 2
$\mathbf{f}(\mathbf{z}_k)$	$f( au_k)f( au_{1,k})$	$f( au_k)f( au_{1,k})$

 Table 5.3 The calculation steps in Case 2

$\mathbf{f}'(\boldsymbol{ au}_k)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)^{2}$
$\mathbf{s}(\mathcal{Q}_k)$	$s\left(O_{k}=1\right)s\left(O_{1,k}=1\right)$	$s(O_k = 1)s(O_{1,k} = 2)$
$\boldsymbol{\phi}(\boldsymbol{\mathfrak{Q}}_{k})$	$\phiig(lpha_{\scriptscriptstyle 1,k}ig)\phiig(lpha_{\scriptscriptstyle 1,1,k}ig)$	$\phiig(lpha_{\scriptscriptstyle 1,k}ig)\phiig(lpha_{\scriptscriptstyle 1,1,k}ig)\phiig(lpha_{\scriptscriptstyle 2,1,k}ig)$
$\mathbf{h}\left(x,\overline{z}_{k},Q_{k},\alpha_{k} t\right)$	$N \begin{pmatrix} \mu_{k}t + \mu_{1,k} (t - \tau_{k}) \\ + \mu_{1,1,k} (t - \tau_{k} - \tau_{1,k}), \\ \sigma_{k}^{2}t + \sigma_{1,k}^{2} (t - \tau_{k}) \\ + \sigma_{1,1,k}^{2} (t - \tau_{k} - \tau_{1,k}) \end{pmatrix}$	$N \begin{pmatrix} \mu_{k}t + \mu_{1,k} (t - \tau_{k}) \\ + (\mu_{1,1,k} + \mu_{2,1,k})(t - \tau_{k} - \tau_{1,k}), \\ \sigma_{k}^{2}t + \sigma_{1,k}^{2} (t - \tau_{k}) \\ (\sigma_{1,1,k}^{2} + \sigma_{2,1,k}^{2})(t - \tau_{k} - \tau_{1,k}) \end{pmatrix}$
$\mathbf{g}\left(x,\underline{\tau}_{k},\underline{O}_{k},\underline{\alpha}_{k}\left t\right.\right)$	Eq. (5.32)	Eq. (5.33)
$\Psi_{\mathbf{r}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{\mathcal{Q}}_{k},\boldsymbol{\boldsymbol{\alpha}}_{k}\left t\right.\right)\right)$	Eq. (5.34)	Eq. (5.35)
$\Omega_{\mathbf{Q}_{k},N_{b,k}=2,N_{g,k}=2}\left(\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\right)\right)$		

# Table 5.3 (Continued)

	Sub-category 3,3'	Sub-category 4,4'
$\mathbf{f}(\mathbf{z}_k)$	$f(\tau_k)f(\tau_{1,k})$	$f( au_k)f( au_{1,k})$
$\mathbf{f}'(\boldsymbol{ au}_k)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)^{2}$
$\mathbf{s}(\mathcal{Q}_k)$	$s(O_k = 2)s(O_{1,k} = 1)$	$s(O_k = 2)s(O_{1,k} = 2)$
$oldsymbol{\phi}(ec{lpha}_k)$	$\phi(\alpha_{\scriptscriptstyle 1,k})\phi(\alpha_{\scriptscriptstyle 2,k})\phi(\alpha_{\scriptscriptstyle 1,1,k})$	$\phi(\alpha_{1,k})\phi(\alpha_{2,k})\phi(\alpha_{1,1,k})\phi(\alpha_{2,1,k})$
$\mathbf{h}\left(x, \mathbf{z}_{k}, \mathbf{Q}_{k}, \mathbf{a}_{k}   t\right)$	$N\begin{pmatrix} \mu_{k}t \\ +(\mu_{1,k}+\mu_{2,k})(t-\tau_{k}) \\ +\mu_{1,1,k}(t-\tau_{k}-\tau_{1,k}), \\ \sigma_{k}^{2}t(t-\tau_{k}) \\ +(\sigma_{1,k}^{2}+\sigma_{2,k}^{2})(t-\tau_{k}) \\ +\sigma_{1,1,k}^{2}(t-\tau_{k}-\tau_{1,k}) \end{pmatrix}$	$N \begin{pmatrix} \mu_{k}t + (\mu_{1,k} + \mu_{2,k})(t - \tau_{k}) \\ + (\mu_{1,1,k} + \mu_{2,1,k})(t - \tau_{k} - \tau_{1,k}), \\ \sigma_{k}^{2}t + (\sigma_{1,k}^{2} + \sigma_{2,k}^{2})(t - \tau_{k}) \\ + (\sigma_{1,1,k}^{2} + \sigma_{2,1,k}^{2})(t - \tau_{k} - \tau_{1,k}) \end{pmatrix}$

$\mathbf{g}(x t, \underline{\tau}_k, \underline{O}_k, \underline{lpha}_k)$	Eq. (5.36)	Eq. (5.37)
$\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\mathbf{t}_{k},\mathbf{\mathcal{O}}_{k},\mathbf{\mathbf{\mathcal{Q}}}_{k}\right t\right)\right)$	Eq. (5.38)	Eq. (5.39)
$\Omega_{\mathbf{Q}_{k},N_{b,k}=2,N_{g,k}=2}\left(\Psi_{\mathbf{T}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\right)\right)$	Eq. (5.40)	

$$\mathbf{g}(x, \underline{\tau}_{k,1}, \underline{Q}_{k,1}, \underline{\alpha}_{k,1} | t) = \begin{pmatrix} f(\tau_{k}) f(\tau_{1,k}) (1 - F(t - \tau_{k} - \tau_{1,k})) \\ \times s(O_{k} = 1) s(O_{1,k} = 1) \phi(\alpha_{1,k}) \phi(\alpha_{1,1,k}) \end{pmatrix} N \begin{pmatrix} \mu_{k}t + \mu_{1,k} (t - \tau_{k}) \\ + \mu_{1,1,k} (t - \tau_{k} - \tau_{1,k}), \\ \sigma_{k}^{2}t + \sigma_{1,k}^{2} (t - \tau_{k}) \\ + \sigma_{1,1,k}^{2} (t - \tau_{k} - \tau_{1,k}) \end{pmatrix}$$
(5.32)

$$\mathbf{g}(x, \underline{\tau}_{k,2}, Q_{k,2}, \underline{\alpha}_{k,2} | t) = \begin{pmatrix} f(\tau_k) f(\tau_{1,k}) (1 - F(t - \tau_k - \tau_{1,k}))^2 \\ \times s(O_k = 1) s(O_{1,k} = 1) \phi(\alpha_{1,k}) \phi(\alpha_{1,1,k}) \end{pmatrix} N \begin{pmatrix} \mu_k t + \mu_{1,k} (t - \tau_k) \\ + \mu_{1,1,k} (t - \tau_k - \tau_{1,k}) \\ + \mu_{2,1,k} (t - \tau_k - \tau_{1,k}) \\ + \sigma_{1,1,k}^2 (t - \tau_k) \\ + \sigma_{2,1,k}^2 (t - \tau_k - \tau_{1,k}) \end{pmatrix}$$
(5.33)

$$\Psi_{\underline{\tau}_{k},\underline{q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k,1},\underline{Q}_{k,1},\underline{\alpha}_{k,1}|t\right)\right)$$
  
=  $\int_{0}^{t}\int_{0}^{t-\tau_{k}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\mathbf{g}\left(x,\underline{\tau}_{k,1},\underline{Q}_{k,1},\underline{\alpha}_{k,1}|t\right)d\tau_{k}d\tau_{1,k}d\alpha_{1,k}d\alpha_{1,l,k}$  (5.34)

$$\Psi_{\underline{\tau}_{k},\underline{q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k,2},\underline{Q}_{k,2},\underline{\alpha}_{k,2}|t\right)\right)$$

$$=\int_{0}^{t}\int_{0}^{t-\tau_{1}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\mathbf{g}\left(x,\underline{\tau}_{k,2},\underline{Q}_{k,2},\underline{\alpha}_{k,2}|t\right)d\tau_{k}d\tau_{1,k}d\alpha_{1,k}d\alpha_{1,l,k}d\alpha_{2,l,k}$$
(5.35)

$$\mathbf{g}(x, \underline{\tau}_{k,3}, \underline{Q}_{k,3}, \underline{\alpha}_{k,3} | t)$$

$$= \begin{pmatrix} f(\tau_{k}) f(\tau_{1,k}) (1 - F(t - \tau_{k} - \tau_{1,k})) \\ s(O_{k} = 2) s(O_{1,k} = 1) \phi(\alpha_{1,k}) \phi(\alpha_{2,k}) \phi(\alpha_{1,1,k}) \end{pmatrix} N \begin{pmatrix} \mu_{k} t \\ +(\mu_{1,k} + \mu_{2,k}) (t - \tau_{k}) \\ +\mu_{1,1,k} (t - \tau_{k} - \tau_{1,k}), \\ \sigma_{k}^{2} t(t - \tau_{k}) \\ +(\sigma_{1,k}^{2} + \sigma_{2,k}^{2}) (t - \tau_{k}) \\ +\sigma_{1,1,k}^{2} (t - \tau_{k} - \tau_{1,k}) \end{pmatrix}$$

$$(5.36)$$

$$\mathbf{g}(x, \underline{\tau}_{k,4}, \underline{Q}_{k,4}, \underline{\alpha}_{k,4} | t) = \begin{pmatrix} f(\tau_k) f(\tau_{1,k}) (1 - F(t - \tau_k - \tau_{1,k}))^2 \\ s(O_k = 2) s(O_{1,k} = 2) \phi(\alpha_{1,k}) \\ \phi(\alpha_{2,k}) \phi(\alpha_{1,1,k}) \phi(\alpha_{2,1,k}) \end{pmatrix} N \begin{pmatrix} \mu_k t + (\mu_{1,k} + \mu_{2,k}) (t - \tau_k) \\ + (\mu_{1,1,k} + \mu_{2,1,k}) (t - \tau_k - \tau_{1,k}), \\ \sigma_k^2 t + (\sigma_{1,k}^2 + \sigma_{2,k}^2) (t - \tau_k) \\ + (\sigma_{1,1,k}^2 + \sigma_{2,1,k}^2) (t - \tau_k - \tau_{1,k}) \end{pmatrix} (5.37)$$

$$\Psi_{\underline{\tau}_{k},\underline{q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k,3},\underline{\mathcal{Q}}_{k,3},\underline{\alpha}_{k,3}|t\right)\right)$$

$$=\int_{0}^{t}\int_{0}^{t-\tau_{k}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\mathbf{g}\left(x,\underline{\tau}_{k,3},\underline{\mathcal{Q}}_{k,3},\underline{\alpha}_{k,3}|t\right)d\tau_{k}d\tau_{1,k}d\alpha_{1,k}d\alpha_{2,k}d\alpha_{1,l,k}$$
(5.38)

$$\Psi_{\underline{\tau}_{k},\underline{q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k,4},\underline{\mathcal{Q}}_{k,4},\underline{\alpha}_{k,4}|t\right)\right)$$

$$=\int_{0}^{t}\int_{0}^{t-\tau_{k}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\mathbf{g}\left(x,\underline{\tau}_{k,4},\underline{\mathcal{Q}}_{k,4},\underline{\alpha}_{k,4}|t\right)d\tau_{k}d\tau_{1,k}d\alpha_{1,k}d\alpha_{2,k}d\alpha_{1,1,k}d\alpha_{2,1,k$$

$$\boldsymbol{\Omega}_{\mathbf{Q}_{k},2,2}\left(\boldsymbol{\Psi}_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\left|t\right)\right)\right) \\
= \boldsymbol{\Psi}_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k,1},\boldsymbol{Q}_{k,1},\boldsymbol{q}_{k,1}\left|t\right)\right) + \boldsymbol{\Psi}_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k,2},\boldsymbol{Q}_{k,2},\boldsymbol{q}_{k,2}\left|t\right)\right) \\
+ 2\boldsymbol{\Psi}_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k,3},\boldsymbol{Q}_{k,3},\boldsymbol{q}_{k,3}\left|t\right)\right) + 2\boldsymbol{\Psi}_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k,4},\boldsymbol{Q}_{k,4},\boldsymbol{q}_{k,4}\left|t\right)\right) \right) \tag{5.40}$$

Case 3:

There are 3 sub-categories when  $N_{g,k} = 2$  and  $N_{b,k} = 3$  (sub-categories 2' is a duplicate) as shown in Figure 5.9. The calculation of  $\Omega_{\mathbf{Q}_k,3,2}\left(\Psi_{\mathfrak{T}_k,\mathfrak{q}_k}\left(\mathbf{g}\left(x,\mathfrak{T}_k,\mathcal{Q}_k,\mathfrak{Q}_k|t\right)\right)\right)$  is shown in Table 5.4.



Figure 5.9 The subcategories of Case 3 in Table 5.3

	Sub-category 1	Sub-category 2,2'
$\mathbf{f}(\boldsymbol{ au}_k)$	$f(\tau_k)f(\tau_{1,k})f(\tau_{2,k})$	$f(\tau_k)f(\tau_{1,k})f(\tau_{2,k})$
$\mathbf{f}'(\mathbf{ au}_k)$	$\left(1-F\left(t- au_{k}- au_{1,k} ight) ight)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)$
	$\times \left(1 - F\left(t - \tau_k - \tau_{2,k}\right)\right)$	$\times \left(1 - F\left(t - \tau_k - \tau_{2,k}\right)\right)^2$
$\mathbf{s}(\mathcal{Q}_k)$	$s(O_k = 2)s(O_{1,k} = 1)$	$s\left(O_{k}=2\right)s\left(O_{1,k}=1\right)$
	$\times s(O_{2,k} = 1)$	$\times s(O_{2,k} = 2)$
$\phi(lpha_k)$	$\phi(lpha_{\scriptscriptstyle 1,k})\phi(lpha_{\scriptscriptstyle 2,k})$	$\phi(\alpha_{\scriptscriptstyle 1,k})\phi(\alpha_{\scriptscriptstyle 2,k})\phi(\alpha_{\scriptscriptstyle 1,1,k})$
	$\times \phi\bigl(\alpha_{\scriptscriptstyle 1,1,k}\bigr) \phi\bigl(\alpha_{\scriptscriptstyle 1,2,k}\bigr)$	$\times \phi\bigl(\alpha_{\scriptscriptstyle 1,2,k}\bigr) \phi\bigl(\alpha_{\scriptscriptstyle 2,2,k}\bigr)$

Table 5.4 The calculation steps in Case 3

$\mathbf{h}\left(x, \underline{\tau}_{k}, Q_{k}, \underline{\alpha}_{k}   t\right)$	$N \begin{pmatrix} \mu_{k}t + \mu_{1,k} (t - \tau_{k}) \\ + \mu_{2,k} (t - \tau_{k}) \\ + \mu_{1,1,k} (t - \tau_{k} - \tau_{1,k}) \\ + \mu_{1,2,k} (t - \tau_{k} - \tau_{2,k}), \\ \sigma_{k}^{2}t + \sigma_{1,k}^{2} (t - \tau_{k}) \\ + \sigma_{2,k}^{2} (t - \tau_{k}) \\ + \sigma_{1,1,k}^{2} (t - \tau_{k} - \tau_{1,k}) \\ + \sigma_{1,2,k}^{2} (t - \tau_{k} - \tau_{2,k}) \end{pmatrix}$	$N \begin{pmatrix} \mu_{k}t + (\mu_{1,k} + \mu_{2,k})(t - \tau_{k}) \\ + \mu_{1,1,k}(t - \tau_{k} - \tau_{1,k}) \\ + (\mu_{1,2,k} + \mu_{2,2,k})(t - \tau_{k} - \tau_{2,k}), \\ \sigma_{k}^{2}t + (\sigma_{1,k}^{2} + \sigma_{2,k}^{2})(t - \tau_{k}) \\ + \sigma_{1,1,k}^{2}(t - \tau_{k} - \tau_{1,k}) \\ + (\sigma_{1,2,k}^{2} + \sigma_{2,2,k}^{2})(t - \tau_{k} - \tau_{2,k}) \end{pmatrix}$	
$\mathbf{g}\left(x, \underline{\tau}_{k}, \underline{Q}_{k}, \underline{Q}_{k} \left  t \right.  ight)$	Eq. (5.31)	Eq. (5.32)	
$\Psi_{\mathtt{I}_{k}, \mathtt{Q}_{k}}\left( \mathbf{g}\left(x, \mathtt{Z}_{k}, \mathtt{Q}_{k}, \mathtt{Q}_{k}   t\right) \right)$	Eq. (5.34)	Eq. (5.35)	
$\Omega_{\mathbf{Q}_{k},N_{b,k}=3,N_{g,k}=2}\left(\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\right)\right)$	Eq. (5.37)		

# Table 5.4 (Continued)

	Sub-category 3
$\mathbf{f}(\boldsymbol{\mathfrak{t}}_k)$	$f(\tau_k)f(\tau_{1,k})f(\tau_{2,k})$
$\mathbf{f}'(\mathbf{z}_k)$	$\left(1-F\left(t-\tau_{k}-\tau_{1,k}\right)\right)^{2}\times\left(1-F\left(t-\tau_{k}-\tau_{2,k}\right)\right)^{2}$
$\mathbf{s}(\mathcal{Q}_k)$	$s(O_k = 2)s(O_{1,k} = 2)$ ×s(O <sub>2,k</sub> = 2)
$\mathbf{\phi}(\mathbf{ ilde{a}}_k)$	$\phi(lpha_{{\scriptscriptstyle 1},k})\phi(lpha_{{\scriptscriptstyle 2},k})\phi(lpha_{{\scriptscriptstyle 1},1,k}) \  imes \phi(lpha_{{\scriptscriptstyle 2},1,k})\phi(lpha_{{\scriptscriptstyle 1},2,k})\phi(lpha_{{\scriptscriptstyle 2},2,k})$

$\mathbf{h}\left(x, \underline{\tau}_{k}, \underline{\mathcal{Q}}_{k}, \underline{\mathbf{\alpha}}_{k}   t\right)$	$N\begin{pmatrix} \mu_{k}t + (\mu_{1,k} + \mu_{2,k})(t - \tau_{k}) \\ + (\mu_{1,1,k} + \mu_{2,1,k})(t - \tau_{k} - \tau_{1,k}) \\ + (\mu_{1,2,k} + \mu_{2,2,k})(t - \tau_{k} - \tau_{2,k}), \\ \sigma_{k}^{2}t + (2\sigma_{1,k}^{2} + 2\sigma_{2,k}^{2})(t - \tau_{k}) \\ + (\sigma_{1,1,k}^{2} + \sigma_{2,1,k}^{2})(t - \tau_{k} - \tau_{1,k}) \\ + (\sigma_{1,2,k}^{2} + \sigma_{2,2,k}^{2})(t - \tau_{k} - \tau_{2,k}) \end{pmatrix}$
$\mathbf{g}\left(x, \mathbf{z}_{k}, \mathbf{\mathcal{Q}}_{k}, \mathbf{\mathbf{\mathcal{Q}}}_{k} \left  t  ight)$	Eq. (5.33)
$\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\mathbf{t}_{k},\mathcal{Q}_{k},\mathbf{q}_{k}\left t\right.\right)\right)$	Eq. (5.36)
$\Omega_{\mathbf{Q}_{k},N_{b,k}=3,N_{g,k}=2}\left(\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\right)\right)$	Eq. (5.37)

$$\mathbf{g}\left(x, \underline{\tau}_{k,1}, Q_{k,1}, \underline{\alpha}_{k,1} \middle| t\right) = \begin{pmatrix} f\left(\tau_{k}\right) f\left(\tau_{1,k}\right) f\left(\tau_{2,k}\right) \left(1 - F\left(t - \tau_{k} - \tau_{1,k}\right)\right) \\ \left(1 - F\left(t - \tau_{k} - \tau_{2,k}\right)\right) s\left(O_{k} = 2\right) s\left(O_{1,k} = 1\right) \\ s\left(O_{2,k} = 1\right) \phi\left(\alpha_{1,k}\right) \phi\left(\alpha_{2,k}\right) \phi\left(\alpha_{1,1,k}\right) \phi\left(\alpha_{1,2,k}\right) \end{pmatrix} \\ N \begin{pmatrix} \mu_{k} t + \left(\mu_{1,k} + \mu_{2,k}\right) \left(t - \tau_{k}\right) \\ + \mu_{1,1,k} \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \mu_{1,2,k} \left(t - \tau_{k} - \tau_{2,k}\right), \\ \sigma_{k}^{2} t + \left(\sigma_{1,k}^{2} + \sigma_{2,k}^{2}\right) \left(t - \tau_{k}\right) \\ + \sigma_{1,1,k}^{2} \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \sigma_{1,2,k}^{2} \left(t - \tau_{k} - \tau_{2,k}\right) \end{pmatrix}$$
(5.41)

$$\mathbf{g}\left(x, \underline{\tau}_{k,2}, Q_{k,2}, \underline{\alpha}_{k,2} | t\right) \\
= \begin{pmatrix} f\left(\tau_{k}\right) f\left(\tau_{1,k}\right) f\left(\tau_{2,k}\right) \\ \left(1 - F\left(t - \tau_{k} - \tau_{1,k}\right)\right) \\ \left(1 - F\left(t - \tau_{k} - \tau_{2,k}\right)\right)^{2} s\left(O_{k} = 2\right) \\ s\left(O_{1,k} = 1\right) s\left(O_{2,k} = 2\right) \phi\left(\alpha_{1,k}\right) \\ \phi\left(\alpha_{2,k}\right) \phi\left(\alpha_{1,1,k}\right) \phi\left(\alpha_{1,2,k}\right) \phi\left(\alpha_{2,2,k}\right) \end{pmatrix} \\
N \begin{pmatrix} \mu_{k} t + \left(\mu_{1,k} + \mu_{2,k}\right) \left(t - \tau_{k}\right) \\ + \mu_{1,1,k} \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \left(\mu_{1,2,k} + \mu_{2,2,k}\right) \left(t - \tau_{k} - \tau_{2,k}\right), \\ \sigma_{k}^{2} t + \left(\sigma_{1,k}^{2} + \sigma_{2,k}^{2}\right) \left(t - \tau_{k}\right) \\ + \sigma_{1,1,k}^{2} \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \left(\sigma_{1,2,k}^{2} + \sigma_{2,2,k}^{2}\right) \left(t - \tau_{k} - \tau_{2,k}\right) \end{pmatrix}$$
(5.42)

$$\mathbf{g}\left(x, \underline{\tau}_{k,3}, Q_{k,3}, \underline{\alpha}_{k,3} | t\right) = \begin{pmatrix} f\left(\tau_{k}\right) f\left(\tau_{1,k}\right) f\left(\tau_{2,k}\right) \\ \left(1 - F\left(t - \tau_{k} - \tau_{1,k}\right)\right)^{2} \\ \left(1 - F\left(t - \tau_{k} - \tau_{2,k}\right)\right)^{2} s\left(O_{k} = 2\right) \\ s\left(O_{1,k} = 2\right) s\left(O_{2,k} = 2\right) \phi\left(\alpha_{1,k}\right) \phi\left(\alpha_{2,2,k}\right) \end{pmatrix} \\ N \begin{pmatrix} \mu_{k}t + \left(\mu_{1,k} + \mu_{2,1,k}\right) \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \left(\mu_{1,2,k} + \mu_{2,2,k}\right) \left(t - \tau_{k} - \tau_{2,k}\right), \\ \sigma_{k}^{2}t + \left(\sigma_{1,k}^{2} + \sigma_{2,k}^{2}\right) \left(t - \tau_{k}\right) \\ + \left(\sigma_{1,1,k}^{2} + \sigma_{2,1,k}^{2}\right) \left(t - \tau_{k} - \tau_{1,k}\right) \\ + \left(\sigma_{1,2,k}^{2} + \sigma_{2,2,k}^{2}\right) \left(t - \tau_{k} - \tau_{2,k}\right) \end{pmatrix}$$
(5.43)

$$\Psi_{\mathfrak{x}_{k},\mathfrak{g}_{k}}\left(\mathbf{g}\left(x,\mathfrak{z}_{k,1},\mathcal{Q}_{k,1},\mathfrak{Q}_{k,1},\mathfrak{q}_{k,1}|t\right)\right)$$

$$=\int_{0}^{t}\int_{0}^{t-\tau_{k}}\int_{0}^{t-\tau_{k}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\mathbf{g}\left(x,\mathfrak{z}_{k,1},\mathcal{Q}_{k,1},\mathfrak{q}_{k,1}|t\right)d\tau_{k}d\tau_{1,k}d\tau_{2,k}d\alpha_{1,k}d\alpha_{2,k}d\alpha_{1,1,k}d\alpha_{1,2,k}$$

$$(5.44)$$

$$\Psi_{\underline{\tau}_{k},\underline{q}_{k}}\left(\mathbf{g}\left(x,\underline{\tau}_{k,2},Q_{k,2},\underline{q}_{k,2}|t\right)\right)$$

$$=\int_{0}^{t}\int_{0}^{t-\tau_{k}}\int_{0}^{t-\tau_{k}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}\int_{\alpha_{l}}^{\alpha_{u}}g\left(x,\underline{\tau}_{k,2},Q_{k,2},\underline{q}_{k,2}|t\right)$$

$$\times d\tau_{k}d\tau_{1,k}d\tau_{2,k}d\alpha_{1,k}d\alpha_{2,k}d\alpha_{1,1,k}d\alpha_{1,2,k}d\alpha_{2,2,k}$$
(5.45)

$$\Psi_{\mathbf{t}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\mathbf{t}_{k,3},\mathbf{Q}_{k,3},\mathbf{x}_{k,3}|t\right)\right) = \int_{0}^{t} \int_{0}^{t-\tau_{k}} \int_{0}^{t-\tau_{1,k}} \int_{0}^{t-\tau_{2,k}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \int_{\alpha_{l}}^{\alpha_{u}} \mathbf{g}\left(x,\mathbf{t}_{k,3},\mathbf{Q}_{k,3},\mathbf{x}_{k,3}|t\right)$$

$$\times d\tau_{k} d\tau_{1,k} d\tau_{2,k} d\tau_{1,1,k} d\tau_{2,1,k} d\tau_{1,2,k} d\tau_{2,2,k} d\alpha_{1,k} d\alpha_{2,k} d\alpha_{1,1,k} d\alpha_{2,1,k} d\alpha_{2,2,k} d\alpha_{2,2,k}$$
(5.46)

$$\boldsymbol{\Omega}_{\boldsymbol{Q}_{k},3,2}\left(\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{Q}_{k},\boldsymbol{Q}_{k},\boldsymbol{\mathfrak{q}}_{k}\left|t\right)\right)\right) \\ = \boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{I}}_{k,1},\boldsymbol{Q}_{k,1},\boldsymbol{\mathfrak{q}}_{k,1}\left|t\right)\right) + 2\boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{I}}_{k,2},\boldsymbol{Q}_{k,2},\boldsymbol{\mathfrak{q}}_{k,2}\left|t\right)\right) \\ + \boldsymbol{\Psi}_{\boldsymbol{\mathfrak{I}}_{k},\boldsymbol{\mathfrak{q}}_{k}}\left(\boldsymbol{g}\left(x,\boldsymbol{\mathfrak{I}}_{k,3},\boldsymbol{Q}_{k,3},\boldsymbol{\mathfrak{q}}_{k,3}\left|t\right)\right)\right) \qquad (5.47)$$

The PDF of the degradation of tree k is

$$\tilde{h}_{k}\left(x|t\right) = \Omega_{\mathbf{Q}_{k},1,1}\left(\Psi_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\left|t\right)\right)\right) + \Omega_{\mathbf{Q}_{k},2,2}\left(\Psi_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\left|t\right)\right)\right) + \Omega_{\mathbf{Q}_{k},2,2}\left(\Psi_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\left|t\right)\right)\right)\right)$$

$$+ \Omega_{\mathbf{Q}_{k},3,2}\left(\Psi_{\mathbf{I}_{k},\mathbf{q}_{k}}\left(\mathbf{g}\left(x,\boldsymbol{\tau}_{k},\boldsymbol{Q}_{k},\boldsymbol{q}_{k}\left|t\right)\right)\right)$$

$$(5.48)$$

#### 5.4.3 Comparison of the Simulations and the Theoretical Results

In this section, we show the comparison of the numerical solutions and the simulation results of the case study. The details of the simulation steps are shown in 5.4.3.1 and the comparison results are shown in 5.4.3.2.

#### 5.4.3.1 The simulation steps of GDDBP

In each realization of the GDDBP, the following steps are performed:

- (1) Randomly generate the total number of trees  $n_{tr}$
- (2) Let k = 1
- (3) Grow tree k. For each branch in tree k that has not branched, let τ<sub>0</sub> denote the initiation time of this branch. Generate a random time τ, if τ<sub>0</sub> + τ < t a random number of O branches is initiated. Otherwise, no further branchings occur until time t. For each newly initiated branch, generate an angle α and the growth of this branch from time τ<sub>0</sub> + τ to time t follows a Brownian motion with drift μ<sub>0</sub>(η<sub>0</sub> + η<sub>1</sub>α) and diffusion σ<sub>0</sub>(η<sub>0</sub> + η<sub>1</sub>α)<sup>0.5</sup> where μ<sub>0</sub> and σ<sub>0</sub> are its parent's drift and diffusion respectively. Repeat these procedures for all the branches and their offsprings.

- (4) Let k = k + 1, if  $k \le n_{tr}$ , go to step (3). Otherwise, go to step (5).
- (5) Record the sum of the degradation of all the branches.

## 5.4.3.2 Comparison of the numerical and the simulations results

Figure 5.10 shows the PDF of the total degradation obtained by the numerical solutions and the histogram of the simulated total degradation at t = 10. As shown in Figure 5.10, they are closely matched. Figure 5.11 shows a typical realization of the simulated branched trees. We observe that by time t = 10 different trees show different growth patterns as expected: (1) some trees branch once while others branch more than once, (2) some branchings initiate one branch while others initiate two branches.



Figure 5.10 The comparison of the PDF of total degradation obtained by numerical

solution and histogram of the simulations at t = 10



Figure 5.11 A typical simulation of the branched trees

Figure 5.12 shows the evolution of the PDF of the total degradation of all the branches from t=1 to t=10. We observe that as time increases the PDFs become flat, indicating that the variance of the total degradation increases over time. This can also be concluded from Figure 5.13, which shows the means and variances of the total degradation from t=1to t=10 obtained by the simulations and the numerical solutions respectively. It is observed that the numerical solutions accurately match the simulation results. As expected, the mean and variance increase over time. Compared with the growth of the mean of total degradation, the growth of variance increases more sharply due to the uncertainty of the total number of branches and the variance of degradation of the branches.



Figure 5.12 The PDFs of the total degradation over time obtained by numerical solution



Figure 5.13 The means and variances obtained with simulations and numerical solutions



Figure 5.14 The reliability of the system obtained by numerical solution and simulations

When the failure threshold c = 40, The reliability functions of the system obtained by simulations and numerical solutions are shown in Figure 5.14. The results obtained by the numerical solutions accurately match the simulation results.

# 5.5 Summary

We develop a stochastic model that characterizes the GDDBP. The model is general because it captures the randomness of the number of branches and their angles. The relationship between the degradation rate and the angles is also investigated. The statistical properties of the process, such as the PDF of the degradation and the related reliability metrics such as failure time distribution and reliability function are presented. The model is validated with simulations.

#### **CHAPTER 6**

#### **CORROSION MONITORING AND PREDICTION**

#### 6.1 Introduction

In the corrosion monitoring process, under the conditions where missing data exist, missing data interpolation approaches are provided to interpolate them. We provide three approaches for data interpolation: the nonlinear Brownian bridge, the nonlinear Gamma Bridge and the inverse Gaussian Bridge when the underlying degradation paths follow Brownian motion processes, Gamma process and inverse Gaussian process respectively. The stochastic bridges capture the nonlinearity and uncertainties of the degradation processes. When multiple sensors are used, we propose a model that integrates observations from multiple sensors. The data integration approach is nonparametric and considers the dynamic clustering of sensors' observations. Sensor weights (contribution) are determined based on its performance in the previous several time instants and the variance (if the data is interpolated with stochastic bridges and bootstrap). It utilizes a moving time window to capture the switching of the sensors between clusters with time so that the weights are adjusted accordingly. The models are validated with a real crack growth dataset.

#### 6.1 Missing Data Interpolation Models

Three stochastic bridges are provided in this section to interpolate missing data in the degradation process. They are nonlinear Brownian Bridge, nonlinear Gamma Bridge and inverse Gaussian Bridge.

#### 6.1.1 Nonlinear Brownian Bridge

In this section, we develop a nonlinear Brownian bridge to interpolate missing data when the underlying process is a nonlinear Brownian motion. For simplicity, we call the data that are missing in the degradation process as "missing data" and those that are observed as "observed data."

#### 6.1.1.1 Standard Brownian Bridge

Suppose *a* and *b* are the degradation values of sensor *i* at times  $t_{\tau}$  and  $t_{m}$  respectively, where  $t_{\tau} < t_{m}$ . The degradation data between times  $t_{\tau}$  and  $t_{m}$  are missing. Let  $\{B_{i}\}$  be a one-dimensional Brownian motion. The process  $BB_{i}^{a\to b}(t)$  is defined as the standard Brownian bridge (BB) of sensor *i* from *a* to *b* between times  $t_{\tau}$  and  $t_{m}$  (Shreve, 2004) is expressed as

$$BB_{i}^{a \to b}(t) = a + \frac{t - t_{\tau}}{T}(b - a) + \left(B_{t - t_{\tau}} - \frac{t - t_{\tau}}{T}B_{t_{m} - t_{\tau}}\right), \ t_{\tau} \le t \le t_{m}$$
(6.1)

where  $T = t_m - t_\tau$ . The expectation of the standard Brownian Bridge is

$$E\left[BB_{i}^{a\to b}\left(t\right)\right] = a + \frac{t - t_{\tau}}{T}(b - a), \ t_{\tau} \le t \le t_{m}$$

$$(6.2)$$

Let  $t_{\tau} \leq s, t \leq t_m$ , the covariance of the two missing values at times s and t within the bridge is

$$c_{i}^{a \to b}\left(s, t\right) = E\left[\left(BB_{i}^{a \to b}\left(s\right) - m_{i}^{a \to b}\left(s\right)\right)\left(BB_{i}^{a \to b}\left(t\right) - m_{i}^{a \to b}\left(t\right)\right)\right]$$
  
$$= \min\left\{s - t_{\tau}, t - t_{\tau}\right\} - \frac{\left(s - t_{\tau}\right)\left(t - t_{\tau}\right)}{T}, t_{\tau} \le s, t \le t_{m}$$
(6.3)

 $BB_{i}^{a \to b}(t) \text{ follows a normal distribution with mean } a + \frac{t - t_{\tau}}{T}(b - a) \text{ and variance}$  $t - t_{\tau} - \frac{(t - t_{\tau})^{2}}{T} \text{ as follows}$  $BB_{i}^{a \to b}(t) \sim N\left(a + \frac{t - t_{\tau}}{T}(b - a), t - t_{\tau} - \frac{(t - t_{\tau})^{2}}{T}\right), t_{\tau} \leq t \leq t_{m}$ (6.4)

where  $N(\cdot)$  denotes normal distribution. However, the volatility of the data is not captured in the standard Brownian Bridge. Moreover, according to Eq. (6.2), it is assumed that the mean degradation path is a linear function of time, which is impractical in many situations (McCuen *et al.*, 1992, Soares *et al.*, 2009). We modify the standard Brownian Bridge so that the mean function captures nonlinearity of the degradation processes and the variance of interpolated values is consistent with observed data. A piecewise linear function is proposed to capture the nonlinearity of degradation mean path. Let  $k_j$  denote the slope of the line segment between time  $t_j$  and  $t_{j+1}$  as shown in Figure 6.1. Because the sum of the projections of all these line segments on the vertical direction is b-a, we have

$$a + \sum_{j=\tau}^{m-1} k_j \Delta t = b \tag{6.5}$$

We assume that the degradation data are observed at discrete times with equal time interval length  $\Delta t$ .



Figure 6.1 Comparison of the piecewise linear mean function and linear mean function

We also assume that the slopes of the segments of the degradation path change gradually and linearly with time, as described in Eq. (6.6).

$$k_{i} = c_{0} + c_{1}t_{i} \tag{6.6}$$

where  $c_0$  and  $c_1$  are constants. Substituting Eq. (6.6) into Eq. (6.5) results in

$$a + \sum_{j=\tau}^{m-1} (c_0 + c_1 t_j) \Delta t = b$$
(6.7)

 $c_{\scriptscriptstyle 1}\,$  is estimated by considering the observed data. Using least squares, the estimate  $\,\hat{c}_{\scriptscriptstyle 1}\,$  is

$$\hat{c}_{1} = \frac{k_{\tau-1} - k_{0} + k_{M-1} - k_{m}}{\Delta t \left(\tau + M - m - 2\right)}$$
(6.8)

where M is time index of the last observation in the time frame being considered.  $\hat{c}_0$  is obtained with

$$\hat{c}_0 = \frac{b - a - \hat{c}_1 \Delta t \sum_{j=\tau}^{m-1} t_j}{(m - \tau) \Delta t}$$
(6.9)

Figure 6.2 is a schematic diagram of the slopes and observation times, where missing data are shown with white diamonds.



Figure 6.2 Estimate the parameter  $c_1$  by considering the observed data

The predicted slopes of the mean path segment between time  $t_j$  and  $t_{j+1}$  is

$$\hat{k}_{t} = (\hat{c}_{0} + \hat{c}_{1}t), \ t_{j} < t \le t_{j+1}$$
(6.10)

The mean of the nonlinear Brownian bridge is

$$m_{i}^{a \to b}(t) = \begin{cases} a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})(t - t_{\tau}), t_{\tau} < t \le t_{\tau+1} \\ a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})\Delta t + (\hat{c}_{0} + \hat{c}_{1}t_{\tau+1})(t - t_{\tau+1}), t_{\tau+1} < t \le t_{\tau+2} \\ \vdots \\ a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})\Delta t + \dots + (\hat{c}_{0} + \hat{c}_{1}t_{m-1})(t - t_{m-1}), t_{m-1} < t \le t_{m} \end{cases}$$
(6.11)

The proposed nonlinear Brownian bridge from a to b between time  $t_{\tau}$  and  $t_m$  is

$$BB_{i}^{a \to b}\left(t\right) = m_{i}^{a \to b}\left(t\right) + \sigma\left(B_{t-t_{\tau}} - \frac{t-t_{\tau}}{T}B_{t_{m}-t_{\tau}}\right), \ t_{\tau} \le t \le t_{m}$$
(6.12)

where  $T = t_m - t_r$ ,  $\sigma$  is the parameter that controls the volatility, which is estimated as follows: assuming the underlying degradation process is a nonlinear Brownian motion with time-dependent drift  $\hat{k}_t$  and diffusion  $\sigma$  as shown in Eq. (6.13), the diffusion  $\sigma$  is estimated with the observed data using the Maximum Likelihood Estimation (MLE) method.

$$dD_i(t) = \hat{k}_i dt + \sigma dB_i \tag{6.13}$$

Note that  $m_i^{a\to b}(t_{\tau}) = a$ ,  $m_i^{a\to b}(t_m) = b$ . The variance of the nonlinear Brownian Bridge is

$$Var\left(BB_{i}^{a\to b}\left(t\right)\right) = \sigma^{2}\left(t - t_{\tau} - \frac{\left(t - t_{\tau}\right)^{2}}{T}\right), \ t_{\tau} \le t \le t_{m}$$
(6.14)

From Eq. (6.14), we observe the variance first increases from one end of the "bridge" with a starting value zero until it reaches the peak in the middle of the bridge and then decreases until it reaches the other end of the bridge with a value zero. Figure 6.3 shows a schematic plot of the variance of a nonlinear Brownian bridge with  $\tau = 5$ , m = 11,  $t_{\tau} = 5$ ,  $t_m = 11$  and  $\sigma = 1$ . This property is consistent with the intuition that missing data close to the observed data in the time frame are easier and more accurate to predict, while those far from the observed data are less accurate and difficult to predict.



Figure 6.3 The plot of the variance of a nonlinear Brownian bridge

The interpolated data with a nonlinear Brownian bridge are correlated and follow a multivariate Normal distribution. The interpolated data at different times are correlated. Let  $t_x \leq s$ ,  $t \leq t_m$ , the covariance function is expressed as

$$c_{i}^{a \to b}\left(s, t\right) = E\left[\left(BB_{i}^{a \to b}\left(s-t_{\tau}\right)-m_{i}^{a \to b}\left(s-t_{\tau}\right)\right)\left(BB_{i}^{a \to b}\left(t-t_{\tau}\right)-m_{i}^{a \to b}\left(t-t_{\tau}\right)\right)\right]$$
  
$$= \sigma^{2}\left(\min\left\{s-t_{\tau}, t-t_{\tau}\right\}-\frac{\left(s-t_{\tau}\right)\left(t-t_{\tau}\right)}{T}\right)$$
(6.15)

#### 6.1.2 Nonlinear Gamma Bridge

The proposed nonlinear Brownian Bridge in section 6.1.1 is useful especially when degradation increments are Gaussian distributed. However, there are plenty of cases in real life where degradation increments are positive. Better alternatives in this situation are the nonlinear Gamma Bridge and the inverse Gaussian Bridge. We discuss the development of nonlinear Gamma Bridge in this section by reviewing the standard Gamma Bridge.

#### 6.1.2.1 Standard Gamma Bridge

Assume the degradation process follows a Gamma process with mean  $\mu t$  and variance  $\mu t$  (scale parameter=1). The process  $GB_i^{a\to b}(t)$  is defined as the Gamma bridge of sensor i between time  $t_{\tau}$  and  $t_m$ . The problem now turns to find the probability density function of  $\{D_i(t)\}, t_{\tau} \leq t \leq t_m$  under the condition that  $D_i(t_{\tau}) = a, D_i(t_m) = b$ . Let X(t) denote the degradation increment between  $t_{\tau}$  and t, Y(t) denote the increment between t and

$$t_m$$
, Z denote an increment between  $t_\tau$  and  $t_m$ . Let  
 $X(t) = D_i(t) - D_i(t_\tau) \sim Ga(\mu(t - t_\tau), 1), t_\tau \le t \le t_m,$   
 $Y(t) = D_i(t_m) - D_i(t) \sim Ga(\mu(t_m - t), 1), t_\tau \le t \le t_m, Z = X(t) + Y(t)$ . According to  
 $Y(t)$ 

Ribeiro *et al.* (2004),  $\frac{X(t)}{Z}$  follows a Beta distribution with

$$\frac{X(t)}{Z} \sim B(\mu(t-t_{\tau}), \ \mu(t_m-t)), \ t_{\tau} \le t \le t_m$$
(6.16)



Figure 6.4 The schematic diagram of a Gamma Bridge

#### 6.1.2.2 Nonlinear Gamma Bridge

However, for standard Gamma Bridge, the underlying mean function is assumed to be linear, which is not useful in missing data interpolation, especially in the nonlinear degradation process. Moreover, the scale parameter of the Gamma process is 1, which is not general in applications. We propose a nonlinear Gamma bridge that captures nonlinearity of the missing data based on the Gamma process.

The mean function  $m_i^{a \to b}(t)$  is the same as that in section 6.1.1. Rather than modeling the degradation as a standard Gamma process, we let  $X(t) = D_i(t) - D_i(t_\tau) \sim Ga\left(\frac{m_i^{a \to b}(t) - a}{\theta}, \theta\right)$ , ,  $Y(t) = D_i(t_m) - D_i(t) \sim Ga\left(\frac{b - m_i^{a \to b}(t)}{\theta}, \theta\right)$ , Z = X(t) + Y(t), where  $\theta$  is a scale

parameter. Given the starting time  $t_{\tau}$  and the degradation value a, the ending time  $t_m$ and degradation value b, the conditional probability density function of X(t)|Z is derived as follows

$$\begin{split} f_{X(t)|Z}(x(t)|z) &= P(D_{i}(t) - D_{i}(t_{\tau}) = x(t)|D_{i}(t_{m}) - D_{i}(t_{\tau}) = z) \\ &= \frac{P(D_{i}(t) - D_{i}(t_{\tau}) = x(t), D_{i}(t_{m}) - D_{i}(t_{\tau}) = z)}{P(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z)} \\ &= \frac{P(D_{i}(t) - D_{i}(t_{\tau}) = x(t), D_{i}(t_{m}) - D_{i}(t) = z - x(t))}{P(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z)} \\ &= \frac{P(D_{i}(t) - D_{i}(t_{\tau}) = x(t))P(D_{i}(t_{m}) - D_{i}(t) = z - x(t))}{P(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z)} \\ &= \frac{x(t)^{\frac{m_{i}^{a \to b}(t) - a}{\theta}} \exp\left(-\frac{x(t)}{\theta}\right)}{\Gamma\left(\frac{m_{i}^{a \to b}(t) - a}{\theta}\right)} \frac{(z - x(t))^{\frac{b - m_{i}^{a \to b}(t)}{\theta} - 1}}{\Gamma\left(\frac{b - m_{i}^{a \to b}(t)}{\theta}\right)} \frac{1}{P(t)^{\frac{b - m_{i}^{a \to b}(t)}{\theta}}} \frac{$$

We get

$$f_{\frac{X(t)}{Z}|z}\left(\frac{x(t)}{z}|z\right) = \frac{\Gamma\left(\frac{b-a}{\theta}\right)}{\Gamma\left(\frac{m_i^{a\to b}(t)-a}{\theta}\right)\Gamma\left(\frac{b-m_i^{a\to b}(t)}{\theta}\right)} \left(\frac{x(t)}{z}\right)^{\frac{m_i^{a\to b}(t)-a}{\theta}-1} \left(1-\frac{x(t)}{z}\right)^{\frac{b-m_i^{a\to b}(t)}{\theta}-1}$$

$$(6.18)$$

which can be proved as follows:

$$P(Y \le y) = P(cX \le y)$$
  
=  $P\left(X \le \frac{y}{c}\right) = F_{X}\left(\frac{y}{c}\right)$  (6.19)

$$f_Y(y) = \frac{1}{c} f_X\left(\frac{y}{c}\right) \tag{6.20}$$

Change the variables to 
$$\beta_i(t) = \left(\frac{X(t)}{Z}|Z\right)$$
, then

$$\beta_i(t) = \left(\frac{X(t)}{Z}|Z\right) \sim B\left(\frac{m_i^{a \to b}(t) - a}{\theta}, \frac{b - m_i^{a \to b}(t)}{\theta}\right) \text{ has a beta distribution with}$$

parameters  $\frac{m_i^{a \to b}(t) - a}{\theta}$  and  $\frac{b - m_i^{a \to b}(t)}{\theta}$ . Given  $D_i(t_\tau) = a, D_i(t_m) = b$ , the

nonlinear Gamma bridge is defined as

$$GB_i^{a \to b}(t) = a + \beta_i(t)(b - a)$$
(6.21)

where  $\beta_i(t) \sim B\left(\frac{m_i^{a \to b}(t) - a}{\theta}, \frac{b - m_i^{a \to b}(t)}{\theta}\right), \theta$  is estimated from observed data as

follows: assuming the degradation process is a nonlinear Gamma process with time-

dependent shape parameter  $\frac{m_i^{a\to b}(t_j) - m_i^{a\to b}(t_{j-1})}{\theta}$  and scale parameter  $\theta$ , as shown in Eq. (6.23), the scale parameter  $\theta$  is estimated with observed data using MLE. Eq. (6.22) shows the mean function of the nonlinear Gamma process of the observed degradation

process.

$$m_{i}^{a \to b}(t) = \begin{cases} a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})(t - t_{\tau}), t_{\tau} < t \le t_{\tau+1} \\ a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})\Delta t + (\hat{c}_{0} + \hat{c}_{1}t_{\tau+1})(t - t_{\tau+1}), t_{\tau+1} < t \le t_{\tau+2} \\ \vdots \\ a + (\hat{c}_{0} + \hat{c}_{1}t_{\tau})\Delta t + \dots + (\hat{c}_{0} + \hat{c}_{1}t_{m-1})(t - t_{m-1}), t_{m-1} < t \le t_{m} \end{cases}$$
(6.22)

The variance of the nonlinear Gamma bridge is

$$D_{i}(t_{j}) - D_{i}(t_{j-1}) \sim G\left(\frac{m_{i}^{a \to b}(t_{j}) - m_{i}^{a \to b}(t_{j-1})}{\theta}, \theta\right)$$
(6.23)

$$Var(GB_{i}^{a\to b}(t)) = \frac{\left(m_{i}^{a\to b}(t) - a\right)\left(b - m_{i}^{a\to b}(t)\right)}{\left(\frac{b-a}{\theta} + 1\right)}$$
(6.24)

From Eq. (6.24), we find that the variance first increases from one end of the "bridge" with value zero and then decreases until it reaches the other end with value zero. Figure 6.5 shows the variance plot of a Gamma bridge with  $\tau = 5$ , m = 11,  $t_{\tau} = 5$ ,  $t_m = 11$ ,  $\theta = 10$ , a = 1, b = 2,  $m_i^{a \to b}(t) = \frac{1}{6}t + \frac{1}{6}$ . Like that of the nonlinear Brownian Bridge, this property is consistent with the fact that missing data far from the other observations in the time frame has higher uncertainties and is more difficult to interpolate with high accuracy.



Figure 6.5 The variance plot of a Gamma bridge

## 6.1.3 Inverse Gaussian Bridge

When degradation increments are observed to follow inverse Gaussian distribution, an inverse Gaussian bridge is suitable to interpolate the missing values. Therefore, the inverse Gaussian Bridge is a more suitable alternative for estimating the missing data in these situations. The mean function  $m_i^{a\to b}(t)$  is the same as that in section 6.1.1. We let

$$X(t) = D_i(t) - D_i(t_{\tau}) \sim Ig\left(m_i^{a \to b}(t) - a, \lambda\left(m_i^{a \to b}(t) - a\right)^2\right), t_{\tau} \leq t \leq t_m,$$
  

$$Y(t) = D_i(t_m) - D_i(t) \sim Ig\left(b - m_i^{a \to b}(t), \lambda\left(b - m_i^{a \to b}(t)\right)^2\right), t_{\tau} \leq t \leq t_m, \quad Z = X(t) + Y(t).$$

where  $\lambda$  is a constant that controls the volatility. Given the starting time  $t_{\tau}$  and degradation data *a*, the ending time  $t_m$  and ending degradation data *b*, the conditional probability density function of X(t)|Z is

$$\begin{split} f_{X(t)|Z}(x(t)|z) \\ &= P\left(D_{i}(t) - D_{i}(t_{\tau}) = x(t) | D_{i}(t_{m}) - D_{i}(t_{\tau}) = z\right) \\ &= \frac{P\left(D_{i}(t) - D_{i}(t_{\tau}) = x(t), D_{i}(t_{m}) - D_{i}(t_{\tau}) = z\right)}{P\left(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z\right)} \\ &= \frac{P\left(D_{i}(t) - D_{i}(t_{\tau}) = x(t), D_{i}(t_{m}) - D_{i}(t) = z - x(t)\right)}{P\left(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z\right)} \\ &= \frac{P\left(D_{i}(t) - D_{i}(t_{\tau}) = x(t)\right) P\left(D_{i}(t_{m}) - D_{i}(t) = z - x(t)\right)}{P\left(D_{i}(t_{m}) - D_{i}(t_{\tau}) = z\right)} \\ &= \sqrt{\frac{\lambda}{2\pi}} \left(\frac{z}{x(t)z - x(t)^{2}}\right)^{3} \left(\frac{\left(m_{i}^{a \to b}(t) - a\right)\left(b - m_{i}^{a \to b}(t)\right)}{b - a}\right)^{2}}{b - a}^{2} \times \\ \exp\left[-\lambda \left(\frac{\left(x(t) - \left(m_{i}^{a \to b}(t) - a\right)\right)^{2}}{2x(t)\left(m_{i}^{a \to b}(t) - a\right)^{2}} + \frac{\left(z - x(t) - \left(b - m_{i}^{a \to b}(t)\right)\right)^{2}}{2(z - x(t))\left(b - m_{i}^{a \to b}(t)\right)^{2}} - \frac{\left(z - \left(b - a\right)\right)^{2}}{2z(b - a)^{2}}\right)\right]$$
(6.25)

where  $\lambda$  is estimated from the observed data as follows: assuming the degradation process is an inverse Gaussian process with time-dependent mean  $m_i^{a\to b}(t)$  and shape parameter and the increments follow inverse Gaussian distributions as shown in Eq. (6.26), the parameter  $\lambda$  is estimated using MLE.

$$D_{i}(t_{j}) - D_{i}(t_{j-1}) \sim Ig\left(m_{i}^{a \to b}(t_{j}) - m_{i}^{a \to b}(t_{j-1}), \ \lambda\left(m_{i}^{a \to b}(t_{j}) - m_{i}^{a \to b}(t_{j-1})\right)^{2}\right) \lambda\left(m_{i}^{a \to b}(t)\right)^{2}$$

$$(6.26)$$

Let  $x(t)|z^*$  be a random number generated from the PDF in Eq. (6.25), the inverse Gaussian Bridge is

$$IGB_i^{a \to b}\left(t\right) = a + x\left(t\right) \left|z\right|^*$$
(6.27)

The variance of the inverse Gaussian Bridge is

$$Var\left(IGB_{i}^{a\to b}\left(t\right)\right) = E\left(\left(IGB_{i}^{a\to b}\left(t\right)\right)^{2}\right) - \left(E\left(IGB_{i}^{a\to b}\left(t\right)\right)\right)^{2}$$
(6.28)

The variance plot of an inverse Gaussian bridge is similar to that of the nonlinear Brownian Bridge.

#### 6.1.4 Models Validation

6.1.4.1 Models Validation with Corrosion Data

The three proposed bridging models are validated with corrosion data (corrosion pit depth) by (Guedes Soares *et al.*, 2011). The corrosion data is randomized by adding Gaussian noises. Assuming a proportion of data is missing, we interpolate them with the three bridging models respectively and predict future corrosion growth. Figures 6.6-6.8 show the observed data and interpolated data using the proposed nonlinear Brownian Bridge, nonlinear Gamma Bridge, and inverse Gaussian Bridge respectively when the data are assumed to be missing from  $t_r = 5$  to  $t_m = 11$ . The solid lines denote observed data while diamonds denote the interpolated data.



Figure 6.6 The interpolated data with the nonlinear Brownian Bridge


Figure 6.7 The interpolated data with the nonlinear Gamma Bridge



Figure 6.8 Interpolated data with the inverse Gaussian Bridge

After data interpolation, we censor the data at time 10 and estimate the parameters of the IIG process as described in section 3.2. We predict corrosion growth after time 10. Figures 6.9-6.11 show the predicted corrosion growth with three bridges respectively. Intuitively, we see that all these bridging methods predict missing data well and the interpolated data are consistent with observed data.



Figure 6.9 The corrosion growth prediction with data interpolated by the nonlinear

Brownian Bridge



Figure 6.10 The corrosion growth prediction with data interpolated by the nonlinear



Gamma Bridge

Figure 6.11 Corrosion growth prediction with data interpolated by the inverse Gaussian

Bridge

To quantitatively analyze the data interpolation results, we compare two indicators: failure time prediction precision and mean corrosion growth prediction precision. First, we compare the predicted mean failure time and 98% confidence interval with the true failure time. The threshold is set as 1.0mm. Figures 6.12-6.14 show the failure time prediction results when 40%, 20% and 10% data are missing and interpolated with the three bridging approaches. In the three figures, the predicted mean values are denoted by filled dots while the 98% confidence intervals are represented by the horizontal bars. The horizontal line represents the true failure time. As expected, the confidence intervals narrow down as the missing proportion decreases. All the predicted confidence intervals cover the true failure time. In this case, where degradation monotonically increases, nonlinear Gamma bridge and inverse Gaussian bridge perform better than the nonlinear Brownian bridge regarding confidence length and overlap of confidence intervals with that of complete data.



Figure 6.12 Failure time prediction confidence intervals with 40% missing data



interpolated from the bridges

Figure 6.13 Failure time prediction confidence intervals with 20% missing data

interpolated from the bridges



Figure 6.14 Failure time prediction confidence intervals with 10% missing data interpolated from the bridges

Second, another indicator, the sum of squared percent errors (SSPE) are calculated and compared. SSPE is defined as follows

$$SSPE = \sum_{j=t_c+1}^{T_j} \left(\frac{d_j - \overline{d}_j}{d_j}\right)^2$$
(6.29)

where  $\overline{d}_{j}$  is the predicted mean value of the  $j^{th}$  observation,  $d_{i}$  is the true degradation value,  $T_{f}$  is the failure time,  $t_{c}$  is the censoring time. Figure 6.15 shows the SSPE using three bridging methods when 10%, 20%, 30% and 40% data are missing. We find the nonlinear Gamma bridge and inverse Gaussian bridge are better than the nonlinear

Brownian bridge regarding SSPE. The SSPE increases as missing data size increases. However, the SSPE is small for all cases.



Figure 6.15 SSPE using three bridges when different proportions of data are missing

# 6.1.4.2 Models Validation with Crack Growth Data

In this section, a fatigue-crack-growth dataset is utilized to evaluate the results of the proposed models numerically. The dataset is obtained from a fatigue testing experiment conducted by the authors in the NDT&E lab of Harbin Institute of Technology. A specimen, made of 6061 aluminum, is subject to fatigue of the 70Hz-frequency cyclic load until fracture. Meanwhile, the length of the crack on the specimen is measured regularly by nine

Non-Destructive Testing (NDT) sensors. Detailed descriptions of the sensors are summarized in Table 6.1.

Sensor symbol	Descriptions			
TC	Thermal camera			
OC	Optical camera			
X-ray	X-ray machine			
TOFD-1	2MHz Time-of-flight diffraction probe			
TOFD-2	5MHz Time-of-flight diffraction probe			
PA-1	5MHz Ultrasonic phased array probe			
PA-2	10MHz Ultrasonic phased array probe			
EC-1	Eddy current probe 1			
EC-2	Eddy current probe 2			

Table 6.1 Descriptions of the sensors

Each sensor has a total of 21 observations. Since the metallographic microscope can observe the micro-cracks with a width of  $1\mu m$ , these observations are regarded as the actual crack length. The starting crack length is 3mm. The threshold is 23mm. The complete dataset is plotted against time in Figure 6.16. We randomly pick some observations to be missing. The missing data pattern is shown in Figure 6.17. The shaded

color indicates that the corresponding observation is observed. The white color indicates that the corresponding observation is missing. The missing data are interpolated using the three stochastic bridges. One interpolated dataset is shown in Figure 6.18 where all the interpolated data points are represented with dashed lines, while the observed data points are plotted with solid lines. In the data interpolation process using stochastic bridges, it shows that the degradation data observed by sensor OC, X-ray and TOFD-1 follow the Brownian motion processes, those observed by sensors TOFD-2 and PA-1 follow the nonlinear Gamma bridge and those observed by sensors PA-2, EC-1 and EC-2 follow the inverse Gaussian process. The variance plot of the missing data and the observed values is shown in Figure 6.19.



Figure 6.16 The complete dataset







Figure 6.18 An interpolated missing dataset



Figure 6.19 The variance plot

To evaluate the effectiveness of the proposed data interpolation model, we use the data interpolation approaches modified from the k nearest neighbor imputation (kNNI) by Batista *et al.* (2003) and those obtained by the bootstrap approach by Guo *et al.* (2018) as a benchmark and compare the mean imputation results. The kNNI is developed to interpolate the missing attribute values in a multivariate dataset. We modify kNNI so that it can be used to interpolate the univariate missing observations within a multi-sensor time-series dataset. The bootstrap approach in (Guo *et al.*, 2018) is used to predict degradation increments and is modified to interpolate the missing observations within the dataset. The details of the modified kNNI and bootstrap approaches are found in Appendices G and H respectively. The interpolated datasets with the two methods are shown in Figures 6.20 and 6.21. The mean squared error (MSE) of mean interpolated values (compared with the actual observations in the missing positions) for both methods are summarized in Table 6.2. It is

observed that, of all the 24 interpolated missing data, the mean of the stochastic bridges is more accurate than that of the benchmark approaches.

Approaches	MSE
Mean of the Stochastic bridges	0.248
Bootstrap + modified kNNI	1.146

Table 6.2 Mean squared errors of interpolated values



Figure 6.20 Mean interpolated data by stochastic bridges (dashed lines denote

interpolated data)



Figure 6.21 Mean interpolated data by modified bootstrap and kNNI (dashed lines denote

interpolated data)

# 6.2 Multiple Sensors Data Integration

In this section, we introduce a non-parametric data integration methodology. As discussed earlier, the main challenge in data integration is the correlation between observations obtained by different sensors. We propose a dynamic cluster-based data integration model that integrates multiple sensor observations (including interpolated data) in consecutive times, rather than in individual time instants. At each observation time, weights of observations are calculated. The weighted sum of the observations is the integrated value of all observations at that time instant. At each observation time, when we assign the sensors' data to a fixed number of clusters, we may observe that some sensors switch clusters from one time instant to another while others remain in the same cluster throughout the observation times. Those sensors that switch frequently are considered as "erratic" and "unreliable" and are assigned lower weights and vice versa.

# 6.2.1 Clustering of the Sensors' Observations

Let  $d_i(t)$  represent the observation of sensor i at time t. At each observation time, we apply k-means clustering to assign the N (the total number of sensors) measurements into k clusters. We define the switching of sensors from one cluster to another as follows: if there are  $p \ge q$  sensors belong to the same cluster at two times  $t_1$  and  $t_2$ , these psensors do not switch between clusters (q is a predefined value). The relationship is represented as  $C_i(t_1) = C_i(t_2), i \in \{i_1, ..., i_p\}$ , where  $C_i(t)$  is the cluster where sensor ibelongs at time t and  $\{i_1, ..., i_p\}$  is the set of indexes of the p sensors. Suppose we are considering a time window at time t with length l, where the time window is defined as the set of observations and interpolated data of all N sensors over the period (t-l, t). There are cases where p < q measurements of the same sensors do not switch between cluster over the entire time window. In other words, those sensors do not switch between clusters. We define the diameter of a cluster as the distance between the two farthest observations in the cluster. If the distances between the centroid of this cluster and other centroids are larger than the diameter of the largest cluster, those measurements are treated as outliers and are assigned zero weights. Figure 6.22 shows an example with N = 8sensors during a time window with length l = 5. The eight measurements are assigned to k = 3 clusters. The measurements are shown in Table 6.3. In this example, q = 2. Because there are p = 2 sensors (sensors 1 and 5) in the same cluster both at times  $t_1$ and  $t_2$ , we have  $C_i(t_1) = C_i(t_2), i \in \{1, 5\}$ . Similarly,  $C_i(t_1) = C_i(t_2), i \in \{3, 6\}$  and  $C_i(t_1) = C_i(t_2), i \in \{7, 8\}$ . In comparison,  $C_i(t_1) \neq C_i(t_2), i \in \{2\}$ . This is because there are less than q sensors assigned to the same cluster that contains sensor 2 both at time  $t_1$ and  $t_2$ . The measurements of sensor 4 in the time window are determined as outliers because only those measurements are continuously assigned to an individual cluster of the time window and the distance between the centroid of the cluster and other centroids are all larger than the diameter of the largest cluster.

Sensors	<i>t</i> <sub>1</sub>	<i>t</i> <sub>2</sub>	<i>t</i> <sub>3</sub>	$t_4$	$t_5$	
Sensor 1	1.4	2.0	2.5	3.0	3.5	
Sensor 2	1.4	1.5	2.0	2.5	3.0	
Sensor 3	1.6	1.9	2.5	3.1	3.4	

 Table 6.3 Example of sensors data observed during five-time instants

Sensor 4	1.0	1.0	1.0	1.0	1.0
Sensor 5	1.3	2.1	2.4	2.9	3.6
Sensor 6	1.7	2.0	2.4	3.1	3.4
Sensor 7	1.7	1.4	1.8	2.5	3.1
Sensor 8	1.6	1.6	1.9	2.6	3.1



Figure 6.22 Cluster switching example (sensors 2, 3 and 6 switched clusters at time 2)

# 6.2.2 Sensors' Data Integration

After data interpolation, the dataset now includes observed data and interpolated data. Since the interpolated data are not observed directly but predicted based on observations, their weights should be adjusted according to their variances. As discussed in section 6.1, the variance of interpolated data depends on the position in the stochastic bridge. Accordingly, we assign higher weights to observed data and interpolated data that are close to the two ends of the bridge and lower weights to those in the middle of the bridge. In other words, we assign higher weights to data with smaller variance and lower weights to data with larger variance. The same goes for data interpolated with bootstrap. The variance of the missing data obtained with bootstrap refers to the sample variance. Because the observed data have no uncertainty, we treat their variances as zero. For simplicity, we use  $\sigma_i^*(t)^2$  to denote the variance of the missing data of sensor *i* at time *t*. The dynamic observation weight of sensor *i* at time *t* is as follows:

$$w_{i}(t) = \frac{\frac{1}{\exp(\sigma_{i}^{*}(t)^{2})}}{\sum_{i=1}^{N} \left(\frac{1}{\exp(\sigma_{i}^{*}(t)^{2})}\right)} \frac{g(dp(d_{i}(t)))\max(R_{i}(t) - P_{i}(t), 0)}{\sum_{i=1}^{N} g(dp(d_{i}(t)))\max(R_{i}(t) - P_{i}(t), 0))}$$
(6.30)

Where

$$R_i(t) = \sum_{C_i(t-h)=C_i(t-l), 1 \le h \le l} r \exp(-h\lambda)$$
(6.31)

$$P_{i}(t) = \sum_{C_{i}(t-h)\neq C_{i}(t-h-1), 1 \le h \le l} o \exp(-h\mu)$$
(6.32)

 $w_i(t)$  is the weight of sensor *i* at time *t*. *r*, *o*,  $\lambda$  and  $\mu$  are all positive constants. *r* is the reward for each time a sensor stays in its initial cluster assigned at the first location of the time window and *o* is the penalty for each time a sensor switches between clusters in the consecutive times in the time window. The notation  $R_i(t)$  is a weighted sum of the rewards of sensor *i* in the time window at time *t*, while  $P_i(t)$  is the weighted sum of the penalties of sensor *i* in the time window at time *t*. The notation *h* indicates the location of observation in the time window while  $\exp(-h\lambda)$  and  $\exp(-h\mu)$  are the corresponding weights of reward and penalty. The weights of rewards and penalties are negatively related to h so that the closer the location of the observation to the first position of the time window, the higher the weight (of the reward or penalty). The parameters  $\lambda$  and  $\mu$  are weight coefficients of the reward and penalty respectively and  $dp(d_i(t))$  is the depth of observation  $d_i(t)$ . Depth is proposed by Tukey (1975) to measure the distance of the observation from the center of the cluster. The function  $g(dp(d_i(t)))$  is a monotonically increasing function of  $dp(d_i(t))$ . It takes forms such as polynomial, exponential and others. Without loss of generality, we assume  $g(dp(d_i(t)))$  to be

$$g\left(dp\left(d_{i}(t)\right)\right) = c \cdot dp(d_{i}(t)) \tag{6.33}$$

where c is a positive constant. Eq. (6.30) proposes that the weight of a sensor at a specific time is positively related to the weighted sum of the time it stays in the initial cluster minus the weighted sum of the times it switches between clusters in the time window. The values of k, l, o, q and c are determined empirically. Normalizing the weights, we have:

$$W_{i}(t) = \frac{W_{i}(t)}{\sum_{i=1}^{N} W_{i}(t)}$$
(6.34)

The integrated value of all sensor observations at time t is

$$\tilde{d}(t) = \sum_{i=1}^{N} W_i(t) d_i(t)$$
(6.35)

# 6.2.3 Case Study

# 6.2.3.1 Multiple Sensors' Data Integration

We apply the data integration model proposed to the observed and interpolated dataset in Figure 6.18. The weights obtained by considering the variance of missing values and cluster switching are provided in Figures 6.23 (a) and 6.23 (b) respectively. The final normalized weights are shown in Figure 6.24. It is observed that the integrated path lies close to the majority of the sensors' paths. The weights of sensors change over time dynamically. Weights of TC and OC are close to zero, which corresponds to the fact that the observations of the two sensors are far from the majority of other sensors and are identified as outliers. When sensors TOFD-1 and TOFD-2 deviate from the majority of the other sensors at times t = 15 and t = 19, 20 in Figure 5.18, they are identified as "unreliable" or "erratic" sensors and their weights are reduced almost immediately as shown in Figure 6.23 (b). Meanwhile, the weights are adjusted considering whether the data is interpolated or observed. For example, as sensor EC-2 has four missing values, the variances in the middle of the bridge are higher thus their weights are lower than other sensors. The performance of sensors is evaluated dynamically, and the weights are adjusted accordingly. Consequently, consistent (sensors that switch less frequently) and reliable (sensors that have fewer missing observations) sensors are assigned higher weights. Parameters used in the case study are shown in Table 6.4.

k	l	М	Ν	0	q	r	λ	μ
3	5	21	9	1	2	1	0.5	0.8

Table 6.4 Summary of parameters used in the case study



Figure 6.23 (a) Weights of the sensors by variance (b) Weights of the sensors by cluster



shifting

Figure 6.24 Final normalized weights of the sensors



Figure 6.25 Integrated paths and mean of the incomplete data

To numerically evaluate the performance of the proposed models, we compare the Mean Squared Errors (MSE) of the individual sensors and the four integrated paths: the integrated path obtained with proposed data integration model of the complete dataset, the integrated path obtained with proposed data integration model of the interpolated dataset obtained with stochastic bridges, the integrated path obtained with Shi *et al.* (2012) of the interpolated dataset obtained with stochastic bridges and mean of the missing dataset. Figure 6.25 shows the integrated paths of the interpolated datasets. Table 6.5 shows the MSE comparison results. Under the condition that missing data exist, the proposed approach that combines stochastic bridges and data integration method provides the most accurate results. Its MSE is smaller than that obtained with Shi's model and is close to that

obtained with the complete dataset. All the integrated paths are better than those obtained by taking the mean of the incomplete dataset. This shows the feasibility of using the combination of stochastic bridges and the proposed data integration approach to integrating information from multiple sensors, even if there are missing data within the dataset.

Sensor symbol	MSE
TC	65.8712
OC	6.9456
X-ray	0.1245
TOFD-1	0.3452
TOFD-2	0.3896
PA-1	0.2641
PA-2	0.1659
EC-1	0.1312
EC-2	0.1732
The integrated path with the complete dataset	0.0809
The integrated path with the proposed model	0.1173
The integrated path by Shi's model	0.1302
Mean of the incomplete dataset	0.2751

 Table 6.5 MSE of individual sensors and the integrated paths

In this section, we use the improved inverse Gaussian (IIG) process to predict the remaining life of a unit, assuming the degradation data are censored before failure. Chi-square goodness-of-fit test is performed with data from all sensors (integrated data are treated as new sensor's data) to ensure that the degradation increments follow IIG distribution. The model parameters are estimated by MLE. We compare the expectation of the remaining life of each sensor (including integrated path) with actual remaining life. Let  $T_D$  represent the actual failure time of the specimen. The percentage error between predicted and the actual failure time is evaluated as well with Eq. (6.36), where  $\tilde{T}_i(t)$  is the remaining life

$$err_{i}(t) = \frac{\left(t + \tilde{T}_{i}(t)\right) - T_{D}}{T_{D}}$$
(6.36)

To compare the situations when different percentages of actual lifetimes are remaining, we censor the data by 10 and 25 percent. Table 6.6 demonstrates the detailed comparison results. It shows that the integrated path with interpolated data outperforms all individual sensor paths and the mean of the incomplete dataset in predicting remaining life. Moreover, the prediction error is only slightly higher than that of the complete dataset.

Sensor symbol	25%	10%
TC	-	-
OC	6.87	12.48
X-ray	9.07	3.45
TOFD-1	26.35	3.92
TOFD-2	16.37	5.45
PA-1	2.92	3.36
PA-2	4.32	3.42
EC-1	9.43	5.04
EC-2	3.27	2.91
The integrated path with the complete dataset	1.57	2.21
The integrated path with the interpolated dataset	1.60	2.24
Mean of the incomplete dataset	6.62	7.46

 Table 6.6 Comparison results of absolute percentage error by using each sensor and

# integrated paths with different percentage of actual remaining lifetime

# 6.3 Summary

We develop three stochastic bridges to interpolate the missing data in the degradation monitoring process. The stochastic bridges are developed based on obtaining the probability density function of the missing values conditioned on the fixed (observed) values on two bridge ends. They capture the nonlinearity of the underlying degradation process, as well as the uncertainty of the missing values. The data integration model assigns weights based on cluster switching of sensors and the influence of the variance (if the data is interpolated). The models are validated with actual fatigue crack growth data. Results show that the stochastic bridges outperform the benchmark models, which include modified *k*NNI and bootstrap. The proposed data integration model also outperforms the benchmark data integration model stated in this chapter. The failure times predicted with the integrated interpolated dataset obtained with the proposed models are more accurate compared with other models. The proposed models provide accurate predictions of the missing data and the true degradation path.

### **CHAPTER 7**

# **CONCLUSIONS AND FUTURE RESEARCH**

### 7.1 Conclusions

There are several major contributions of this dissertation, they are the stochastic modeling of corrosion growth, missing data imputation and multi-sensor data integration in the corrosion monitoring process and the stochastic modeling of degradation branching processes. More specifically

1. We propose two stochastic models to capture the corrosion depth growth and corrosion volume growth respectively. The improved inverse Gaussian (IIG) process is developed as a stochastic degradation model to describe the corrosion depth growth. The model captures the dependency between the corrosion increments and corrosion depth and is more accurate and robust than the inverse Gaussian (IG) model. The corrosion pit volume growth model is developed by assuming that the corrosion pit growth in both radial and depth directions follow IIG processes. It captures the phenomenon where a critical amount of volume loss of material leads to failures of a component although the corrosion pit depth is noncritical. In addition, we develop a physics-based stochastic model to characterize the corrosion pits propagation. The physics factors include the spatial and size distributions of the particles and the

influence of corroded pits, are incorporated into the pits' initiation model. The influence of stresses that include temperature, relative humidity and pH is investigated based on chemical reactions of the corrosion process. We develop three stochastic bridges to interpolate the missing data in the corrosion monitoring process. They are developed based on obtaining the PDF of the missing values conditioned on the fixed (observed) values on two bridge ends. They capture the nonlinearity of the underlying degradation process, as well as the uncertainty of the missing values. When multiple sensors are used to monitor the corrosion growth process, a data integration model is proposed to assign weights based on cluster switching of the sensors and the influence of the variance (if the data is interpolated). The models are validated with actual corrosion data.

2. We develop stochastic models for the corrosion propagation to other locations, which we refer to as the degradation branching processes (DBP). As the propagation of the corrosion proceeds, it branches into two branches and it is terminated upon system failure. This is a general model which is also applicable to the propagation of the cracks in materials. Besides, the degradation process of each branch can be applicable to any stochastic process. Two types of DBPs are identified based on the branching criteria: in Type I DBP a new branch is initiated when the original branch's degradation reaches its threshold while in Type II DBP the branching time is determined by other related

physics-based processes independent of the degradation of the branches. The statistical properties of the process, such as the PDF, the mean and the variance of the total degradation, the expected number of branches, the remaining life prediction, the reliability function and the MTTF are investigated. Closed-form expressions of mean, variance and the MTTF are obtained for special cases. The model has been validated using real-life data.

3. We generalize the stochastic degradation branching models by considering the random direction of a branch from its parent branch, which is referred to as the general directional degradation branching process. The model considers two factors during the branching process from a parent branch: the first is that the number of potential new branches is a random variable and the second is that the angles of these potential branches also are expressed as a random variable. The branching angles have an influence on the total degradation by affecting the branches' degradation growth rate. We develop efficient algorithms to estimate the total degradation of all the branches as well as the number of branches in the entire branching process. The reliability metrics such as the reliability, the PDF of residual life, the MTTF and others are obtained. The proposed model is validated with simulations.

# 7.2 Future Research

We have identified two areas to be investigated further, they are:

First area: In Chapters 4 and 5, the degradation branching processes and the general directional degradation branching processes under constant stress conditions are discussed. The stresses such as loading, humidity and temperature have significant influences on the degradation branching process of fuel cell membrane cracks' growth (Singh *et al.*, 2017). Sakai *et al.* (1991) also state that loading affects the degradation growth rate of the branches in the glass degradation process. In the future, we intend to develop stochastic models to incorporate such factors and design the corresponding acceleration degradation testing plans.

Second area: In Chapter 5, we investigate the influence of branching directions on the total degradation. However, as the direction of a crack surface in the 3D space is determined by two angles with respect to the two perpendicular axes in space (Kagan, 1994), the assumption made in the proposed model in Chapter 5 that the offspring branches only rotate around one axis with respect to the parent branch may result in inaccurate degradation estimation in some conditions. In the future, we intend to relax this assumption and investigate the crack propagation in three dimensions and estimate the corresponding system reliability.

# APPENDIX A PARAMETERS ESTIMATION IN IIG AND IG PROCESSES

The parameters of the IIG process include  $\{\mu_0, \mu_1, \lambda\}$ . The maximum likelihood estimation method is used to estimate these parameters. Let  $D = \{d(t_1), d(t_2), \dots d(t_{\tau})\}$  be the observed degradation values. The initial degradation state is  $d(t_0) = 0$ . In this case, the degradation process has been observed  $\tau$  times. The log-likelihood function is given by

$$\begin{split} &\ln l(\mu_{0}, \mu_{1}, \lambda) \\ &= \sum_{j=1}^{\tau} \ln f\left(d(t_{j}) - d(t_{j-1})\right) \\ &= \sum_{j=1}^{\tau} \ln \left(\sqrt{\frac{\lambda \Lambda^{2}\left(d(t_{j}) - d(t_{j-1})\right)}{2\pi\left(d(t_{j}) - d(t_{j-1})\right)^{3}}} \exp\left[-\frac{\lambda\left(\left(d(t_{j}) - d(t_{j-1})\right) - \Lambda\left(d(t_{j-1})\right)\right)^{2}\right]}{2\left(d(t_{j}) - d(t_{j-1})\right)}\right] \right] \\ &= \sum_{j=1}^{\tau} \ln \left(\sqrt{\frac{\lambda\left(\mu_{o} + \mu_{1}d(t_{j-1})\right)^{2}}{2\pi\left(d(t_{j}) - d(t_{j-1})\right)^{3}}} \exp\left[-\frac{\lambda\left(\left(d(t_{j}) - d(t_{j-1})\right) - \left(\mu_{o} + \mu_{1}d(t_{j-1})\right)\right)^{2}\right]}{2\left(d(t_{j}) - d(t_{j-1})\right)}\right] \right] \\ &= \sum_{j=1}^{\tau} \left[\frac{1}{2} \ln \frac{\lambda\left(\mu_{o} + \mu_{1}d(t_{j-1})\right)^{2}}{2\pi\left(d(t_{j}) - d(t_{j-1})\right)^{3}} - \frac{\lambda\left(\left(d(t_{j}) - d(t_{j-1})\right) - \left(\mu_{o} + \mu_{1}d(t_{j-1})\right)\right)^{2}\right]}{2\left(d(t_{j}) - d(t_{j-1})\right)}\right] \end{split}$$

The parameters can be obtained by solving the following equation:

$$\frac{\partial \ln l}{\partial \mu_0} = 0, \quad \frac{\partial \ln l}{\partial \mu_1} = 0, \quad \frac{\partial \ln l}{\partial \lambda} = 0$$

The R codes are provided as follows:

Here y denotes the degradation data and dy denotes the degradation increments data. n is the size of degradation dataset. miu0 denotes  $\mu_0$ , miu1 denotes  $\mu_1$  and lambda denotes  $\lambda$ .

```
IGA.para=zeros(nx=3,ny=1)
vecy=as.vector(y[1:n,])
vecdy=as.vector(dy[1:n,])
```

y1<-data.frame(vecy) y2<-data.frame(vecdy)

```
iig.lik<-function(theta,y){
  miu0=theta[1]
  miu1=theta[2]
  lambda=theta[3]</pre>
```

# l=0

IGA.para=optim(c(40,1,10), iig.lik,y=yt)\$par

Parameters of IG process model can be estimated using maxLik in R. Here dy denotes the degradation increments. miu denotes m and lambda denotes /.

```
vecdy=as.vector(dy[1:n,])
LL <- function( param) {
    miu=param[1]
    lambda=param[2]
    R=dinvgauss(vecdy, miu, lambda*(miu)^2, log = FALSE)
    sum(log(R))
}
numericGradient(LL,c(lambda=0.07,miu=0.6))
estll <- maxLik( LL, start = c(lambda=0.07,miu=0.6))</pre>
```

#### Cycles (in million) Path 0.00 0.01 0.02 0.03 0.04 0.08 0.09 0.10 0.11 0.05 0.06 0.07 0.12 1.12 1.19 1.27 1.35 1 0.9 0.95 1 1.05 1.48 1.64 0.9 0.94 0.98 1.03 1.08 1.14 1.21 1.28 1.37 1.47 2 1.6 3 0.9 0.94 0.98 1.03 1.08 1.13 1.19 1.26 1.35 1.46 1.58 1.77 4 1.25 0.9 0.94 0.98 1.03 1.07 1.12 1.19 1.34 1.43 1.55 1.73 5 0.94 0.98 1.24 1.34 0.9 1.03 1.07 1.12 1.19 1.43 1.55 1.71 6 1.12 1.18 1.23 1.33 1.41 1.51 1.68 0.9 0.94 0.98 1.03 1.07 7 1.11 1.17 1.23 1.32 1.41 1.52 0.9 0.94 0.98 1.02 1.07 1.66 1.39 8 0.9 0.93 0.97 1.11 1.17 1.23 1.3 1.49 1.62 1 1.06 9 1.44 0.9 0.92 0.97 1.01 1.05 1.09 1.15 1.21 1.28 1.36 1.55 1.72 1.13 1.19 1.26 1.34 1.52 10 0.9 0.92 0.96 1 1.04 1.08 1.42 1.67 1.24 1.31 1.39 1.49 1.65 11 0.9 0.93 0.96 1 1.04 1.08 1.13 1.18 1.29 12 0.9 0.93 0.97 1 1.03 1.07 1.1 1.16 1.22 1.37 1.48 1.64 13 0.9 0.92 0.97 0.99 1.03 1.06 1.1 1.14 1.2 1.26 1.31 1.4 1.52 14 0.9 0.93 0.96 1.03 1.07 1.12 1.16 1.2 1.26 1.3 1.37 1.45 1 1.27 15 0.9 0.92 0.96 0.99 1.03 1.06 1.1 1.16 1.21 1.33 1.4 1.49 1.22 16 0.9 0.92 0.95 0.97 1 1.03 1.07 1.11 1.16 1.26 1.33 1.4 17 0.96 1.05 1.11 1.16 1.2 1.24 1.32 0.9 0.93 0.97 1 1.08 1.38 18 0.9 0.92 0.94 0.97 1.01 1.04 1.07 1.09 1.14 1.19 1.23 1.28 1.35 0.99 19 0.9 0.92 0.94 0.97 1.02 1.05 1.08 1.12 1.16 1.2 1.25 1.31 20 0.9 0.92 0.94 0.97 0.99 1.02 1.05 1.08 1.12 1.16 1.19 1.24 1.29 21 0.9 0.92 0.94 0.97 0.99 1.02 1.04 1.07 1.11 1.14 1.18 1.22 1.27

 Table B.1 Fatigue-crack-growth data from Bogdanoff and Kozin (1985)

APPENDIX B THE CRACK GROWTH DATA

# APPENDIX C CORROSION VOLUME GROWTH MODELING BASED ON IMPROVED GAMMA PROCESS

One of the challenges of modeling corrosion volume growth is that morphologies of corrosion pits are highly stochastic and irregular (Horner et al., 2011, Kioumarsi et al., 2016, Trzaskoma, 1990). We may observe that the corrosion propagates with different rates in all directions as shown in Figure C.1 (a). It is inappropriate to assume the shape of a pit is a perfect smooth hemisphere or half ellipsoid as presented in (Kondo, 1989, Yu et al., 2015) when estimating the pit volume growth. Kariyawasam et al. (2012) show in Figure C.2 that there is a dependency between the corrosion depth growth rate and the original corrosion depth. Al-Amin et al. (2016) also state that in the corrosion growth process, the corrosion growth rate decreases because of the buildup of the corrosion product, which inhibits the transport of the reactants to or from the surface. Moreover, if there is a cathodic reactant being consumed by corrosion and it is being depleted from the surrounding environment, a decline of the corrosion rate is expected. In the Gamma processes, usually time-dependent mean functions are assumed, which do not capture the physics of the corrosion reactions in the corrosion process. As the corrosion growth is a nonlinear process, we implement the state-dependent Gamma process to capture the physical dependency between the corrosion depth growth rate and original corrosion depth. Fan et al. (2015) demonstrate that the state-dependent Gamma process is superior to Gamma process

especially in the nonlinear degradation process. Specifically, the power-law mean function, which is widely used in corrosion modeling, is used by the Gamma process in the comparison. By using the physics-based state-dependent Gamma process model, the influence of stresses can be incorporated, as demonstrated later in this appendix. In addition, as the corrosion depth growth rates in different locations of a pit are different (newly corroded areas grow faster), the state-dependent Gamma process can be efficiently applied to capture the depth growth of those locations simultaneously.



Figure C.1 (a) Three-dimensional X-ray micro-tomographic image of a corrosion pit by Horner *et al.* (2011) (notice the morphologies of the corrosion pit) (b) A simplified schematic diagram of the corrosion pit growth



**Figure C.2** The original depth versus the depth growth rate adapted from (Kariyawasam *et al.*, 2012) (negative rates are abandoned as we focus on the general monotonically increasing corrosion processes)



Figure C.3 Projected view of the area growth of a corrosion pit
To characterize the corrosion pit volume growth, we consider the pit growth in both radial and depth directions as independent state-dependent Gamma processes. As shown in a simplified diagram of the corrosion pit (Figure C.1(b)), the corrosion pit is divided into  $n_s$ sectors. We assume that the pit consists of many small squares (pixels), as projected in Figure C.3. It is evident that the pixel depth varies from one place to another. The volume growth process is described in Figure C.3.

#### C.1 Corrosion Pit Area Growth

It is widely accepted that as corrosion progresses, the produced corrosion layer prevents the atmosphere from contacting new metal and decreases corrosion growth rate (Vanaei *et al.*, 2017). In the proposed state-dependent Gamma process (Wang *et al.*, 2020a), we let the shape parameter be a function of the corrosion state (pixel depth or radial length), such that corrosion growth rate decreases with time. The corroded area grows as corrosion pit grows. As shown in Figure C.4 (a), suppose a pit is divided into  $n_s$  sectors, where the center of the pit corresponds to the location of the deepest pixel in the pit (which can be obtained by slicing the image of the pit horizontally).



Figure C.4 (a) The sectors of a corrosion pit (b) The area growth approximation of a

#### sector

The angle of each sector is  $\xi = \frac{2\pi}{n_s}$ . Suppose that at time *t* the radius of the *j*<sup>th</sup> sector

is  $r_j(t)$ . We use Ga(x) to denote the probability density function (PDF) of the Gamma distribution. The radius growth  $\Delta R_j(t)$  with starting value  $r_j(t)$  follows a state-dependent Gamma process with parameters  $\eta_0 > 0$ ,  $\eta_1 > 0$ ,  $\omega > 0$  as

$$\Delta R_{j}(t) \sim Ga\left(\Delta r_{j}(t) \middle| \frac{1}{\eta_{0} + \eta_{1}r_{j}(t)}, \omega\right)$$
(C.1)

where  $\frac{1}{\eta_0 + \eta_1 r_j(t)}$  captures the dependency between the corrosion growth rate and the

corrosion value such that larger  $r_j(t)$  leads to smaller corrosion growth rate. The radii of the sectors grow independently. As shown in Figure C.4 (b), by simple geometry, the length of the arc is  $2r_j(t)\sin\frac{\xi}{2}$ . Since the sectors are in the same corrosion pit, the parameters  $\eta_0$ ,  $\eta_1$ ,  $\omega$  are assumed to be the same for all the sectors. All the newly corroded pixels begin to grow immediately in the depth direction after they are corroded. As proved in section C.6, when the radius increment is small (growth time is small) and the angle is small, the area growth can be approximated precisely with a rectangle as shown in Figure C.4 (b). The area growth of sector j in (t, t+1) is  $\Delta S_j(t) = 2r_j(t) \sin \frac{\xi}{2} \cdot \Delta R_j(t)$  and  $\Delta S_j(t)$  follows a Gamma distribution as shown in Eq. (C.2).

$$\Delta S_{j}(t) \sim Ga\left(\Delta s_{j}(t) \middle| \frac{1}{\eta_{0} + \eta_{1}r_{j}(t)}, \ 2r_{j}(t)\sin\frac{\xi}{2} \cdot \omega\right)$$
(C.2)

The area of a pixel is  $s_0$ . The number of corroded pixels in sector j in (t, t+1) is

$$\Delta N_{j}(t) = \frac{\Delta S_{j}(t)}{s_{0}} \sim Ga\left(\Delta n_{j}(t) \left| \frac{1}{\eta_{0} + \eta_{1}r_{j}(t)}, 2r_{j}(t)\sin\frac{\xi}{2} \cdot \frac{\omega}{s_{0}} \right. \right)$$
(C.3)

Correspondingly, the increase of the number of corroded pixels in (t, t+1) in the pit is

$$\Delta N(t) = \sum_{j=1}^{n_s} \Delta N_j(t). \text{ Let } \phi_j(t) = \frac{1}{\eta_0 + \eta_1 r_j(t)}, \quad \varphi_j(t) = 2r_j(t) \sin \frac{\xi}{2} \cdot \frac{\omega}{s_0}.$$

$$\left\{ \Delta N_j(t), \ 1 \le j \le n_s \right\} \text{ are Gamma random variables with shape parameters } \left\{ \phi_j(t) \right\}. \text{ and scale parameters } \left\{ \varphi_j(t) \right\}.$$

$$f_{\Delta N(t)}(\Delta n(t)) = \frac{C(t) \sum_{k=0}^{\infty} \delta_k(t) \Delta n^{\rho(t)+k-1}(t) e^{-\Delta n(t)/\varphi_1(t)}}{\left[ \Gamma(\rho(t)+k) \varphi_1^{\rho(t)+k}(t) \right]}$$
(C.4)

where

$$C(t) = \prod_{j=1}^{n_s} \left( \varphi_1(t) / \varphi_j(t) \right)^{\phi_i(t)} , \qquad \gamma_k(t) = \sum_{j=1}^{N_s} \frac{\phi_j(t) \left( 1 - \varphi_1(t) / \varphi_j(t) \right)^k}{k}, \ k = 1, 2, \cdots ,$$
  
$$\delta_{k+1}(t) = \frac{1}{k+1} \sum_{j=1}^{k+1} j \gamma_j(t) \delta_{k+1-j}(t), \ \delta_0(t) = 1, \ \rho(t) = \sum_{j=1}^{n_s} \phi_j(t).$$

Note that the pixels corroded in the next time increment grow with an initial depth of 0.

#### C.2 Volume Growth of the Corroded Pixels

The corrosion pit volume growth is the sum of the volume growth of all the corroded pixels. Let  $d_i(t)$  denote the corrosion depth of corroded pixel *i* at time *t*. The depth of pixel *i* at t+1 is  $D_i(t+1) = \Delta D_i(t) + d_i(t)$ , where  $\Delta D_i(t)$  is the corrosion depth increment of pixel *i* in (t, t+1). The PDF of  $\Delta D_i(t)$  is

$$f_{\Delta D_{i}(t)}(\Delta d_{i}(t)) = Ga\left(\Delta d_{i}(t) \middle| \frac{1}{\mu_{0} + \mu_{1}d_{i}(t)}, \lambda\right)$$

$$= \frac{\Delta d_{i}^{\frac{1}{\mu_{0} + \mu_{1}d_{i}(t)}^{-1}}(t)e^{-\frac{\Delta d_{i}(t)}{\lambda}}I_{(0,\infty)}(\Delta d_{i}(t))}{\Gamma\left(\frac{1}{\mu_{0} + \mu_{1}d_{i}(t)}\right)\lambda^{\frac{1}{\mu_{0} + \mu_{1}d_{i}(t)}}}$$
(C.5)

where parameters  $\mu_0 > 0$ ,  $\mu_1 > 0$ ,  $\lambda > 0$ . They are the same for all the pixels.  $I_A(x)$  is an indicator function defined as follows:

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Note that the parameters  $\mu_0$ ,  $\mu_1$  and  $\eta_0$ ,  $\eta_1$  jointly determine the shape of the pit. In general, a sheer-sided pit is formed whenever  $\frac{\eta_0}{\eta_1} < \frac{\mu_0}{\mu_1}$  where the pit depth growth rate decreases slowly while the radial growth rates decrease relatively rapidly over time. In this scenario, the pixels in the center have a larger depth growth rate compared with that of the radial sectors and lead to deep pits. By contrast, a shallow pit is formed whenever  $\frac{\eta_0}{\eta_1} > \frac{\mu_0}{\mu_1}$  where the pit depth growth rate decreases rapidly while the radial growth rates decrease relatively slowly over time. In this scenario, the pixels in the center have a smaller depth growth rate decreases rapidly while the radial growth rates decrease relatively slowly over time. In this scenario, the pixels in the center have a smaller depth growth rate compared with that of the radial sectors and the pit tends to be flat.

Suppose that the area of all the existing corroded pixels (plan view area) at time t is s(t), the total number of corroded pixels is  $n(t) = \frac{s(t)}{s_0}$ . As discussed in section C.1, the increase in the number of pixels in (t, t+1) is  $\Delta N(t)$ . Let  $\Delta V_i(t)$  denote the volume growth of pixel *i*, the overall incremental volume of the pit in the next time increment caused by all the corroded pixels is

$$\Delta V(t) = \sum_{i=1}^{n(t)+\Delta N(t)} \Delta V_i(t) = s_0 \sum_{i=1}^{n(t)+\Delta N(t)} \Delta D_i(t)$$
(C.6)

where  $\Delta N(t)$  is a random variable. When  $\Delta N(t) = \Delta n(t)$ ,  $\sum_{i=1}^{n(t)+\Delta n(t)} \Delta D_i(t)$  is a Gamma

random variable with shape parameter  $\sum_{i=1}^{n(t)+\Delta n(t)} \left(\frac{1}{\mu_0 + \mu_1 d_i(t)}\right)$  and scale parameter  $\lambda$ .

By the properties of Gamma distribution,  $s_0 \sum_{i=1}^{n(t)+\Delta n(t)} \Delta D_i(t)$  is a Gamma random variable

with shape parameter  $\sum_{i=1}^{n(t)+\Delta n(t)} \left(\frac{1}{\mu_0 + \mu_1 d_i(t)}\right) = \frac{\Delta n(t)}{\mu_0} + \sum_{i=1}^{n(t)} \frac{1}{\mu_0 + \mu_1 d_i(t)} \quad \text{and scale}$ 

parameter  $\lambda s_0$ . As  $\Delta N(t)$  is a random variable, the PDF of  $\Delta V(t)$  is

$$f_{\Delta V(t)}\left(\Delta v(t)\right) = \int_{0}^{\infty} Ga\left(\Delta v(t) \left| \frac{\Delta n(t)}{\mu_{0}} + \sum_{i=1}^{n(t)} \frac{1}{\mu_{0} + \mu_{1}d_{i}(t)}, s_{0}\lambda\right) f_{\Delta N(t)}\left(\Delta n(t)\right) d\Delta n(t) \quad (C.7)$$

#### C.3 The Incremental Volume with Covariates

Corrosion is significantly affected by the atmospheric conditions, which include temperature, relative humidity Rh, pH levels and others (Soares *et al.*, 2009, VanOverloop, 1990, Yang *et al.*, 2009). We incorporate the effects of these covariates on corrosion volume growth in this section. Hughes *et al.* (2016) state that the corrosion volume growth rate (metal dissolution rate) is related to corrosion current density. Corrosion current arises when there is charge transfer between the anode (metal) and the cathode (Hughes *et al.*, 2016). The corrosion reaction current  $i_c$  is approximately described as

$$i_{c} = i_{0} \exp\left[-\frac{\alpha n F\left(E - E_{rev}\right)}{RT}\right]$$
(C.8)

where *n* is the charge on the ion in equivalents/mol. *F* is the Faraday constant (96,487C/equivalent), *R* is the gas constant (8.314 J/mol-K),  $E_{rev}$  is the reversible potential, *E* is the applied potential,  $\alpha$  is the charge transfer coefficient,  $i_0$  is the exchange current, which is the current of a single electrode at equilibrium when the electrode material experiences no loss or gain. It is a function of soluble species concentration. Eyring (1935) shows that the rate of reaction is proportional to the concentration of reagent  $H^+$  and the water vapor in the atmosphere. Since  $c(H^+)=10^{-pH}$ , we rewrite Eq. (C.8) as

$$i_{c} = \beta_{0} e^{\beta_{1} R h} e^{\beta_{2} p H} \exp\left[\frac{\beta_{3}}{T}\right]$$
  
=  $\beta_{0} \exp\left(\beta_{1} R h + \beta_{2} p H + \frac{\beta_{3}}{T}\right)$  (C.9)

where  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  are constants.  $\beta_0$  reflects the corrosion rate in equilibrium, which is similar to the meaning of  $i_0$  in Eq. (C.8). Eq. (C.8) describes the base corrosion (volume) rate, which corresponds to the base corrosion current density (exchange current rate). We incorporate the effect of temperature, *pH* level, and *Rh* by modifying Eq. (C.7) as Eq. (C.10).

$$f_{\Delta V(t)}(\Delta v(t)) = \int_{0}^{\infty} Ga \left( \Delta v(t) \right| \frac{\Delta n(t)}{\mu_{0}} + \sum_{i=1}^{n(t)} \frac{1}{\mu_{0} + \mu_{1}d_{i}(t)},$$
$$s_{0}\lambda \exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right) \int f_{\Delta N(t)}(\Delta n(t))d\Delta n(t) \quad (C.10)$$

Similarly, the increase in depth is described as

$$\Delta D_i(t) \sim Ga\left(\Delta d_i(t) \middle| \frac{1}{\mu_0 + \mu_1 d_i(t)}, \lambda \exp\left(\beta_1 Rh + \beta_2 pH + \frac{\beta_3}{T}\right)\right)$$
(C.11)

The parameters can be estimated by MLE, as shown in section C.5

#### C.4 Reliability and Remaining Life Prediction

Failure occurs when either the maximum depth of the corrosion pit reaches the depth threshold or the volume of the corrosion pit reaches the volume threshold as shown in Figure C.5, we develop the reliability models based on the two failure modes respectively.



Figure C.5 (a) Failure caused by corrosion volume growth (b) Failure caused by

corrosion depth growth

We use  $GA(\cdot)$  to denote the cumulative distribution function (CDF) of a Gamma distribution. We define the reliability in this situation as the probability that the component (structure) survives in time t+1 given that it has not failed at time t. Let  $T_f$  denote the failure time, the reliability based on corrosion volume growth is

$$R_{\nu}(t+1) = P_{\nu}(T_{f} > t+1)$$

$$= P(\Delta V(t) < c_{\nu} - \nu(t))$$

$$= \int_{0}^{\infty} GA \left( c_{\nu} - \nu(t) \right) \frac{\Delta n(t)}{\mu_{0}} + \sum_{i=1}^{n(t)} \frac{1}{\mu_{0} + \mu_{1}d_{i}(t)},$$

$$s_{0}\lambda \exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right) f_{\Delta N(t)}(\Delta n(t)) d\Delta n(t)$$
(C.12)

Theoretically, the maximum depth of the pit is the depth of the deepest pixel in the pit and the reliability should be estimated accordingly. However, as the center of the pit is corroded first and has the largest corrosion time, for simplicity, we assume that the maximum depth is the depth of the central pixel. However, in the estimation of the reliability regarding both the corrosion depth growth and volume growth simultaneously as discussed later, we will consider the depth of every pixel since the volume of a pixel is the area of the pixel times its depth. Let  $d^*(t)$  represent the maximum depth of a corrosion pit. Let  $c_d$  denote the depth threshold. For the depth failure mode, the reliability is given by:

$$R_{d}(t+1) = P_{d}(T_{f} > t+1)$$
  
=  $P(\Delta D^{*}(t) < c_{d} - d^{*}(t))$  (C.13)  
= $GA\left(c_{d} - d^{*}(t) \middle| \frac{1}{\mu_{0} + \mu_{1}d^{*}(t)}, \lambda \exp\left(\beta_{1}Rh + \beta_{2}pH + \frac{\beta_{3}}{T}\right)\right)$ 

As the corrosion volume growth is dependent on the depth growth of all the pixels and area growth of the sectors, we propose the following model to estimate the reliability regarding the two failure modes simultaneously:

$$R(t+1) = 1 - P(\exists d_{i}(t+1) \ge c_{d}, 1 \le i \le n(t)) -P(v(t) \ge c_{v}, \forall d_{i}(t+1) < c_{d}, 1 \le i \le n(t)) = \prod_{i=1}^{n(t)} GA_{i}(c_{d} - d_{i}(t)) - \int_{0}^{\infty} \int_{0}^{\infty} \dots \int_{0}^{\infty} \left( \left( 1 - GA_{v}\left(c_{v} - v(t) - \sum_{i=1}^{n(t)} s_{0} \Delta d_{i}(t)\right)\right) \right) f_{\Delta N(t)}(\Delta n(t)) \prod_{i=1}^{n(t)} Ga_{i}(\Delta d_{i}(t)) \right) \times d(\Delta n(t)) \prod_{i=1}^{n(t)} d(\Delta d_{i}(t))$$
(C.14)

where

$$GA_{i}\left(c_{d}-d_{i}\left(t\right)\right) = GA\left(c_{d}-d_{i}\left(t\right)\left|\frac{1}{\mu_{0}+\mu_{1}d_{i}\left(t\right)}, \ \lambda \exp\left(\beta_{1}Rh+\beta_{2}pH+\frac{\beta_{3}}{T}\right)\right)\right),$$

$$Ga_{i}\left(\Delta d_{i}\left(t\right)\right) = GA\left(\Delta d_{i}\left(t\right)\left|\frac{1}{\mu_{0}+\mu_{1}d_{i}\left(t\right)}, \ \lambda \exp\left(\beta_{1}Rh+\beta_{2}pH+\frac{\beta_{3}}{T}\right)\right)\right),$$

$$GA_{v}\left(c_{v}-v\left(t\right)-\sum_{i=1}^{n(t)}s_{0}\Delta d_{i}\left(t\right)\right) = GA\left(c_{v}-v\left(t\right)-\sum_{i=1}^{n(t)}s_{0}\Delta d_{i}\left(t\right)\left|\frac{\Delta n\left(t\right)}{\mu_{0}}, s_{0}\lambda \exp\left(\beta_{1}Rh+\beta_{2}pH+\frac{\beta_{3}}{T}\right)\right)\right)$$

The distribution of the remaining life is obtained using an iterative computational method as described next. Assume we have a series of corrosion pit growth measurements (3-D images) up to time t. The corrosion volume at that time is v(t), the maximum depth is  $d^{*}(t)$  and the number of corroded pixels is n(t). The remaining life is predicted as follows:

- (1) Set k = t.
- (2) Generate a random number  $\Delta n(k)$  using Eq. (C.4), update  $\{r_j(k)\}$  with  $\{r_j(k+1)\}$ . Generate a random number  $\Delta d^*(k)$  according to Eq. (C.5).
- (3) Generate a random number  $\Delta v(k)$  using Eq. (C.10).
- (4) Obtain  $d^*(k+1) = d^*(k) + \Delta d^*(k)$  and  $v(k+1) = v(k) + \Delta V(k)$ . If  $d^*(k+1) < c_d$ and  $v(k+1) < c_v$ , set k = k+1, go to step (2). Otherwise, the failure time is k+1and the remaining life T = k+1-t.
- (5) Repeat steps (2)-(4)  $N_p$  times.

Let  $T_m$  denote the *m* percentile remaining life predicted. The obtained 90% confidence interval of predicted remaining life is  $(T_{0.05}, T_{0.95})$ .

#### **C.5** Parameters Estimation

The parameters are estimated with MLE. The parameters of the proposed model include  $\{\mu_0, \mu_1, \lambda, \eta_0, \eta_1, \omega, \beta_1, \beta_2, \beta_3\}$ , where  $\{\eta_0, \eta_1, \omega\}$  correspond to the parameters that determine the corrosion growth in the radial directions,  $\{\mu_0, \mu_1, \lambda\}$  correspond to parameters that determine the corrosion growth in the depth direction,  $\{\beta_1, \beta_2, \beta_3\}$  is the

vector of the covariate parameters. Since we assume that the corrosion growth in both the radial directions and the depth direction to be independent, we obtain the likelihood functions for { $\mu_0$ ,  $\mu_1$ ,  $\lambda$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ } and { $\eta_0$ ,  $\eta_1$ ,  $\omega$ } independently. The estimation of these parameters requires the corrosion depth growth data and corrosion radius growth data for different pits. Assume we have M pits that are observed under different covariate sets. Let  $r_{mj}(t)$  be the radius in the  $j^{th}$  sector of the  $m^{th}$  pit at time t. Let  $d_{mi}(t)$  denote the depth of the  $i^{th}$  pixel of the  $m^{th}$  pit at time t, where  $1 \le m \le M$  and  $1 \le i \le n_m(t)$ ,  $n_m(t)$  is the total number of pixels of the  $m^{th}$  pit at time t. Let  $\left\{ \mathbf{x}(m) = \left\{ x_1(m), x_2(m), x_3(m) \right\}^t \right\}$  be the covariate vector values for the  $m^{th}$  pit. The variable  $x_1(m)$  is the *Rh* value of the  $m^{th}$  pit,  $x_2(m)$  is the *PH* value of the  $m^{th}$  pit

variable  $x_1(m)$  is the *Rh* value of the  $m^m$  pit,  $x_2(m)$  is the *pH* value of the  $m^m$  pit and  $x_3(m)$  is the temperature value of the  $m^{th}$  pit. For simplicity, let

$$f(\mathbf{x}(m)) = \exp\left(\beta_1 x_1(m) + \beta_2 x_2(m) + \frac{\beta_3}{x_2(m)}\right)$$
(C.15)

The likelihood function of  $\{\eta_0, \eta_1, \omega, \beta_1, \beta_2, \beta_3\}$  with corrosion increments in radial directions is given by

$$l_{d}\left(d_{11}(1),...,d_{mj}(h),...,d_{Mn_{M}(t)}(t)\middle|\mu_{0},\mu_{1},\lambda,\beta_{1},\beta_{2},\beta_{3}\right)$$

$$=\prod_{m=1}^{M}\prod_{h=1}^{t-1}\prod_{i=1}^{n_{m}(t)}\left(\frac{\left(d_{mi}(h+1)-d_{mi}(h)\right)^{\frac{1}{\mu_{0}+\mu_{1}d_{mi}(h)}-1}}{\Gamma\left(\frac{1}{\mu_{0}+\mu_{1}d_{mi}(h)}\right)\left(\lambda f\left(\mathbf{x}(m)\right)\right)^{\frac{1}{\mu_{0}+\mu_{1}d_{mi}(h)}}}\exp\left(-\frac{\left(d_{mi}(h+1)-d_{mi}(h)\right)}{\lambda f\left(\mathbf{x}(m)\right)}\right)\right)$$

(C.16)

Taking the natural logarithm of both sides yields

$$\ln l_{d} \left( d_{11}(1), \dots, d_{mj}(h), \dots, d_{Mn_{M}(t)}(t) \middle| \mu_{0}, \mu_{1}, \lambda, \beta_{1}, \beta_{2}, \beta_{3} \right)$$

$$= \sum_{m=1}^{M} \sum_{h=1}^{t-1} \sum_{i=1}^{n_{m}(t)} \left\{ \left( \frac{1}{\mu_{0} + \mu_{1}d_{mi}(h)} - 1 \right) \ln \left( d_{mi}(h+1) - d_{mi}(h) \right) - \left( \frac{1}{\mu_{0} + \mu_{1}d_{mi}(h)} \right) \ln \left( \lambda f\left( \mathbf{x}(m) \right) \right) \right\} - \left( \frac{d_{mi}(h+1) - d_{mi}(h)}{\lambda f\left( \mathbf{x}(m) \right)} - \ln \left( \Gamma \left( \frac{1}{\mu_{0} + \mu_{1}d_{mi}(h)} \right) \right) \right)$$
(C.17)

The MLE estimators are obtained by solving the following equations set

$$\frac{\partial \ln l_d}{\partial \mu_0} = 0, \quad \frac{\partial \ln l_d}{\partial \mu_1} = 0, \quad \frac{\partial \ln l_d}{\partial \lambda} = 0,$$

$$\frac{\partial \ln l_d}{\partial \beta_1} = 0, \quad \frac{\partial \ln l_d}{\partial \beta_2} = 0, \quad \frac{\partial \ln l_d}{\partial \beta_3} = 0$$
(C.18)

The likelihood function of  $\{\mu_0, \mu_1, \lambda\}$  is given by

$$l_{r}\left(r_{11}(1),...,r_{mj}(h),...,r_{Mn_{s}}(t)\big|\eta_{0},\eta_{1},\omega\right)$$

$$=\prod_{m=1}^{M}\prod_{h=1}^{t-1}\prod_{j=1}^{n_{s}}\left(\frac{1}{\Gamma\left(\frac{1}{\eta_{0}+\eta_{1}r_{mj}(h)}\right)\left(\omega f\left(\mathbf{x}(m)\right)\right)^{\frac{1}{\eta_{0}+\eta_{1}r_{mj}(h)}}}{\times\left(r_{mj}(h+1)-r_{mj}(h)\right)^{\frac{1}{\eta_{0}+\eta_{1}r_{mj}(h)^{-1}}}e^{-\frac{\left(r_{mj}(h+1)-r_{mj}(h)\right)}{\omega f\left(\mathbf{x}(m)\right)}}\right)$$
(C.19)

Taking the natural logarithm of both sides yields

$$\ln l_{r}\left(r_{11}(1), \dots, r_{mj}(h), \dots, r_{Mn_{s}}(t) \middle| \eta_{0}, \eta_{1}, \omega\right)$$

$$= \sum_{m=1}^{M} \sum_{j=1}^{t-1} \sum_{j=1}^{n_{s}} \left\{ \frac{\left(\frac{1}{\eta_{0} + \eta_{1}r_{mj}(h)} - 1\right) \ln\left(r_{mj}(h+1) - r_{mj}(h)\right) - \frac{\left(r_{mj}(h+1) - r_{mj}(h)\right)}{\omega f\left(\mathbf{x}(m)\right)} \right\} (C.20)$$

$$- \left(\frac{1}{\eta_{0} + \eta_{1}r_{mj}(h)}\right) \ln\left(\omega f\left(\mathbf{x}(m)\right)\right) - \ln\left(\Gamma\left(\frac{1}{\eta_{0} + \eta_{1}r_{mj}(h)}\right)\right) \right\}$$

The MLE estimators are obtained by solving the equations

$$\frac{\partial \ln l_r}{\partial \eta_0} = 0, \quad \frac{\partial \ln l_r}{\partial \eta_1} = 0, \quad \frac{\partial \ln l_r}{\partial \omega} = 0 \tag{C.21}$$

#### C.6 True Area Growth and Recommendations to Select Number of Sectors

In this section, we prove that the area growth of a sector proposed in Eq. (C.2) is a reasonable approximation of the true area growth and provide recommendations to select the number of sectors in a pit. The actual area growth of sector j is represented by Eq. (C.22), where the superscript denotes that it is obtained by considering the area growth as part of an annulus. In comparison, the area growth of a sector obtained by considering the area growth as a rectangle is represented by Eq. (C.23)

$$\Delta S_{j}^{\text{sec}}(t) = \frac{\pi}{n} \left( r_{j}^{2}(t+1) - r_{j}^{2}(t) \right)$$

$$= \frac{\xi}{2} \left( r_{j}^{2}(t+1) - r_{j}^{2}(t) \right)$$

$$\Delta S_{j}^{\text{rec}}(t) = 2r_{j}(t) \sin \frac{\xi}{2} \cdot \left( r_{j}(t+1) - r_{j}(t) \right)$$
(C.23)



Figure C.6 The ratio of half the angle and its sine value versus the angle of the sector

As shown in Figure C.6, the smaller the value of  $\xi$ , the closer the ratio  $\frac{\xi}{2} / \sin \frac{\xi}{2}$  to 1. When  $\xi \to 0$ ,  $\frac{\xi}{2} / \sin \frac{\xi}{2} \to 1$ . However, smaller  $\xi$  means higher computation cost. The choice of  $\xi$  depends on the need of precision and computation cost. As shown in section C.1,  $E(\Delta r_j(t)) = \omega / (\eta_0 + \eta_1 r_j(t))$ , where  $\eta_1 > 0$ ,  $\omega > 0$ . As given in Eq. (C.25),  $\frac{E((r_j(t+1)+r_j(t)))}{2r_j(t)}$  decreases as  $r_j(t)$  increases. When  $r_j(t) \to \infty$ ,  $\frac{E((r_j(t+1)+r_j(t)))}{2r_j(t)} \to 1$ .  $\frac{E((r_j(t+1)+r_j(t)))}{2r_j(t)} \to 1$ .  $\frac{E((r_j(t+1)+r_j(t)))}{2r_j(t)} = \frac{E(\Delta r_j(t))}{2r_j(t)} + 1$  (C.25) Accordingly,

$$\frac{\Delta S_{j}^{\text{sec}}(t)}{\Delta S_{j}^{\text{rec}}(t)} \rightarrow 1 \text{ as } r_{j}(t) \rightarrow \infty$$
(C.26)

In this appendix, with the data by Kariyawasam *et al.* (2012), the parameters in the radial directions are obtained as  $\{\hat{\eta}_0, \hat{\eta}_1, \hat{\omega}\} = \{0.314, 0.127, 0.083\}$ . Let  $S_j^{\text{sec}}(t)$  denote the cumulative area of sector j at time t,  $S_j^{\text{rec}}(t)$  denote the approximated cumulative area of sector j at time t. When t=150, the ratio  $S_j^{\text{sec}}(t)/S_j^{\text{rec}}(t)$  is plotted against the total number of sectors as shown in Figure C.7.



Figure C.7 The ratio of areas under different total number of sectors

It is observed that  $S_j^{\text{sec}}(t)/S_j^{\text{rec}}(t)$  decreases as  $n_s$  increases. However, as larger  $n_s$  indicates higher computation costs, we choose  $n_s = 36$ ,  $\xi = \frac{\pi}{18}$ , such that  $\frac{\xi}{2}/\sin\frac{\xi}{2} = 1.005$  and  $S_j^{\text{sec}}(t)/S_j^{\text{rec}}(t) = 1.015$ , which means the overall error is only 1.5%. The area growth obtained with Eq. (C.22) and Eq. (C.23) respectively are shown in Figure

C.8. In summary, the area growth of a sector approximated with Eq. (C.23) is accurate. Moreover, the growth rates over time are shown in Figure C.9.



Figure C.8 The comparison of the actual growth obtained with Eq. (C.22) and area

growth approximated with a rectangle obtained with Eq. (C.23)



Figure C.9 The corrosion growth rate over time

#### C.7 Case Study

In general, to accurately estimate the reliability of the corroded component with the proposed model, 3D image data of corrosion pits are required as inputs. These data can be

obtained with NDT devices such as laser scanning microscope (Krawczyk *et al.*, 2017), phased array and other 3D scanners (Turcotte *et al.*, 2016). Figure C.10 shows the representative images of the corrosion pits obtained with a laser scanning microscope (Krawczyk *et al.*, 2017). However, 3D images of pits are costly to obtain than longitudinal data. In the case where corrosion growth rate in both radial and depth directions are similar, depth data (longitudinal data) can be used to give an approximate estimate of the reliability metrics as an alternative.



**Figure C.10** Representative images of corrosion pits obtained with the (a) laser intensity image and (b) the corresponding 3D topography image (Krawczyk *et al.*, 2017)

In this case study, we demonstrate the performance of the model using corrosion depth growth rate data and depth growth data since the volume growth data are not readily available. The two datasets are the corrosion depth growth rates data by Kariyawasam *et al.* (2012) as shown in Figure C.2 and the corrosion depth growth data of ships by Soares *et al.* (2009). In the first dataset, the measurements are obtained under one operating

condition only. In the second dataset, the corrosion data are obtained under six different operating conditions. We describe the datasets and the corresponding results as follows.

#### C.7.1 Kariyawasam's dataset

To illustrate the use of the model, we fit the corrosion depth growth rates data with the depth growth model given in Eq. (C.5) and estimate its parameters. We censor the data at depth d(t) = 3mm and estimate the parameters  $\{\hat{\mu}_0, \hat{\mu}_1, \hat{\lambda}\} = \{0.314, 0.127, 0.083\}$ using MLE as shown in section C.5. Assuming the corrosion growths in the depth and the radial directions are determined by the of same set parameters, i.e.,  $\{\mu_0, \mu_1, \lambda\} = \{\eta_0, \eta_1, \omega\}$ , we simulate the 3-D growth data of the corrosion pit with the estimated parameters as shown in Figure C.11. Suppose we are considering the metal loss of a pipe (Phillips & Johnston) with  $\sigma_{flow} = 200Mpa$ ,  $t_p = 4mm$ , D = 19mm,  $\underline{L} = 10mm$ ,  $v_0$ =400 *cubic mm* and the operating pressure of the pipe is  $P_0 = 79.83Mpa$ . The critical corrosion pit volume  $c_v = 70 \ cubic \ mm$  and the critical depth threshold is  $c_d = 4mm$ .

With these estimated parameters, the expectation and the 90% confidence intervals of the maximum depth growth and volume growth are obtained as shown in Figures C.12 and C.13 respectively. We observe that the corrosion pit maximum depth and volume growth behave differently: the corrosion pit maximum depth growth rate is largest at the start of the corrosion process and decreases with time while the corrosion volume grows gradually,

and the volume growth rate increases over time. To evaluate the accuracy of the proposed model, we compare the reliability function of the pit's maximum depth estimated with the proposed model (Eq. (C.5)) and the other four widely used models: the Gamma process (Van Noortwijk, 2009), the Brownian Motion process (Wang, 2010), the inverse Gaussian process (Ye *et al.*, 2014) and the Markov Chains (Caleyo *et al.*, 2009). Figure C.14 shows the comparison results of the five approaches and the real corrosion pits' failure data, which are obtained by bootstrapping the corrosion growth incremental data in Figure C.2. We use the data in Figure C.2 with an original depth smaller than 5mm. The parameters of the four approaches are estimated with bootstrapped data by time t = 15.

As shown in Figure C.14, the proposed model has the most accurate reliability estimation (based on the mean squared error (MSE) as shown in Table C.1). There are a few time instants when the MC and inverse Gaussian process have better estimates. Although the proposed model can accurately estimate the time that the maximum corrosion depth reaches its threshold, the reliability estimation of the corroded unit by solely relying on considering the depth growth is insufficient and may lead to inaccurate results as shown in Figure C.15. We observe that by time t = 24 the reliability regarding depth growth is high with a value of 0.75 and the condition of the unit can be considered as relatively "safe". However, if we consider the failure caused by the corrosion pit volume growth, we find that the reliability is significantly smaller with a value of 0.16 and the unit is considered as

"unsafe". It is evident that the decision made solely on the depth growth is deficient and should be improved by considering the two failure modes simultaneously as shown in Figure C.15.

Models	Gamma	Brownian Inverse		Markov	The	
	process	Motion Gaussian		Chain	proposed	
		process	process		model	
MSE	0.0126	0.0166	0.0094	0.0104	0.0014	

Table C.1 The MSEs of the reliability estimated using the five approaches



**Figure C.11** Selected simulated corrosion pit 3-D images from t=15 to t=45



Figure C.12 The predicted maximum depth growth



Figure C.13 The predicted volume growth



Figure C.14 The comparison of the proposed model with others when estimating the

reliability regarding corrosion pit's maximum depth



Figure C.15 Reliability when considering different failure modes



Figure C.16 The comparison of the proposed model and the physical model

To validate the proposed corrosion volume growth model, we compare the corrosion volume growth results with the physics-based models (Harlow *et al.*, 1998, Kondo *et al.*, 1989). According to Kondo *et al.* (1989) and Harlow *et al.* (1998), the corrosion volume growth rate is a constant determined by physical parameters of the corrosion process. From Figure C.13, we observe that in the later stage of the corrosion volume growth process, as time increases the corrosion volume growth shows a linear relationship with time. We fit the corrosion volume growth data from t = 25 years to t = 45 years with a linear regression model as shown in Figure C.16. This shows that the results from the stochastic model match that of the physics-based model.

#### C.7.2 Soares's Dataset

The model is also validated with the corrosion depth growth data obtained from deck plating above the ballast tanks of the crude oil tankers during the six ships' service life in some particular routes through the Pacific Ocean as described by Soares et al. (2009). The operating conditions of the 6 ships are shown in Table C.2. The data sample size is increased by generating corrosion data with the deterministic model in the paper (Soares et al., 2009) and randomized with the Brownian motion process with drift, where the mean of the Brownian motion process is the same as that of the corrosion data. We censor the data at time 10 and estimate the parameters of the model as shown in section C.5. The corrosion data, the predicted expectation and 95% confidence interval of the corrosion pit depth in the 6 ships under different operating conditions are shown in Figure C.17. The means of the corrosion data are denoted in blue dash-dot lines. We observe that the model accurately predicts the expectation of the corrosion depth under different operating conditions. The MSE of the predicted mean path and the percent error (PE) of failure times are calculated and shown in Table C.2. Let  $t_0$  denote the censoring time, the PE are defined as follows

$$PE = \frac{T_f - T_r}{T_r}$$

where  $d^*(k)$  is the maximum depth of the pit at time k.  $\hat{d}^*(k)$  is the predicted mean maximum depth of the pit at time k.  $T_f$  and  $T_r$  are the observed and true failure time.

	Ship 1	Ship 2	Ship 3	Ship 4	Ship 5	Ship 6
Stresses	<i>T</i> =283.85	<i>T</i> =289.75	<i>T</i> =291.55	<i>T</i> =293.35	<i>T</i> =293.55	<i>T</i> =296.05
	<i>Rh</i> =85.5%	Rh=82.8%	<i>Rh</i> =83.3%	Rh=79.8%	<i>Rh</i> =82.3%	<i>Rh</i> =81.6%
	<i>pH</i> = 8.16	<i>pH</i> =8.15	<i>pH</i> =8.15	<i>pH</i> =8.14	<i>pH</i> =8.18	<i>pH</i> =8.18
PE of T	1.685	1.163	-1.625	3.870	-0.132	-1.644
MSE	0.000067	0.000130	0.000133	0.000406	0.000056	0.000063

Table C.2 The time to failure prediction accuracy



**Figure C.17** The corrosion depth prediction when data are censored at t=10



Figure C.18 The corrosion volume prediction when data are censored at t=10



Figure C.19 The reliability when considering both depth growth and volume growth

From Table C.2, it is shown that the proposed model accurately predicts the failure times of the corroded components subject to corrosion. The predicted corrosion volume growth paths are shown in Figure C.18. The reliability function considering both the corrosion depth growth and volume growth is shown in Figure C.19 when the maximum depth is 1.2 and the volume threshold is 2.0.

#### **C.8** Conclusions

The existing corrosion pit growth models are limited as they only capture the depth of the pit and may result in inaccurate life predictions of the corroded units. In this appendix, we propose a general degradation model that describes the corrosion pit volume growth over time. This is the first stochastic model that captures the corrosion volume growth when a critical amount of volume loss leads to failures of a component. The corrosion growth of the pit is modeled as a state-dependent Gamma process. The reliability model considers both the corrosion volume growth and depth growth and is more accurate than existing models. The influence of environmental factors is incorporated into the model by considering the corrosion chemistry reactions. Case studies show that the proposed model predicts the remaining life and reliability of the corroded units more accurately than existing models such as Markov Chains, Gamma process, inverse Gaussian process and Brownian Motion process. The agreement of the proposed model with a physics-based

model is verified with real data. The limitation of the work is that the pits overlap and the coalescence process is not thoroughly discussed or quantitively validated. We will address these topics in future research.

#### APPENDIX D THE CORROSION VOLUME THRESHOLD DERIVATION

Corrosion may lead to failure in different ways, including leakage, burst or sudden rupture. For example, leakage occurs when the corrosion pit depth of a pipeline equals its wall thickness. When the remaining strength of the pipeline wall is too small to resist inner fluid or gas pressure, the burst will happen even when the corrosion pit depth is smaller than the thickness of the pipeline wall (Amirat *et al.*, 2006, Jianan *et al.*, 1997). Extensive research has been conducted on predicting the time for a corrosion pit to reach the wall thickness threshold. The work on predicting the failure time when the volume of the corrosion reaches a threshold is scarce. We consider two thresholds: the depth threshold  $c_d$  and the volume threshold  $c_y$ . Ma *et al.* (2011) estimate the remaining strength of a pipeline as

$$P = \frac{\sigma_{flow} \cdot 2 \cdot t_p}{D} \left[ \frac{1 - \frac{A}{A_0}}{1 - \frac{A}{A_0} \cdot \frac{1}{M}} \right]$$
(D.1)

where

$$M = \sqrt{1 + 0.6275 \frac{L^2}{Dt_p} - 0.003375 \left(\frac{L^2}{Dt_p}\right)^2}$$
(D.2)

where  $\sigma_{flow}$  is the yield strength of the pipe material.  $t_p$  is the wall thickness of the pipe. *D* is the outer diameter of the pipe. *A* is the local area of metal loss in the longitudinal plane,  $A_0$  is the local original metal area, *M* is the bulging stress

magnification factor (Folias factor).  $P_0$  is the operating pressure. When  $P \le P_0$ , the pipe fails. For the pipe, the failure threshold for the remaining strength is  $P_0$ . When  $P=P_0$ , the critical corrosion loss area is

$$c_{A} = \frac{A_{0}(\sigma_{flow} \cdot 2 \cdot t_{p} - PD)}{\sigma_{flow} \cdot 2 \cdot t_{p} - \frac{PD}{M}}$$
(D.3)

In the proposed model, we are concerned with volume loss within a fixed local area. Using the criteria above, after modifications of Eq. (D.3), we propose the critical volume loss threshold as

$$c_{v} = \frac{v_{0}(\sigma_{flow} \cdot 2 \cdot t_{p} - PD)}{\sigma_{flow} \cdot 2 \cdot t_{p} - \frac{PD}{M}}$$
(D.4)

where  $\underline{M}$  is adjusted accordingly as

$$\underline{M} = \sqrt{1 + 0.6275 \frac{\underline{L}^2}{Dt_p} - 0.003375 \left(\frac{\underline{L}^2}{Dt_p}\right)^2}$$
(D.5)

where  $\underline{L}$  is the length of the defined area. For example, if we are interested in the volume loss of a whole pipeline,  $\underline{L}$  is the length of the pipeline.  $v_0$  is the local original metal volume.

## APPENDIX E PARAMETERS ESTIMATION IN THE CORROSION VOLUME GROWTH PROCESS BASED ON IIG PROCESS

The parameters are estimated with MLE. The parameters of the proposed model include  $\{\mu_0, \mu_1, \lambda, \eta_0, \eta_1, \omega, \beta_1, \beta_2, \beta_3\}$ , where  $\{\eta_0, \eta_1, \omega\}$  correspond to parameters that control corrosion growth in radial directions,  $\{\mu_0, \mu_1, \lambda\}$  correspond to parameters that control corrosion growth in the depth direction.  $\{\beta_1, \beta_2, \beta_3\}$  is the vector of covariates parameters. Since we assume that the corrosion growth in both the radial directions and the depth direction to be independent, we obtain the likelihood functions for  $\{\eta_0, \eta_1, \omega, \beta_1, \beta_2, \beta_3\}$  and  $\{\mu_0, \mu_1, \lambda\}$  independently. The estimation of these parameters requires the corrosion area growth data and corrosion volume growth data for different pits. Assume we have M pits which are observed under different covariate sets. Let  $r_{mj}(t)$  be the radius in the  $j^{th}$  sector of the  $m^{th}$  pit at time t. Let  $\delta_{mj}(t) = r_{mj}(t+1) - r_{mj}(t)$  be the corrosion growth increment of the  $j^{th}$  sector of the  $m^{th}$ pit in (t, t+1). Let  $\left\{ \mathbf{x}(m) = \left\{ x_1(m), x_2(m), x_3(m) \right\}' \right\}$  be the covariate set values for the  $m^{th}$  pit. The variable  $x_1(m)$  is the Rh value of the  $m^{th}$  pit,  $x_2(m)$  is the pH value of the  $m^{th}$  pit and  $x_3(m)$  is the temperature value of the  $m^{th}$  pit. For simplicity, let

$$f(\mathbf{x}(m)) = \exp\left(\beta_1 x_1(m) + \beta_2 x_2(m) + \frac{\beta_3}{x_2(m)}\right)$$
(E.1)

The likelihood function of  $\{\eta_0, \eta_1, \omega, \beta_1, \beta_2, \beta_3\}$  with corrosion increments in radial directions is given by

$$l_{r}\left(\delta_{11}(1), ..., \delta_{mj}(h), ..., \delta_{Mn_{s}}(t) \middle| \eta_{0}, \eta_{1}, \omega, \beta_{1}, \beta_{2}, \beta_{3}\right)$$

$$=\prod_{m=1}^{M} \prod_{h=1}^{t-1} \prod_{j=1}^{n_{s}} \left\{ \frac{\frac{\omega^{1/2}\left( \left(\eta_{0} + \eta_{1}r_{j}(t)\right) f\left(\mathbf{x}(m)\right)\right)}{(2\pi)^{1/2} \delta_{mj}(h)^{3/2}} \times \exp\left( \frac{-\omega\left(\delta_{mj}(h) - \left(\eta_{0} + \eta_{1}r_{j}(t)\right) f\left(\mathbf{x}(m)\right)\right)^{2}}{2\left(\left(\eta_{0} + \eta_{1}r_{j}(t)\right) f\left(\mathbf{x}(m)\right)\right)^{2} \delta_{mj}(h)} \right) \right\}$$
(E.2)

Taking the natural logarithm of both sides yields:

$$\ln l_{r} \left( \delta_{11}(1), \dots, \delta_{mj}(h), \dots, \delta_{Mn_{s}}(t) \middle| \eta_{0}, \eta_{1}, \omega, \beta_{1}, \beta_{2}, \beta_{3} \right) \\ = \sum_{m=1}^{M} \sum_{h=1}^{r_{s}} \sum_{j=1}^{n_{s}} \left\{ \frac{1}{2} \ln \omega + \ln \left( \left( \eta_{0} + \eta_{1} r_{j}(t) \right) f \left( \mathbf{x}(m) \right) \right) - \frac{1}{2} \ln 2\pi - \frac{3}{2} \ln \delta_{mj}(h) \right\} \\ + \frac{-\omega \left( \delta_{mj}(h) - \left( \eta_{0} + \eta_{1} r_{j}(t) \right) f \left( \mathbf{x}(m) \right) \right)^{2}}{2 \left( \left( \eta_{0} + \eta_{1} r_{j}(t) \right) f \left( \mathbf{x}(m) \right) \right)^{2} \delta_{mj}(h)} \right\}$$
(E.3)

The MLE estimators are obtained by solving the following equations set

$$\frac{\partial \ln l_r}{\partial \eta_0} = 0, \quad \frac{\partial \ln l_r}{\partial \eta_1} = 0, \quad \frac{\partial \ln l_r}{\partial \omega} = 0, \quad \frac{\partial \ln l_r}{\partial \beta_1} = 0, \quad \frac{\partial \ln l_r}{\partial \beta_2} = 0, \quad \frac{\partial \ln l_r}{\partial \beta_3} = 0$$
(E.4)

# APPENDIX F PROOFS OF THE DEGRADATION BRANCHING

### EQUATIONS

#### F.1 The PDF of the Total Degradation, the Reliability and the Expected Number of

#### **Branches**

In this section, the PDF of the total degradation, the expected number of branches and reliability at time t are derived.

Let 
$$h_i\left(x\left|t-\sum_{j=1}^i\tau_j\right)\right)$$
 denote the conditional PDF of  $X_i\left(t-\sum_{j=1}^i\tau_j\right)$ , the degradation of the

 $i^{th}$  branch, given the initiation times of the first n branches are known as  $\{\tau_i, 1 \le i \le n\}$ . Let  $h(x|t, \{\tau_i, 1 \le i \le n\})$  denote the conditional PDF of the total degradation of these

branches and the first degraded unit at time t, which is  $\sum_{i=1}^{n} X_i \left( t - \sum_{j=1}^{i} \tau_j \right) + X_0(t)$ .

Accordingly, given there are n branches at time t, the conditional PDF of the total degradation is

$$h^{*}(x|n,t) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{i=1}^{n-1}\tau_{i}} \prod_{i=1}^{n} f(\tau_{i}) \frac{1}{g(n|t)} \left(1 - F\left(t - \sum_{i=1}^{n}\tau_{i}\right)\right) h(x|t, \{\tau_{i}, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_{i}$$

where  $\prod_{i=1}^{n} f(\tau_i)$  is the probability that the initiation times of the *n* branches are  $\{\tau_i, 1 \le i \le n\}, \left(1 - F\left(t - \sum_{i=1}^{n} \tau_i\right)\right)$  is the probability that no branches are initiated in

 $\left(\sum_{i=1}^{n} \tau_{i}, t\right), g(n|t)$  is the probability mass function (PMF) of the number of branches n

at time t. Let  $n_b$  denote the limit of the number of branches, the PDF of the total degradation is

$$h^{*}(x|t) = \sum_{n=0}^{n_{b}} g(n|t)h^{*}(x|n,t)$$

$$= \sum_{n=0}^{n_{b}} g(n|t)\int_{0}^{t} \cdots \int_{0}^{t-\sum_{i=1}^{n-1}\tau_{i}} \prod_{i=1}^{n} f(\tau_{i}) \frac{1}{g(n|t)} \left(1 - F\left(t - \sum_{i=1}^{n}\tau_{i}\right)\right) h(x|t, \{\tau_{i}, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_{i}$$

$$= \sum_{n=0}^{n_{b}} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{i=1}^{n-1}\tau_{i}} \prod_{i=1}^{n} f(\tau_{i}) \left(1 - F\left(t - \sum_{i=1}^{n}\tau_{i}\right)\right) h(x|t, \{\tau_{i}, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_{i}$$
(F.1)

Q.E.D.

Eq. (4.7) in Chapter 4 is derived as follows. When the threshold is c, the reliability is

$$\begin{split} R(t) &= \int_{x=-\infty}^{c} h^{*}(x|t) \\ &= \int_{x=-\infty}^{c} \sum_{n=0}^{n_{b}} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{i=1}^{n-1} \tau_{i}} \prod_{i=1}^{n} f(\tau_{i}) \left( 1 - F\left(t - \sum_{i=1}^{n} \tau_{i}\right) \right) h(x|t, \{\tau_{i}, 1 \le i \le n\}) \prod_{i=1}^{n} d\tau_{i} dx \\ &= \sum_{n=0}^{n_{b}} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{i=1}^{n-1} \tau_{i}} \prod_{i=1}^{n} f(\tau_{i}) \left( 1 - F\left(t - \sum_{i=1}^{n} \tau_{i}\right) \right) \int_{x=-\infty}^{c} h(x|t, \{\tau_{i}, 1 \le i \le n\}) dx \prod_{i=1}^{n} d\tau_{i} \end{split}$$
(F.2)

where  $H(x|t, \{\tau_i, 1 \le i \le n\})$  denotes the conditional CDF of the total degradation given  $\{\tau_i, 1 \le i \le n\}$  are known.

Q.E.D.

When  $n_b = \infty$ , the expected number of branches by time t, which is obtained with Eq.

(4.2) in Chapter 4, is proved as follows. First, let

N|t| = the number of branches initiated by time t,

M(t) = the expected number of branches at time t = E(N|t),

 $S_N(t)$  = the time by the initiation of the  $N^{th}$  branch,  $S_N(t) = \tau_1 + \ldots + \tau_N = \sum_{i=1}^N \tau_i$ 

The probability that there are N branches by t is the same as the probability that time t lies between the initiation times of the  $N^{th}$  branch and the  $N + 1^{th}$  branch. Let  $\Theta_N(t)$  denote the CDF of  $S_N(t)$ , thus

$$P(N|T < t) = 1 - \Theta_N(t)$$
$$P(N|t) = P(N|T = t)$$
$$= \Theta_N(t) - \Theta_{N+1}(t)$$

The expected value of N|t is then
$$M(t) = \sum_{N=0}^{\infty} NP(N|t)$$
$$= \sum_{N=0}^{\infty} N(\Theta_N(t) - \Theta_{N+1}(t))$$
$$= \Theta(t) + \sum_{N=1}^{\infty} \Theta_{N+1}(t)$$

As  $F(t) = \Theta_1(t)$ ,  $\Theta_{N+1}(t)$  can be obtained with the following equation

$$\Theta_{N+1}(t) = \int_{0}^{t} \Theta_{N}(t-x) f(x) dx$$
$$M(t) = \Theta(t) + \sum_{N=1}^{\infty} \int_{0}^{t} \Theta_{N}(t-x) f(x) dx$$
$$= \Theta(t) + \int_{0}^{t} \sum_{N=1}^{\infty} \Theta_{N}(t-x) f(x) dx$$
$$= \Theta(t) + \int_{0}^{t} M(t-x) f(x) dx$$

Let

$$m(t) = \frac{dM(t)}{dt}$$

Elsayed (2012) obtains m(t) as follows. First  $m^*(s)$  is obtained as

$$m^{*}(s) = \frac{f^{*}(s)}{1 - f^{*}(s)}$$
(F.3)

Take the inverse Laplace transform, m(t) can be obtained with

$$m(t) = \mathcal{L}^{-1}\left(\frac{f^{*}(s)}{1 - f^{*}(s)}\right)$$
(F.4)

Accordingly, the expected number of branches at time t is

$$M(t) = \int_{0}^{t} m(\tau) d\tau$$
 (F.5)

Q.E.D.

#### F.2 The Expectation of the Total Degradation and the Mean Residual Life

Eqs. (4.3), (4.4) and (4.5) in Chapter 4 are proved in this Appendix. When the times to initiate new branches follow a distribution with PDF of  $f(\tau)$ , we propose a model to capture the expectation of the corresponding DBP. This method can be used to estimate the expectation of any DBP if the expectation of the number of branches by time t is available. Given there are n branches at time t, the total degradation increment  $dX^*(t)$  in (t, t+dt) is

$$dX^*(t|n) = \sum_{i=0}^n dX^*_i(t)$$

As the degradation growth of the n branches are independent, we have

$$E\left(dX^{*}(t|n)\right) = \sum_{i=0}^{n} E\left(dX_{i}^{*}(t)\right)$$

We take dt to be infinitely small such that no new branches are initiated in (t, t+dt). The expectation of the total degradation increment  $dX^*(t)$  is

$$E\left(X^{*}\left(t\right)\right) = \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} g\left(n|\tau\right) E\left(dX^{*}\left(\tau|n\right)\right)\right)$$
  
$$= \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} \left(g\left(n|\tau\right) \left(\sum_{i=0}^{n} E\left(dX_{i}^{*}\left(\tau\right)\right)\right)\right)\right)$$
(F.6)

$$E\left(X^{*}(t)\right) = \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} \left(g\left(n|\tau\right)\left(\sum_{i=0}^{n} E\left(dX_{i}^{*}(\tau)\right)\right)\right)\right)$$
$$= \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} \left(g\left(n|\tau\right)(n+1)\mu d\tau\right)\right)$$
$$= \int_{\tau=0}^{t} \left(\sum_{n=0}^{n_{b}} g\left(n|\tau\right)n\mu d\tau E\left(N|\tau\right) + \sum_{n=0}^{n_{b}} g\left(n|\tau\right)\mu d\tau\right)$$
$$= \mu \left(\int_{\tau=0}^{t} E\left(N|\tau\right)d\tau + t\right)$$
(F.7)

Q.E.D.

The MRL is derived as follows. First, suppose by time t the total degradation is x(t), the total number of branches is n(t) and the last branch is initiated at  $t_i$ . The expectation of the total degradation at time  $t_r > t$  is

$$\begin{split} &E\left(X^{*}\left(t_{r}|t,\ t_{l},\ n(t)\right)\right) = \sum_{i=0}^{n(t)} E\left(X^{*}_{i}\left(t_{r}|t,\ t_{l},\ n(t)\right)\right) + E\left(X^{*}_{n(t)+}\left(t_{r}|t,\ t_{l},\ n(t)\right)\right) \\ &= \sum_{i=0}^{n(t)} E\left(X^{*}_{i}\left(t_{r}|t,\ t_{l},\ n(t)\right)\right) + \sum_{n=1}^{n_{b}-n(t)} \left[\sum_{k=1}^{n} \left(\sum_{j=1}^{t_{r}-t_{l}} \cdots \int_{0}^{t_{r}-t_{l}} \prod_{j=1}^{k} f\left(\tau_{j}\right)g\left(n-k\left|t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}\right|\right)\right) \\ &\times \frac{1}{g\left(0|t-t_{l}\right)} E\left(X_{k}\left(t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}\left|n\right)\right)\prod_{j=1}^{k} d\tau_{j}\right)\right) \\ &= \int_{\tau=t}^{t_{r}} \sum_{i=0}^{n(t)} E\left(dX^{*}_{i}\left(\tau\right)\right) + \sum_{n=1}^{n_{b}-n(t)} \left[\sum_{k=1}^{n} \left(\sum_{j=1}^{t_{r}-t_{l}} \cdots \int_{0}^{t_{r}-t_{l}} \prod_{j=1}^{k-1} f\left(\tau_{j}\right)g\left(n-k\left|t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}\right|n\right)\right)\prod_{j=1}^{k} d\tau_{j}\right)\right) \\ &\times \frac{1}{g\left(0|t-t_{l}\right)} E\left(X_{k}\left(t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}\left|n\right)\right)\prod_{j=1}^{k} d\tau_{j}\right)\right) \\ &\times \frac{1}{g\left(0|t-t_{l}\right)} E\left(X_{k}\left(t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}\left|n\right)\right)\prod_{j=1}^{k} d\tau_{j}\right)\right) \end{split}$$

(F.	8)
(1.	0)

where  $E(X_i^*(t_r|t, t_l, n(t)))$  denotes the expectation of the degradation of branch *i* in  $(t, t_r)$  given that the total degradation is x(t) at time *t*, the total number of branches is n(t) and the last branch is initiated at  $t_l$  and  $E(X_{n(t)+}^*(t_r|t, t_l, n(t)))$  denotes the expectation of the total degradation of the  $n(t)^{th}$  and above branches. The MRL is obtained with the following equation.

$$MRL = \inf \left\{ t_{r} \int_{\tau=t}^{t_{r}} \sum_{i=0}^{n(t)} E(dX_{i}^{*}(\tau)) + \int_{\tau=t_{l}}^{t_{r}-t_{l}} \int_{\tau=t_{l}}^{t_{r}-t_{l}} \cdots \int_{0}^{t_{r}-t_{l}} \int_{j=1}^{t_{r}} f(\tau_{j})g(n-k|t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}) \\ \times \frac{1}{g(0|t-t_{l})} E\left(X_{k}\left(t_{r}-t_{l}-\sum_{j=1}^{k}\tau_{j}|n\right)\right) \prod_{j=1}^{k} d\tau_{j}\right) \right\} = c-x(t) \right\}$$
(F.9)

Q.E.D.

## F.3 The Variance of the Total Degradation

The proof of Eq. (4.6), which is the variance of the total degradation, is provided in this Appendix. Given there are *n* branches at time *t*, let  $X_{+}^{*}(t|n) = \sum_{i=1}^{n} X_{i}^{*}(t|n)$ , according to the rule of squared sum, the conditional variance of the degradation of the *n* branches

is

$$Var(X_{+}^{*}(t|n)) = Var\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)$$
  
=  $E\left(\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)^{2}\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2}$  (F.10)  
=  $\sum_{i=1}^{n} \left(E\left(X_{i}^{*}(t|n)\right)\right)^{2} + 2\sum_{i_{1} < i_{2}}^{n} E\left(X_{i_{1}}^{*}(t|n)X_{i_{2}}^{*}(t|n)\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2}$ 

where

$$\begin{split} \sum_{i=1}^{n} \left( E\left(X_{i}^{*}\left(t|n\right)\right) \right)^{2} \\ &= \sum_{i=1}^{n} \int_{-\infty}^{\infty} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} h_{i}\left(x\left|t-\sum_{j=1}^{i}\tau_{j}\right) g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right)x^{2} \prod_{j=1}^{i} d\tau_{j} dx \\ &= \sum_{i=1}^{n} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right) E\left(X_{i}^{2}\left(t-\sum_{j=1}^{i}\tau_{j}\left|n\right)\right) \prod_{j=1}^{i} d\tau_{j} \right) (F.11) \\ &= \sum_{i=1}^{n} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right|\right) \\ &\qquad \times \left( Var\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\left|n\right)\right) + \left(E\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\left|n\right)\right)\right)^{2}\right) \prod_{j=1}^{i} d\tau_{j} \right) \end{split}$$

$$\begin{split} &\sum_{i=1}^{n} \left( E\left(X_{i}^{*}\left(t|n\right)\right) \right)^{2} \\ &= \sum_{i=1}^{n} \int_{-\infty}^{\infty} \int_{0}^{i} \cdots \int_{0}^{i-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right)\right) h_{i}\left(x\left|t-\sum_{j=1}^{i}\tau_{j}\right) x^{2} \prod_{j=1}^{i} d\tau_{j} dx \\ &= \sum_{i=1}^{n} \int_{0}^{i} \cdots \int_{0}^{i-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right)\right) E\left(X_{i}^{2}\left(t-\sum_{j=1}^{i}\tau_{j}\right|n\right)\right) \prod_{j=1}^{i} d\tau_{j} \quad (F.12) \\ &= \sum_{i=1}^{n} \int_{0}^{i} \cdots \int_{0}^{i-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right|\right) \\ &\qquad \times \left( Var\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right|n\right)\right) + \left(E\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right|n\right)\right)\right)^{2} \right) \prod_{j=1}^{i} d\tau_{j} \end{split}$$

For  $\exists l \le i_1 < i_2 \le i$ , let  $x_{i_1}$  and  $x_{i_2}$  denote the corresponding degradation value variable, we have

$$\begin{split} E\left(X_{i_{1}}^{*}\left(t|n\right)X_{i_{2}}^{*}\left(t|n\right)\right) \\ &= \int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{0}^{t}\cdots\int_{0}^{t-\sum_{j=1}^{j-1}\tau_{j}}\prod_{j=1}^{i_{2}-\tau_{j}}f\left(\tau_{j}\right)\frac{1}{g\left(n|t\right)}g\left(n-i_{2}\left|t-\sum_{j=1}^{i_{2}}\tau_{j}\right.\right)\right) \\ &\times h_{i_{1}}\left(x_{i_{1}}\left|t-\sum_{j=1}^{i_{1}}\tau_{j}\right.\right)h_{i_{2}}\left(x_{i_{2}}\left|t-\sum_{j=1}^{i_{2}}\tau_{j}\right.\right)x_{i_{1}}x_{i_{2}}\prod_{j=1}^{i_{2}}d\tau_{j}dx_{i_{1}}dx_{i_{2}}\right. \end{split}$$
(F.13)  
$$&= \int_{0}^{t}\cdots\int_{0}^{t-\sum_{j=1}^{j-1}\tau_{j}}\prod_{j=1}^{i_{2}-\tau_{j}}f\left(\tau_{j}\right)\frac{1}{g\left(n|t\right)}g\left(n-i_{2}\left|t-\sum_{j=1}^{i_{2}}\tau_{j}\right.\right) \\ &\times E\left(X_{i_{1}}\left(t-\sum_{j=1}^{i_{1}}\tau_{j}\right|n\right)\right)E\left(X_{i_{2}}\left(t-\sum_{j=1}^{i_{2}}\tau_{j}\right|n\right)\right)\prod_{j=1}^{i_{2}}d\tau_{j} \end{split}$$

The conditional variance of the degradation of the n branches given there are n branches is

$$\begin{split} &\operatorname{Var}\left(X_{+}^{*}(t|n)\right) = \operatorname{Var}\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right) \\ &= E\left[\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)^{2}\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2} \\ &= \sum_{i=1}^{n} \left(E\left(X_{i}^{*}(t|n)\right)\right)^{2} + 2\sum_{i_{1} < i_{2}}^{n} E\left(X_{i}^{*}(t|n)X_{j}^{*}(t|n)\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2} \\ &= \sum_{i=1}^{n} \left(\sum_{0}^{t} \cdots \sum_{0}^{t-\frac{i-1}{j-1}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i} \tau_{j}\right)\right) \\ &\quad \times \left(\operatorname{Var}\left(X_{i}\left(t-\sum_{j=1}^{i} \tau_{j}\right|n\right)\right) + \left(E\left(X_{i}\left(t-\sum_{j=1}^{i} \tau_{j}\right|n\right)\right)\right)^{2}\right) \prod_{j=1}^{i} d\tau_{j}\right) \\ &+ 2\sum_{\substack{i \ge 1\\i_{i} < i_{2}}}^{n} \left(\sum_{0}^{t} \cdots \sum_{0}^{t-\frac{i-1}{j-1}} \prod_{j=1}^{i-1} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i} \tau_{j}\right|\right) \\ &\quad \times E\left(X_{i_{1}}\left(t-\sum_{j=1}^{i} \tau_{j}\right|n\right)\right) E\left(X_{i_{2}}\left(t-\sum_{j=1}^{i} \tau_{j}\right|n\right)\right) \prod_{j=1}^{i} d\tau_{j}\right) \\ &- \left(\sum_{i=1}^{n} \int_{0}^{t} \cdots \int_{0}^{t-\frac{i-1}{j-1}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \frac{1}{g\left(n|t\right)} g\left(n-i\left|t-\sum_{j=1}^{i} \tau_{j}\right|\right) E\left(X_{i}\left(t-\sum_{j=1}^{i} \tau_{j}\right|n\right)\right) \prod_{j=1}^{i} d\tau_{j}\right)^{2} (F.14) \end{split}$$

The variance of the degradation of all the branches including the first degraded unit at time

t is

$$\begin{split} &\operatorname{Var}\left(X^{*}(t)\right) = \operatorname{Var}\left(X^{*}_{*}(t)\right) + \operatorname{Var}\left(X_{0}(t)\right) \\ &= E\left[\operatorname{Var}\left(X^{*}_{*}(t|n)\right)g\left(n|t\right) + \sum_{n=1}^{\infty}\left[E\left(X^{*}_{*}(t|n)\right)\right]^{2}g\left(n|t\right) + \operatorname{Var}\left(X_{0}(t)\right) \\ &= \sum_{n=1}^{\infty} E\left(X^{*}_{*}(t|n)\right)g\left(n|t\right)\right)^{2} \\ &= \left[\sum_{n=1}^{\infty} E\left(X^{*}_{*}(t|n)\right)g\left(n|t\right)\right]^{2} \\ &= \left[\sum_{n=1}^{\infty} E\left(X^{*}_{*}(t|n)\right)g\left(n|t\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n-t|t|-\sum_{j=1}^{1}\tau_{j}|n\right)\right]\left[\sum_{j=1}^{n}d\tau_{j}\right]^{2} \\ &= \left[\sum_{n=1}^{\infty} E\left(X^{*}_{*}(t|n)\right)g\left(n|t\right)g\left(n-t|t|-\sum_{j=1}^{1}\tau_{j}\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|\right)g\left(n|t|$$

When 
$$\dot{F}_{i}(x) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j}) x \prod_{j=1}^{i} d\tau_{j}$$
,  $g_{n-i} = g\left(n-i\left|t-\sum_{j=1}^{i}\tau_{j}\right|\right)$ ,  $g_{n} = g\left(n\left|t\right)$ ,  
 $m_{0} = Var\left(X_{0}\left(t\right)\right)$ ,  $m_{i}^{n} = E\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right|n\right)\right)$  and  $v_{i}^{n} = Var\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right|n\right)\right)$ , Eq.

(F.15) is revised as

$$Var(X^{*}(t)) = \sum_{n=1}^{n_{b}} \left( g_{n} \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right)^{2} \right) - \left( \sum_{n=1}^{n_{b}} \left( g_{n} \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right) \right)^{2} + v_{0} + \sum_{n=1}^{n_{b}} \left( g_{n} \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} \left( v_{i}^{n} + \left( m_{i}^{n} \right)^{2} \right) \right) \right) - \left( \sum_{i=1}^{n} \left( \dot{F}_{i} \left( \frac{g_{n-i}}{g_{n}} m_{i}^{n} \right) \right) \right)^{2} + 2\sum_{\substack{i_{1} \geq 1 \\ i_{1} < i_{2}}}^{n} \left( \dot{F}_{i_{2}} \left( \frac{g_{n-i_{2}}}{g_{n}} m_{i_{1}}^{n} m_{i_{2}}^{n} \right) \right) \right) \right)$$
(F.16)

Q.E.D.

# F.4 The Expectation and Variance of the Branches

The proofs of the expectation and variance of the degradation of the  $i^{th}$  branch, which are shown in Eqs. (4.9) and (4.10) in Chapter 4, are proposed in this section. The PDF of the degradation of the  $i^{th}$  branch is

$$h_{i}(x|t) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j})g_{i}\left(x|t-\sum_{j=1}^{i}\tau_{j}\right) \prod_{j=1}^{i} d\tau_{j}$$
(F.17)

The expectation of the degradation of the  $i^{th}$  branch is

$$E(X_{i}^{*}(t)) = \int_{-\infty}^{\infty} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j})g_{i}\left(x\left|t-\sum_{j=1}^{i}\tau_{j}\right)x\prod_{j=1}^{i}d\tau_{j}dx\right)$$
$$= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j})\int_{-\infty}^{\infty} g_{i}\left(x\left|t-\sum_{j=1}^{i}\tau_{j}\right)xdx\prod_{j=1}^{i}d\tau_{j}\right)$$
$$= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \prod_{j=1}^{i} f(\tau_{j})E\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right)\right)\prod_{j=1}^{i}d\tau_{j}$$
(F.18)

The variance of the degradation of the  $i^{th}$  branch is

$$\begin{aligned} \operatorname{Var}\left(X_{i}^{*}(t)\right) &= E\left(\left(X_{i}^{*}(t)\right)^{2}\right) - \left(E\left(X_{i}^{*}(t)\right)\right)^{2} \\ &= \int_{-\infty}^{\infty} \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} f\left(\tau_{j}\right) g_{i}\left(x\left|t - \sum_{j=1}^{i}\tau_{j}\right) x^{2} \prod_{j=1}^{i} d\tau_{j} dx \\ &- \left(\int_{-\infty}^{\infty} \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) g_{i}\left(x\left|t - \sum_{j=1}^{i}\tau_{j}\right) x \prod_{j=1}^{i} d\tau_{j} dx \right)^{2} \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \int_{-\infty}^{\infty} g_{i}\left(x\left|t - \sum_{j=1}^{i}\tau_{j}\right) x^{2} dx \prod_{j=1}^{i} d\tau_{j} dx \right)^{2} \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) \int_{-\infty}^{\infty} g_{i}\left(x\left|t - \sum_{j=1}^{i}\tau_{j}\right) x dx \prod_{j=1}^{i} d\tau_{j} d\tau_{j} \right)^{2} \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) E\left(X_{i}^{2}\left(t - \sum_{j=1}^{i}\tau_{j}\right)\right) \prod_{j=1}^{i} d\tau_{j} d\tau_{j} \right)^{2} \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}\tau_{j}} \prod_{j=1}^{i} f\left(\tau_{j}\right) E\left(X_{i}^{2}\left(t - \sum_{j=1}^{i}\tau_{j}\right)\right) \prod_{j=1}^{i} d\tau_{j} d\tau_{j} \right)^{2} \end{aligned}$$
(F.19)

Q.E.D.

# F.5 The Mean and Variance of the Total Degradation where Time to Branch Follows

# an Exponential Distribution

The proofs of Eqs. (4.12) and (4.13) in Chapter 4, which correspond to the mean and variance of the DBP process where time to branch follow an exponential distribution, are discussed in this section. In the case where N follows a Poisson distribution with

$$E(X^{*}(t)) = \mu\left(\int_{\tau=0}^{t} M(\tau)d\tau + t\right)$$
  
=  $\mu\left(\int_{\tau=0}^{t} \lambda \tau d\tau + t\right)$  (F.20)  
=  $\mu\left(\frac{t^{2}\lambda}{2} + t\right)$ 

Q.E.D.

The expectation of the degradation of the  $i^{th}$  branch given there are *n* branches is

$$\begin{split} & E\left(X_{i}^{*}(t|n)\right) \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i-1} \tau_{j}} \lambda e^{-\lambda \tau_{1}} \cdots \lambda e^{-\lambda \tau_{i}} \frac{\left(\lambda\left(t-\sum_{j=1}^{i} \tau_{j}\right)\right)^{n-i} e^{-\lambda\left(t-\sum_{j=1}^{i} \tau_{j}\right)}}{\frac{\left(\lambda t\right)^{n} e^{-\lambda t}}{n!} (n-i)!} \mu\left(t-\sum_{j=1}^{i} \tau_{j}\right) \prod_{j=1}^{i} d\tau_{j} \\ &= \begin{cases} \mu \frac{n!}{t^{n}} \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{i} \tau_{j}} \left(t-\sum_{j=1}^{i} \tau_{j}\right)^{n-i+1} d\tau_{i} \cdots d\tau_{1}, \ i \ge 2 \\ \mu \frac{n!}{t^{n}} \int_{0}^{t} (t-\tau_{1}) d\tau_{1}, \ i = 1 \end{cases} \end{split}$$
(F.21)

Accordingly, the expectation of the total degradation of the n branches given there are n branches is

$$E(X_{+}^{*}(t|n)) = \sum_{i=1}^{n} E(X_{i}^{*}(t|n))$$
  
=  $\sum_{i=1}^{n} \frac{n+1-i}{n+1} \mu t$   
=  $\frac{1}{2} n \mu t$  (F.22)

The variance of the total degradation of the n branches given there are n branches is

$$Var(X_{+}^{*}(t|n)) = E\left(\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)^{2}\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2}$$
$$= \sum_{i=1}^{n} \left(E\left(X_{i}^{*}(t|n)\right)\right)^{2} + 2\sum_{i_{1} < i_{2}}^{n} E\left(X_{i_{1}}^{*}(t|n)X_{i_{2}}^{*}(t|n)\right) - \left(E\left(\sum_{i=1}^{n} X_{i}^{*}(t|n)\right)\right)^{2}$$

$$=\sum_{i=1}^{n} \left( \int_{0}^{t} \dots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \lambda e^{-\lambda\tau_{1}} \dots \lambda e^{-\lambda\tau_{i}} \frac{1}{(\lambda t)^{n} e^{-\lambda t}} \left( \lambda \left( t - \sum_{j=1}^{i} \tau_{j} \right) \right)^{n-i} e^{-\lambda \left[ t - \sum_{j=1}^{i} \tau_{j} \right]} \\ \times \left( \sigma^{2} \left( t - \sum_{j=1}^{i} \tau_{j} \right) + \mu^{2} \left( t - \sum_{j=1}^{i} \tau_{j} \right)^{2} \right) \prod_{j=1}^{i} d\tau_{j} \right) \\ + 2\sum_{\substack{i_{1} \geq 1\\i_{1} < i_{2}}}^{n} \left( \int_{0}^{t} \dots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \dots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \lambda e^{-\lambda\tau_{1}} \dots \lambda e^{-\lambda\tau_{i_{1}}} \dots \lambda e^{-\lambda\tau_{i_{1}}} \dots \lambda e^{-\lambda\tau_{i_{2}}} \frac{\left( \lambda \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right) \right)^{n-i_{2}} e^{-\lambda \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right)}}{\frac{\left( \lambda t \right)^{n} e^{-\lambda t}}{n!}} \\ + 2\sum_{\substack{i_{1} \geq 1\\i_{1} < i_{2}}}^{n} \left( \int_{0}^{t} \dots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \lambda e^{-\lambda\tau_{1}} \dots \lambda e^{-\lambda\tau_{i_{1}}} \dots \lambda e^{-\lambda\tau_{i_{1}}} \dots \lambda e^{-\lambda\tau_{i_{1}}}} \frac{\left( \lambda \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right) \right)^{n-i_{2}}}{\frac{\left( \lambda t \right)^{n} e^{-\lambda t}}{n!}} \\ - \left( \sum_{\substack{i=1\\i=1}}^{n} \int_{0}^{t} \dots \int_{0}^{t-\sum_{j=1}^{i-1}\tau_{j}} \lambda e^{-\lambda\tau_{1}} \dots \lambda e^{-\lambda\tau_{i}} \frac{\left( \lambda \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right) \right)^{(n-i)} e^{-\lambda \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right)}}{\frac{\left( \lambda t \right)^{n} e^{-\lambda t}}{n!}} \mu \left( t - \sum_{j=1}^{i_{2}} \tau_{j} \right) \prod_{j=1}^{i_{1}} d\tau_{j} \right)^{2}$$

$$\begin{split} &= \sum_{i=1}^{n} \left( \frac{n!}{t^{n}} \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}} \left( \sigma^{2} \left( t - \sum_{j=1}^{i} \tau_{j} \right)^{n-i+1} + \mu^{2} \left( t - \sum_{j=1}^{i} \tau_{j} \right)^{n-i+2} \right) \prod_{j=1}^{i} d\tau_{j} \right) \\ &+ 2\sum_{\substack{i, \geq 1\\ i_{i} < i_{i}}}^{n} \left( \frac{n! \mu^{2}}{t^{n}} \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}} \cdots \int_{0}^{t-\frac{j-1}{j-1}} \left( t - \sum_{j=1}^{i} \tau_{j} \right) \left( t - \sum_{j=1}^{i} \tau_{j} \right)^{n-i+2} \prod_{j=1}^{i} d\tau_{j} \right) \\ &- \left( \sum_{\substack{i=1\\ i=1}}^{n} \frac{n! \mu}{t^{n}} \int_{0}^{t} \cdots \int_{0}^{t-\frac{j-1}{j-1}} \left( t - \sum_{j=1}^{i} \tau_{j} \right)^{n-i+1} \prod_{j=1}^{i} d\tau_{j} \right)^{2} \\ &= \sum_{i=1}^{n} \left( \frac{n+1-i}{n+1} t \sigma^{2} + \frac{(n+1-i)(n+2-i)}{(n+1)(n+2)} t^{2} \mu^{2} \right) \\ &+ 2\sum_{\substack{i, \geq 1\\ i_{i} < i_{2}}}^{n} \mu^{2} t^{2} \frac{(n+1-i_{2})(n+2-i_{1})}{(n+1)(n+2)} - \left( \sum_{i=1}^{n} \frac{n+1-i}{n+1} \mu t \right)^{2} \\ &= \frac{1}{2} \sigma^{2} tn + \frac{1}{3} \mu^{2} t^{2} n + \frac{1}{4} \mu^{2} t^{2} n (n-1) - \frac{1}{4} \mu^{2} t^{2} n^{2} \end{split}$$
(F.23)

In the proof of Eq. (F.23), we use the conclusions of Eqs. (F.24) and (F.25) as follows.

$$\begin{split} &\sum_{i=1}^{n} \left( \frac{n+1-i}{n+1} t \sigma^{2} + \frac{(n+1-i)(n+2-i)}{(n+1)(n+2)} t^{2} \mu^{2} \right) \\ &= t \sigma^{2} \left( n - \frac{(1+\ldots+n)}{n+1} \right) + t^{2} \mu^{2} \sum_{i=1}^{n} \left( 1 - \frac{i}{n+1} - \frac{i}{n+2} + \frac{i^{2}}{(n+1)(n+2)} \right) \\ &= t \sigma^{2} \left( n - \frac{\frac{n(1+n)}{2}}{n+1} \right) + t^{2} \mu^{2} \left( n - \frac{\frac{n(1+n)}{2}}{n+1} - \frac{\frac{n(1+n)}{2}}{n+2} + \frac{\frac{n(1+n)(1+2n)}{6}}{(n+1)(n+2)} \right) \end{split}$$
(F.24)  
$$&= \frac{1}{2} \sigma^{2} t n + \frac{1}{3} \mu^{2} t^{2} n \end{split}$$

$$2\sum_{\substack{i_{1}\geq 1\\i_{1}

$$= 2\mu^{2} t^{2} \sum_{\substack{i_{2}=2}}^{n} \sum_{\substack{i_{1}=1}}^{i_{2}-1} \left(1 - \frac{i_{2}}{n+1} - \frac{i_{1}}{n+2} + \frac{i_{1}i_{2}}{(n+1)(n+2)}\right)$$

$$= 2\mu^{2} t^{2} \sum_{\substack{i_{2}=2}}^{n} \left((i_{2}-1) - \frac{i_{2}(i_{2}-1)}{n+1} - \frac{(i_{2}-1)i_{2}}{2(n+2)} + \frac{(i_{2}-1)i_{2}^{2}}{2(n+1)(n+2)}\right) \quad (F.25)$$

$$= 2\mu^{2} t^{2} \left(\frac{n(n-1)}{2} - \frac{n(n-1)}{3} - \frac{n(n+1)(n-1)}{6(n+2)} + \frac{n(n-1)(3n+2)}{24(n+2)}\right)$$

$$= \frac{1}{4}\mu^{2} t^{2} n(n-1)$$$$

The variance of the total degradation of the branches at time t is

$$\begin{aligned} \operatorname{Var}\left(X_{+}^{*}(t)\right) &= E\left(\operatorname{Var}\left(X_{+}^{*}(t|n)\right)\right) + \operatorname{Var}\left(E\left(X_{+}^{*}(t|n)\right)\right) \\ &= \sum_{n=1}^{\infty} \left(\operatorname{Var}\left(X_{+}^{*}(t|n)\right)g\left(n|t\right)\right) + \sum_{n=1}^{\infty} \left(\left(E\left(X_{+}^{*}(t|n)\right)\right)^{2}g\left(n|t\right)\right) - \left(\sum_{n=1}^{\infty} E\left(X_{+}^{*}(t|n)\right)g\left(n|t\right)\right)^{2} \\ &= \sum_{n=1}^{\infty} \left(\frac{1}{2}\sigma^{2}t + \frac{1}{12}\mu^{2}t^{2}\right)ng\left(n|t\right) + \sum_{n=1}^{\infty} \left(\frac{1}{2}\mu tn\right)^{2}g\left(n|t\right) - \left(\sum_{n=1}^{\infty}\frac{1}{2}\mu tng\left(n|t\right)\right)^{2} \\ &= \sum_{n=1}^{\infty} \left(\frac{1}{2}\sigma^{2}t + \frac{1}{12}\mu^{2}t^{2}\right)ng\left(n|t\right) + \left(\frac{1}{2}\sigma^{2}t + \frac{1}{12}\mu^{2}t^{2}\right) \cdot 0 \cdot g\left(0|t\right) + \sum_{n=1}^{\infty} \left(\frac{1}{2}\mu tn\right)^{2}g\left(n|t\right) \\ &\quad + \left(\frac{1}{2}\mu t\cdot 0\right)^{2}g\left(0|t\right) - \left(\sum_{n=1}^{\infty}\frac{1}{2}\mu tng\left(n|t\right)\right)^{2} - \left(\frac{1}{2}\mu t\cdot 0\cdot g\left(0|t\right)\right)^{2} \end{aligned}$$

$$&= \sum_{n=0}^{\infty} \left(\frac{1}{2}\sigma^{2}t + \frac{1}{12}\mu^{2}t^{2}\right)ng\left(n|t\right) + \sum_{n=0}^{\infty} \left(\frac{1}{2}\mu tn\right)^{2}g\left(n|t\right) - \left(\sum_{n=1}^{\infty}\frac{1}{2}\mu tng\left(n|t\right)\right)^{2} \\ &= \left(\frac{1}{2}\sigma^{2}t + \frac{1}{12}\mu^{2}t^{2}\right)E\left(N|t\right) + \frac{1}{4}\mu^{2}t^{2}\left(\operatorname{Var}\left(N|t\right) + E^{2}\left(N|t\right)\right) - \left(\frac{1}{2}\mu tE\left(N|t\right)\right)^{2} \\ &= \frac{1}{3}\mu^{2}\lambda t^{3} + \frac{1}{2}\sigma^{2}\lambda t^{2} \end{aligned}$$

As the growth of the branches (including the first degraded unit) are independent, the variance of the total degradation is the sum of the variances of the branches  $Var(X_{+}^{*}(t))$  and the variance of the first degraded unit  $Var(X_{0}^{*}(t))$  as shown in Eq. (F.27)

$$Var(X^{*}(t)) = Var(X^{*}_{+}(t)) + Var(X^{*}_{0}(t))$$
  
$$= \frac{1}{3}\mu^{2}\lambda t^{3} + \frac{1}{2}\sigma^{2}\lambda t^{2} + \sigma t$$
 (F.27)

Q.E.D.

(F.26)

The MRL as shown in Eq. (4.18) in Chapter 4 is derived as follows. According to Eq. (F.9), for a general degradation process whose mean function and variance function of a branch in (t, t+dt) is  $\mu dt$ , the expectation of the total degradation increment in  $(t, t_r)$  is

$$E\left(X^{*}\left(t_{r}|t,t_{l},n(t)\right)\right)$$

$$=\mu\left(n(t)+1\right)\left(t_{r}-t\right)+\sum_{n=1}^{n_{b}-n(t)}\left(\frac{g\left(n|t_{r}-t_{l}\right)}{g\left(0|t-t_{l}\right)}\sum_{i=1}^{n}\frac{n+1-i}{n+1}\mu\frac{\left(t-t_{l}\right)^{n+1}}{\left(t_{r}-t_{l}\right)^{n}}\right) \quad (F.28)$$

$$=\mu\left(n(t)+1\right)\left(t_{r}-t\right)+\frac{\mu}{2g\left(0|t-t_{l}\right)}\sum_{n=1}^{n_{b}-n(t)}\left(g\left(n|t_{r}-t_{l}\right)n\frac{\left(t-t_{l}\right)^{n+1}}{\left(t_{r}-t_{l}\right)^{n}}\right)$$

According to Eq. (F.28), the MRL is

$$MRL = \inf\left\{t_r \left| \mu(n(t)+1)(t_r-t) + \frac{\mu}{2g(0|t-t_l)} \sum_{n=1}^{n_b-n(t)} \left(g(n|t_r-t_l)n\frac{(t-t_l)^{n+1}}{(t_r-t_l)^n}\right) = c - x(t)\right\}$$
(F.29)

Q.E.D.

# F.6 The Mean and Variance of the Branches where Time to Branch Follows an Exponential Distribution

We prove Eq. (4.19) in Chapter 4 by induction. First, we prove that it holds when i=1. The expectation of the degradation of the first branch can be obtained according to Eq. (F.30) as follows:

$$E\left(X_{1}^{*}(t)\right) = \int_{0}^{t} f\left(\tau_{1}\right) E\left(X_{1}\left(t-\tau_{1}\right)\right) d\tau_{1}$$
  
$$= \mu\left(t\int_{0}^{t} \lambda e^{-\lambda\tau_{1}} d\tau_{1} - \int_{0}^{t} \lambda e^{-\lambda\tau_{1}} \tau_{1} d\tau_{1}\right)$$
  
$$= \mu\left(t\left(1-e^{-\lambda t}\right) - \frac{1-(\lambda t+1)e^{-\lambda t}}{\lambda}\right)$$
  
$$= \mu\left(t-\frac{1}{\lambda} + \frac{e^{-\lambda t}}{\lambda}\right)$$
  
(F.30)

The result is the same as that obtained with Eq. (4.19) as follows

$$E(X_1^*(t)) = \mu\left(\left(t - \frac{1}{\lambda}\right) + \frac{e^{-\lambda t}}{\lambda}\right)$$
(F.31)

Second, assume that Eq. (4.19) holds for i=n, we prove that it also holds for i=n+1. According to Eq. (4.19)

$$E\left(X_{n}^{*}(t)\right) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) E\left(X_{n}\left(t-\sum_{j=1}^{n}\tau_{j}\right)\right) d\tau_{n} \cdots d\tau_{1}$$

$$= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) \mu\left(t-\sum_{j=1}^{i}\tau_{j}\right) d\tau_{n} \cdots d\tau_{1}$$

$$= \mu\left(\int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) \left(t-\sum_{j=1}^{i}\tau_{j}\right) d\tau_{n} \cdots d\tau_{1}$$

$$= \mu\left(\left(t-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1} \frac{(n-j)}{j!} t^{j} \lambda^{j-1}\right) e^{-\lambda t}\right)$$
(F.32)

According to the last two steps in Eq. (F.32), we have

$$\int_{0}^{t} \cdots \int_{0}^{n-1} f(\tau_{1}) \cdots f(\tau_{n}) \left( t - \sum_{j=1}^{n} \tau_{j} \right) d\tau_{n} \cdots d\tau_{1}$$

$$= \left( \left( t - \frac{n}{\lambda} \right) + \left( \sum_{j=0}^{n-1} \frac{(n-j)}{j!} t^{j} \lambda^{j-1} \right) e^{-\lambda t} \right)$$
(F.33)

Let k = j+1 and  $t = t - \tau_1$  where  $\tau_1$  is a constant, we have

$$\int_{0}^{t-\tau_{1}} \cdots \int_{0}^{\tau_{k}} f(\tau_{2}) \cdots f(\tau_{n+1}) \left( t - \tau_{1} - \sum_{k=2}^{n+1} \tau_{k} \right) d\tau_{n+1} \cdots d\tau_{2}$$

$$= \left( \left( t - \tau_{1} - \frac{n}{\lambda} \right) + \left( \sum_{k=0}^{n-1} \frac{(n-k)}{k!} (t - \tau_{1})^{k} \lambda^{k-1} \right) e^{-\lambda(t-\tau_{1})} \right)$$
(F.34)

According to Eqs. (4.9) and (F.34), the expectation of the  $n+1^{th}$  branch is derived as follows:

$$\begin{split} E\left(X_{n+1}^{*}(t)\right) &= \int_{-\infty}^{\infty} \int_{0}^{t-r_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n+1}\right) g_{n+1}\left(x\left|t-\sum_{j=1}^{n+1}\tau_{j}\right.\right) x d\tau_{n+1} \cdots d\tau_{1} dx \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n+1}\right) \mu\left(t-\sum_{j=1}^{n+1}\tau_{j}\right) d\tau_{n+1} \cdots d\tau_{1} \\ &= \mu\int_{0}^{t} f\left(\tau_{1}\right) \int_{0}^{t-\tau_{1}} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f\left(\tau_{2}\right) \cdots f\left(\tau_{n+1}\right) \left(t-\sum_{j=1}^{n+1}\tau_{j}\right) d\tau_{n+1} \cdots d\tau_{2} d\tau_{1} \\ &= \mu\int_{0}^{t} f\left(\tau_{1}\right) \left(\left(t-\tau_{1}-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1} \frac{(n-j)}{j!}(t-\tau_{1})^{j} \lambda^{j-1}\right) e^{-\lambda(t-\tau_{1})}\right) d\tau_{1} \\ &= \mu\left(\int_{0}^{t} f\left(\tau_{1}\right) \left(t-\tau_{1}-\frac{n}{\lambda}\right) d\tau_{1} + \int_{0}^{t} f\left(\tau_{1}\right) \left(\left(\sum_{j=0}^{n-1} \frac{(n-j)}{j!}(t-\tau_{1})^{j} \lambda^{j-1}\right) e^{-\lambda(t-\tau_{1})}\right) d\tau_{1}\right) \\ &= \mu\left(t\int_{0}^{t} \lambda e^{-\lambda \tau_{1}} d\tau_{1} - \int_{0}^{t} \lambda e^{-\lambda \tau_{1}} \left(\tau_{1}+\frac{n}{\lambda}\right) d\tau_{1} + \lambda e^{-\lambda t} \int_{0}^{n-1} \frac{(n-j)}{j!}(t-\tau_{1})^{j} \lambda^{j-1} d\tau_{1}\right) \\ &= \mu\left(\left(t-te^{-\lambda t}\right) - \left(\frac{n+1}{\lambda} - te^{-\lambda t} - \frac{n+1}{\lambda}e^{-\lambda t}\right) + \lambda e^{-\lambda t} \sum_{j=0}^{n-1} \frac{(n-j)}{j!} \lambda^{j-1} \frac{(t-\tau_{1})^{j+1}}{(j+1)}\right) \\ &= \mu\left(\left(t-\frac{n+1}{\lambda}\right) + \frac{n+1}{\lambda}e^{-\lambda t} + \lambda e^{-\lambda t} \sum_{j=0}^{n-1} \frac{(n-j)}{(j+1)!} \lambda^{j-1} t^{j+1}\right) \end{split}$$

In the last step of Eq. (F.35), let k = j + 1 we get

$$E\left(X_{n+1}^{*}\left(t\right)\right) = \mu_{n+1}\left(\left(t - \frac{n+1}{\lambda}\right) + \frac{n+1}{\lambda}e^{-\lambda t} + \lambda e^{-\lambda t}\sum_{k=1}^{n}\frac{\left(n-k+1\right)}{k!}\lambda^{k-2}t^{k}\right)$$
$$= \mu\left(\left(t - \frac{n+1}{\lambda}\right) + \lambda e^{-\lambda t}\frac{n-0+1}{0!}\lambda^{0-2}t^{0} + \lambda e^{-\lambda t}\sum_{k=1}^{n}\frac{\left(n-k+1\right)}{k!}\lambda^{k-2}t^{k}\right)$$
$$(F.36)$$
$$= \mu\left(\left(t - \frac{n+1}{\lambda}\right) + \lambda e^{-\lambda t}\sum_{k=0}^{n}\frac{\left(n-k+1\right)}{k!}\lambda^{k-2}t^{k}\right)$$

which is the same as the result obtained with Eq. (4.19) when i = n+1.

We prove Eq. (4.20) by induction. First, we show that it holds when i=1. Given  $\{\tau_j\}$  are known, the conditional variance of branch i is

$$Var\left(X_{i}\left(t-\sum_{j=1}^{i}\tau_{j}\right)\right)=\sigma^{2}\left(t-\sum_{j=1}^{i}\tau_{j}\right)$$

Then

$$E\left(X_i^2\left(t-\sum_{j=1}^i\tau_j\right)\right) = Var\left(X_i\left(t-\sum_{j=1}^i\tau_j\right)\right) + \left(E\left(X_i\left(t-\sum_{j=1}^i\tau_j\right)\right)\right)^2$$
  
$$= \sigma^2\left(t-\sum_{j=1}^i\tau_j\right) + \mu^2\left(t-\sum_{j=1}^i\tau_j\right)^2$$
(F.37)

When i=1, the variance of the degradation of the first branch can be obtained as follows:

$$\begin{aligned} \operatorname{Var}\left(X_{1}^{*}(t)\right) &= E\left(\left(X_{1}^{*}(t)\right)^{2}\right) - \left(E\left(X_{1}^{*}(t)\right)\right)^{2} \\ &= \int_{0}^{t} \lambda e^{-\lambda \tau_{1}} \left(\sigma^{2}\left(t-\tau_{1}\right) + \mu^{2}\left(t-\tau_{1}\right)^{2}\right) d\tau_{1} - \mu^{2} \left(t-\frac{1}{\lambda} + \frac{e^{-\lambda t}}{\lambda}\right)^{2} \\ &= \frac{1}{\lambda^{2}} \left(\lambda^{2} \mu^{2} t^{2} + \left(\lambda^{2} \sigma^{2} - 2\lambda \mu^{2}\right) t + \left(1 - e^{-\lambda t}\right) \left(2\mu^{2} - \lambda \sigma^{2}\right)\right) - \mu^{2} \left(t-\frac{1}{\lambda} + \frac{e^{-\lambda t}}{\lambda}\right)^{2} \end{aligned}$$
(F.38)  
$$&= \frac{\mu^{2}}{\lambda^{2}} \left(\lambda^{2} t^{2} - 2\lambda t + (1+1)1 - e^{-\lambda t} t^{0} \frac{(1-0+1)(1-0)}{0!} \lambda^{0}\right) \\ &+ \sigma^{2} \left(t-\frac{1}{\lambda} + \frac{(1-0)}{0!} t^{0} \lambda^{-1} e^{-\lambda t}\right) - \mu^{2} \left(t-\frac{1}{\lambda} + \frac{e^{-\lambda t}}{\lambda}\right)^{2} \end{aligned}$$

The variance of the first branch obtained with Eq. (F.38) is the same as that obtained with Eq. (4.20). The first step of the proof is done. In the second step, we prove that if Eq. (4.20) holds for the  $n^{th}$  branch, it also holds for the  $n+1^{th}$  branch. As

$$\begin{aligned} \operatorname{Var}\left(X_{n}^{*}(t)\right) &= E\left(\left(X_{n}^{*}(t)\right)^{2}\right) - \left(E\left(X_{n}^{*}(t)\right)\right)^{2} \\ &= \int_{0}^{t} \cdots \int_{0}^{t-\frac{p-1}{p-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) \left(\sigma^{2}\left(t-\sum_{j=1}^{n}\tau_{j}\right) + \mu^{2}\left(t-\sum_{j=1}^{n}\tau_{j}\right)^{2}\right) d\tau_{n} \cdots d\tau_{1} \\ &- \left(\mu\left(\left(t-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1}\left(\frac{n-j}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right)\right)^{2} \right) d\tau_{n} \cdots d\tau_{1} \\ &+ \int_{0}^{t} \cdots \int_{0}^{t-\frac{p-1}{p-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) \left(\mu^{2}\left(t-\sum_{j=1}^{n}\tau_{j}\right)^{2}\right) d\tau_{n} \cdots d\tau_{1} \\ &+ \int_{0}^{t} \cdots \int_{0}^{t-\frac{p-1}{p-1}\tau_{j}} f\left(\tau_{1}\right) \cdots f\left(\tau_{n}\right) \left(\sigma^{2}\left(t-\sum_{j=1}^{n}\tau_{j}\right)\right) d\tau_{n} \cdots d\tau_{1} \\ &- \left(\mu\left(\left(t-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1}\left(\frac{n-j}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right)\right)^{2} \\ &= \frac{\mu^{2}}{\lambda^{2}} \left(\lambda^{2}t^{2} - 2i\lambda t + (i+1)i - e^{-\lambda t}\sum_{j=0}^{t-1} t^{j} \left(\frac{(i-j+1)(i-j)}{j!}\lambda^{j}\right)\right) + \\ \sigma^{2} \left(\left(t-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1}\left(\frac{n-j}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right) - \left(\mu\left(\left(t-\frac{n}{\lambda}\right) + \left(\sum_{j=0}^{n-1}\left(\frac{n-j}{j!}t^{j}\lambda^{j-1}\right)e^{-\lambda t}\right)\right)^{2} \end{aligned}$$

From the last two steps of Eq. (F.39), we know that

$$\int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n-1} \tau_{j}} f(\tau_{1}) \cdots f(\tau_{n}) \left( t - \sum_{j=1}^{n} \tau_{j} \right)^{2} d\tau_{n} \cdots d\tau_{1}$$

$$= \frac{1}{\lambda^{2}} \left( \lambda^{2} t^{2} - 2i\lambda t + (i+1)i - e^{-\lambda t} \sum_{j=0}^{i-1} t^{j} \left( \frac{(i-j+1)(i-j)}{j!} \lambda^{j} \right) \right)$$
(F.40)

Let k = j+1,  $t = t - \tau_1$ , where  $\tau_1$  is a constant, we have

$$\int_{0}^{t-\tau_{1}} \cdots \int_{0}^{t-\tau_{1}-\sum_{k=2}^{n}\tau_{k}} f(\tau_{2})\cdots f(\tau_{n+1})\left(t-\tau_{1}-\sum_{j=2}^{n+1}\tau_{k}\right)^{2} d\tau_{n+1}\cdots d\tau_{2}$$

$$= \frac{1}{\lambda^{2}} \begin{pmatrix} \lambda^{2} (t-\tau_{1})^{2} - 2n\lambda(t-\tau_{1}) + (n+1)n \\ -e^{-\lambda(t-\tau_{1})} \sum_{j=0}^{n-1} (t-\tau_{1})^{j} \left(\frac{(n-j+1)(n-j)}{j!}\lambda^{j}\right) \end{pmatrix}$$
(F.41)

Then we obtain Eq. (F.42) as follows:

$$E\left(\left(X_{n+1}^{*}(t)\right)^{2}\right) = \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f(\tau_{1})\cdots f(\tau_{n+1})E\left(X_{n+1}^{2}\left(t-\sum_{j=1}^{n+1}\tau_{j}\right)\right)d\tau_{n+1}\cdots d\tau_{1}$$

$$= \mu\int_{0}^{t} f(\tau_{1})\int_{0}^{t-\tau_{1}} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f(\tau_{2})\cdots f(\tau_{n+1})\left(t-\sum_{j=1}^{n+1}\tau_{j}\right)d\tau_{n+1}\cdots d\tau_{2}d\tau_{1}$$

$$= \int_{0}^{t} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f(\tau_{1})\cdots f(\tau_{n+1})\left(\sigma^{2}\left(t-\sum_{j=1}^{n+1}\tau_{j}\right)+\mu^{2}\left(t-\sum_{j=1}^{n+1}\tau_{j}\right)^{2}\right)d\tau_{n+1}\cdots d\tau_{1}$$

$$= \sigma^{2}\int_{0}^{t} f(\tau_{1})\int_{0}^{t-\tau_{1}} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f(\tau_{2})\cdots f(\tau_{n+1})\left(t-\sum_{j=1}^{n+1}\tau_{j}\right)d\tau_{n+1}\cdots d\tau_{2}d\tau_{1}$$

$$+ \mu^{2}\int_{0}^{t} f(\tau_{1})\left(\int_{0}^{t-\tau_{1}} \cdots \int_{0}^{t-\sum_{j=1}^{n}\tau_{j}} f(\tau_{2})\cdots f(\tau_{n+1})\left(t-\sum_{j=1}^{n}\tau_{j}\right)^{2}d\tau_{n+1}\cdots d\tau_{2}d\tau_{1}$$

$$= \sigma^{2}\int_{0}^{t} f(\tau_{1})\left(\left(t-\tau_{1}-\frac{n}{\lambda}\right)+\left(\sum_{j=0}^{n-1}\frac{(n-j)}{j!}(t-\tau_{1})^{j}\lambda^{j-1}\right)e^{-\lambda(t-\tau_{1})}\right)d\tau_{1}$$

$$+ \frac{\mu^{2}}{\lambda^{2}}\int_{0}^{t} f(\tau_{1})\left(\left(\lambda^{2}(t-\tau_{1})^{2}-2n\lambda(t-\tau_{1})+(n+1)n\right)-e^{-\lambda(t-\tau_{1})}\sum_{j=0}^{n-1}(t-\tau_{1})^{j}\left(\frac{(n-j+1)(n-j)}{j!}\lambda^{j}\right)\right)d\tau_{1}$$
(F.42)

where *B* can be decomposed into  $B_1$  and  $B_2$  as follows:

$$B = \int_{0}^{t} f(\tau_{1}) \left(\lambda^{2} (t - \tau_{1})^{2} - 2n\lambda(t - \tau_{1}) + (n + 1)n\right) d\tau_{1}$$
  
- 
$$\int_{0}^{t} f(\tau_{1}) \left(e^{-\lambda(t - \tau_{1})} \sum_{j=0}^{n-1} (t - \tau_{1})^{j} \left(\frac{(n - j + 1)(n - j)}{j!}\lambda^{j}\right)\right) d\tau_{1}$$
  
- 
$$B_{2}$$
  
(F.43)

where

$$B_{1} = \int_{0}^{t} f(\tau_{1}) \Big(\lambda^{2} (t - \tau_{1})^{2} - 2n\lambda(t - \tau_{1}) + (n + 1)n\Big) d\tau_{1}$$
  

$$= \int_{0}^{t} \lambda e^{-\lambda\tau_{1}} \lambda^{2} (t - \tau_{1})^{2} d\tau_{1} - \int_{0}^{t} \lambda e^{-\lambda\tau_{1}} 2n\lambda(t - \tau_{1}) d\tau_{1} + \int_{0}^{t} \lambda e^{-\lambda\tau_{1}} (n + 1)n d\tau_{1} \qquad (F.44)$$
  

$$= \Big(\lambda^{2} t^{2} - 2\lambda t + 2 - 2e^{-\lambda t}\Big) - \Big(2ne^{-\lambda t} + 2n\lambda t - 2n\Big) + \Big(n(n + 1) - n(n + 1)e^{-\lambda t}\Big)$$
  

$$= \lambda^{2} t^{2} - 2(n + 1)\lambda t + (n + 1)(n + 2) - e^{-\lambda t} (n + 1)(n + 2)$$

$$B_{2} = \int_{0}^{t} f\left(\tau_{1}\right) \left(e^{-\lambda(t-\tau_{1})} \sum_{j=0}^{n-1} \left(t-\tau_{1}\right)^{j} \frac{\left(n-j+1\right)(n-j)}{j!} \lambda^{j}\right) d\tau_{1}$$

$$= \lambda e^{-\lambda t} \int_{0}^{t} \left(\sum_{j=0}^{n-1} \left(t-\tau_{1}\right)^{j} \frac{\left(n-j+1\right)(n-j)}{j!} \lambda^{j}\right) d\tau_{1}$$

$$= e^{-\lambda t} \sum_{j=0}^{n-1} \left(\frac{-\left(t-\tau_{1}\right)^{j+1}}{j+1}\right) \int_{0}^{t} \frac{\left(n-j+1\right)(n-j)}{j!} \lambda^{j+1}$$

$$= e^{-\lambda t} \sum_{j=0}^{n-1} \frac{t^{j+1}}{j+1} \frac{\left(n-j+1\right)(n-j)}{j!} \lambda^{j+1}$$

$$= e^{-\lambda t} \sum_{j=0}^{n-1} \frac{\left(n-j+1\right)(n-j)}{\left(j+1\right)!} t^{j+1} \lambda^{j+1}$$
(F.45)

Let k = j+1, Eq. (F.45) changes to

$$B_2 = e^{-\lambda t} \sum_{k=1}^{n} \frac{(n-k+2)(n-k+1)}{k!} t^k \lambda^k$$
(F.46)

Similarly,

$$A = \int_{0}^{t} f(\tau_{1}) \left( t - \tau_{1} - \frac{n}{\lambda} \right) d\tau_{1} + \int_{0}^{t} f(\tau_{1}) \left( \left( \sum_{j=0}^{n-1} \frac{(n-j)}{j!} (t - \tau_{1})^{j} \lambda^{j-1} \right) e^{-\lambda(t-\tau_{1})} \right) d\tau_{1}$$
  
$$= \mu \left( \left( t - \frac{n+1}{\lambda} \right) + \frac{n+1}{\lambda} e^{-\lambda t} + \lambda e^{-\lambda t} \sum_{j=0}^{n-1} \frac{(n-j)}{(j+1)!} \lambda^{j-1} t^{j+1} \right)$$
 (F.47)

Then we get

$$\begin{aligned} \operatorname{Var}\left(X_{n+1}^{*}(t)\right) &= E\left(\left(X_{n+1}^{*}(t)\right)^{2}\right) - E^{2}\left(X_{n+1}^{*}(t)\right) \\ &= \sigma^{2}\left(\left(t - \frac{n+1}{\lambda_{b}}\right) + \lambda e^{-\lambda t} \sum_{k=0}^{n} \frac{(n-k+1)}{k!} \lambda^{k-2} t^{k}\right) \\ &+ \frac{\mu^{2}}{\lambda^{2}} \left( \frac{\lambda^{2} t^{2} - 2(n+1)\lambda t + (n+1)(n+2)}{-e^{-\lambda t} \left(n-0+2\right)(n-0+1)} t^{0} \lambda^{0} - e^{-\lambda t} \sum_{k=1}^{n} \frac{(n-k+2)(n-k+1)}{k!} t^{k} \lambda^{k}\right) \\ &- \left(\mu\left(\left(t - \frac{n+1}{\lambda}\right) + \left(\sum_{j=0}^{n} \frac{(n+1-j)}{j!} t^{j} \lambda^{j-1}\right) e^{-\lambda t}\right)\right)^{2} \\ &= \sigma^{2} \left( \frac{\left(t - \frac{n+1}{\lambda}\right)}{\lambda e^{-\lambda t}} \sum_{k=0}^{n} \frac{(n-k+1)}{k!} \lambda^{k-2} t^{k} \right) \\ &+ \frac{\mu^{2}}{\lambda^{2}} \left( \frac{\lambda^{2} t^{2} - 2(n+1)\lambda t + (n+1)(n+2)}{-e^{-\lambda t} \sum_{k=0}^{n} \frac{(n-k+2)(n-k+1)}{k!} t^{k} \lambda^{k} \right) \\ &- \left( \mu\left(\left(t - \frac{n+1}{\lambda}\right) + \left(\sum_{j=0}^{n} \frac{(n+1-j)}{j!} t^{j} \lambda^{j-1}\right) e^{-\lambda t} \right) \right)^{2} \end{aligned}$$
(F.48)

Which is the same as the result obtained with Eq. (4.20) when i = n+1.

Q.E.D.

#### APPENDIX G MODIFIED K NEAREST NEIGHBOUR IMPUTATION

In this section, we describe a non-parametric data interpolation approach modified from the *k* nearest neighbor imputation (*k*NNI) model (Batista *et al.*, 2003). Suppose that the degradation dataset that includes observations from all the *N* sensors is complete until time  $t_{\tau}$ . The first missing data is noticed at time  $t_{\tau+1}$ . Suppose the first missing data among the *N* sensors (by sensor index) appears from sensor *i*. The interpolation procedures are as follows:

- In each time instant t<sub>κ</sub>, 1≤κ≤τ , assign the N observations {d<sub>u</sub>(t<sub>κ</sub>)}, 1≤u≤N into k clusters using k-means clustering algorithm. At each time instant, we index the clusters by the rank of its centroid from 1 to k. For example, cluster 1 corresponds to the cluster that has the smallest centroid while cluster k corresponds to the cluster that has the largest centroid. For each sensor u, where 1≤u≤N, we have a cluster membership sequence as {C<sub>u</sub>(t<sub>κ</sub>)}, 1≤κ≤τ. C<sub>u</sub>(t<sub>κ</sub>) takes one of the values from {1,...,k}.
- 2) For each  $1 \le \kappa \le \tau$ , compare  $\{C_u(t_\kappa)\}, 1 \le u \le N, u \ne i$  with  $C_i(t_\kappa)$ , obtain the indicator matrix

$$M_{u}(t_{\kappa}) = \begin{cases} 0 \text{ if } C_{u}(t_{\kappa}) \neq C_{i}(t_{\kappa}) \\ 1 \text{ if } C_{u}(t_{\kappa}) = C_{i}(t_{\kappa}) \end{cases}.$$

3) Calculate the similarity score of each sensor  $S_u(t_{\tau+1})$ ,  $1 \le u \le N$  at time  $t_{\tau+1}$  with the following formula

$$S_{u}(t_{\tau+1}) = \sum_{\kappa=1}^{t_{\tau}} e(t_{\kappa}) M_{u}(t_{\kappa})$$

where  $\{e(t_{\kappa})\}$ ,  $1 \le \kappa \le \tau$  are the weights of the similarity scores that follow a geometric series as

$$e(t_{\kappa}) = e(t_1)q^{\kappa-1}$$

where q satisfies

$$e(t_1)q^{\kappa-1}-e(t_1)-q+1=0$$

The weights place higher importance on the recent cluster membership as the missing data is considered to be more closely related to recent measurements.

- 4) Take the mean of the similarity scores of the sensors in each of the k clusters at time  $t_{\tau+1}$ . Take the centroid of the cluster that has the largest mean similarity score at time  $t_{\tau+1}$  as the interpolated data, which is denoted as  $d_i^*(t_{\tau+1})$ .
- 5) Treat the interpolated data  $d_i^*(t_{\tau+1})$  as observed value. Interpolate other missing values with the same procedures until there is no more missing data.

#### APPENDIX H IMPUTATION WITH BOOTSTRAP

In this section, we describe a set of degradation data interpolation approaches modified from the previous work of the authors based on bootstrap (Guo *et al.*, 2018). Bootstrap has been used primarily to obtain confidence intervals of statistical parameters. The general approach is to assume that the degradation increments are independent and identically distributed (*i.i.d.*), and approximate the distribution F of  $\Delta d(t)$  by the empirical distribution  $F^*$  of the observed  $\Delta d(t)$ . If we observe a sufficiently large sample,  $F^*$ is then a reasonable approximation of F. Three scenarios are investigated and the corresponding missing degradation data interpolation models are proposed accordingly.

## H.1 Degradation Increments are i.i.d.

If the degradation increments are approximately *i.i.d.*, we propose to interpolate the missing degradation increments by bootstrapping the observed degradation increments. Suppose sensor i has missing data. The specific interpolation procedures are as follows:

 Suppose we have an observed degradation dataset for sensor *i* as {d<sub>i</sub>(t<sub>j</sub>)}, t<sub>j</sub> ∈ A, where A is the set of time indexes of non-missing observations.
 For simplicity, we define {Δd<sub>i</sub>(t<sub>j</sub>) = d<sub>i</sub>(t<sub>j+1</sub>) - d<sub>i</sub>(t<sub>j</sub>)}, t<sub>j</sub> ∈ A as the set of the corresponding degradation increments. Δd<sub>i</sub>(t<sub>i</sub>) is the increment corresponding to the starting degradation  $d_i(t_j)$ . Let  $\Delta \mathbf{d}_i = \{\Delta d_i(t)\}, t \in A$ . Because the missing data starts at time  $t_{\tau+1}$ , set j = 1.

- 2) Randomly draw a sample data  $\Delta d_i^*(t_{\tau-1+j})$  from set  $\Delta \mathbf{d}$ , the interpolated missing data at time  $t_{\tau+j}$  is  $d_i^*(t_{\tau+j}) = \Delta d_i^*(t_{\tau-1+j}) + d_i^*(t_{\tau-1+j})$ .
- 3) Treat the interpolated data at observed data. Let j = j+1, repeat (2) until all missing data have been interpolated.
- 4) Follow the same procedures, interpolate other missing data.

## H.2 Residuals are Approximately i.i.d.

The method described in section 1 assumes that the degradation increments  $\{d_i(t_j)\}, t_j \in A$  are *i.i.d.* However,  $\{d_i(t_j)\}, t_j \in A$  may not always be *i.i.d.*, because the mean or variance of  $\{d_i(t_j)\}, t_j \in A$  may change with time. In this section we are interested in the case where future degradation increments are dependent on current degradation states. The basic approach is that  $\{\Delta d_i(t_j)\}, t_j \in A$  are first transformed into an *i.i.d.* sequence, which can then be resampled and inversely transformed to produce the new  $\Delta d_i^*(t_j)$  sequence. More specifically, the increments are first fitted with a linear function of the degradation states. If the residuals *i.i.d.*, they are then resampled and the interpolated increments are obtained accordingly. The procedures are as follows:

- 1) Suppose we have an observed degradation dataset of sensor *i* as  $\{d_i(t_j)\}, t_j \in A$ , where *A* is the set of the time indexes of the non-missing observations.  $\{\Delta d_i(t_j) = d_i(t_{j+1}) d_i(t_j)\}, t_j \in A$  is the set of the corresponding degradation increments. Use the degradation increments to fit the linear model:  $\Delta d_i(t_j) = \beta_0 + \beta_1 d_i(t_j) + \varepsilon_i(t_j), t_j \in A$  and obtain the estimate of the parameters  $\hat{\beta} = \{\hat{\beta}_0, \hat{\beta}_1\}$ . Check if the residuals  $\{\varepsilon_i(t_j)\}, t_j \in A$  are *i.i.d*. If the residuals  $\{\varepsilon_i(t_j)\}, t_j \in A$  are identically distributed and independent of  $\{d_i(t_j)\}, t_j \in A$ , let  $\varepsilon_i = \{\varepsilon_i(t_j)\}, t_j \in A$ . Because the missing data starts at time  $t_{r+1}$ , let j = 1.
- 2) Randomly draw a sample data  $\varepsilon_i^*(t_{\tau-1+j})$  from the set  $\varepsilon_i$ , the corresponding interpolated missing data increment in  $(t_{\tau-1+j}, t_{\tau+j})$  is  $\Delta d_i^*(t_{\tau-1+j}) = \hat{\beta}_0 + \hat{\beta}_1 d_i^*(t_{\tau-1+j}) + \varepsilon_i^*(t_{\tau-1+j}).$
- 3) The interpolated missing data at time  $t_{\tau+j}$  is  $d_i^*(t_{\tau+j}) = \Delta d_i^*(t_{\tau-1+j}) + d_i^*(t_{\tau-1+j})$ .
- 4) Treat interpolated data at observed data. Let j = j+1, repeat (2) (3) until all the missing data have been interpolated

#### H.3 Residuals are Non-i.i.d.

In many cases, the residuals are not *i.i.d.*, e.g., they are dependent on  $\{d_i(t_j)\}, t_j \in A$ . Transformations of the residuals to *i.i.d.* values are required such that the bootstrap method can be effectively used. Common transformations include taking logarithms of the residuals, squares of the residuals and roots of the residuals. Without loss of generality, we use square root transformation as an example to illustrate the procedures of interpolating the missing observations.

1) Suppose we have an observed degradation dataset  $\{d_i(t_j)\}, t_j \in A$ , where A is indexes of the set of the time non-missing observations.  $\left\{\Delta d(t_j) = d_i(t_{j+1}) - d_i(t_j)\right\}, t_j \in A$  is the set of the corresponding degradation increments. Use the degradation increments to fit the linear model:  $\Delta d_i(t_i) = \beta_0 + \beta_1 d_i(t_j) + \varepsilon_i(t_j), t_j \in A$  and obtain the estimates of the parameters  $\hat{\boldsymbol{\beta}} = \{\hat{\beta}_0, \hat{\beta}_1\}$ . When the residuals  $\boldsymbol{\varepsilon}_i = \{\varepsilon_i(t_j)\}, t_j \in A$  are dependent on  $\{d_i(t_j)\}, t_j \in A$ , the absolute values of the residuals as well as the preceding degradation states are transformed by taking the square roots. The transformed values are used fit the linear model to  $\sqrt{abs(\varepsilon_i(t_j))} = c_0 + c_1\sqrt{d_i(t_j)} + \zeta_i(t_j), t_j \in A$  and estimate the parameters  $\hat{\mathbf{c}} = \{\hat{c}_0, \hat{c}_1\}$ , where  $E\left(\sqrt{abs\left(\varepsilon_i\left(t_j\right)\right)}\right) = c_0 + c_1\sqrt{d_i\left(t_j\right)}, t_j \in A$ . Let  $\left\{\frac{1}{E\left(\sqrt{abs\left(\varepsilon_{i}\left(t_{i}\right)\right)}\right)}\right\}, t_{j} \in A \text{ be the set of weights, update } \hat{\boldsymbol{\beta}} = \{\hat{\beta}_{0}, \hat{\beta}_{1}\} \text{ with}$ weighted least square model. Now

$$\omega_{i}(t_{j}) = \frac{\sqrt{abs(\varepsilon_{i}(t_{j}))}}{E(\sqrt{abs(\varepsilon_{i}(t_{j}))})} = \frac{abs(\varepsilon_{i}(t_{j}))}{\hat{c}_{0} + \hat{c}_{1}\sqrt{d_{i}(t_{j})}}, t_{j} \in A \quad \text{are approximately } i.i.d. \text{ The}$$

normalized residuals are denoted as a dataset vector  $\boldsymbol{\omega}_i = \{\omega_i(t_j)\}, t_j \in A$ . The missing data starts at time  $t_{r+1}$ . Set j = 1.

2) Randomly draw a sample data  $\omega_i^*(t_{\tau-1+j})$  from set  $\omega_i$ , the corresponding interpolated missing data increment in  $(t_{\tau-1+j}, t_{\tau+j})$  is  $\Delta d_i^*(t_{\tau-1+j}) = \hat{\beta}_0 + \hat{\beta}_1 d_i (t_{\tau-1+j}) + s \left( \omega_{t-1+j}^*(\hat{c}_0 + \hat{c}_1 \sqrt{d_i (t_{\tau-1+j})}) \right)^2$ , where *s* is a

number randomly generated from the set  $\{-1, 1\}$ .

3) The interpolated missing data at time  $t_{\tau+j}$  is  $d_i^*(t_{\tau+i}) = \Delta d_i^*(t_{\tau-1+j}) + d_i^*(t_{\tau-1+j})$ .

Treat the interpolated data as observed data. Let j = j+1, repeat (2) (3) until all the missing data have been interpolated

#### REFERENCES

- ACHTERBOSCH, G. G. and GRZELAK, L. A. 2006. Determination of the corrosion rate of a MIC influenced pipeline using four consecutive pig runs. 2006 International Pipeline Conference. American Society of Mechanical Engineers, 209-217.
- AL-AMIN, M., KARIYAWASAM, S., ZHANG, S. and ZHOU, W. 2016. Non-Linear Corrosion Growth: A More Appropriate and Accurate Model for Predicting Corrosion Growth Rate. 2016 11th International Pipeline Conference. American Society of Mechanical Engineers, V001T03A021-V001T03A021.
- ALLISON, P. D. 1987. Estimation of linear models with incomplete data. *Sociological methodology*, 71-103.
- ALYANNEZHADI, M. M., POUYAN, A. A. and ABOLGHASEMI, V. 2016. An efficient algorithm for multisensory data fusion under uncertainty condition. *Journal of Electrical Systems and Information Technology*.
- AMIRAT, A., MOHAMED-CHATEAUNEUF, A. and CHAOUI, K. 2006. Reliability assessment of underground pipelines under the combined effect of active corrosion and residual stress. *International Journal of Pressure Vessels and Piping*, 83, 107-117.
- AMOLINS, K., ZHANG, Y. and DARE, P. 2007. Wavelet based image fusion techniques—An introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62, 249-263.
- ASSOUL, Y., BACHA, N. and SEMNAR, D. 2008. Probabilistic model for pitting corrosion and fatigue life estimation of turbine blades. *Structural Integrity and Life*, 8, 3-12.
- AYDILEK, I. B. and ARSLAN, A. J. I. S. 2013. A hybrid method for imputation of missing values using optimized fuzzy c-means with support vector regression and a genetic algorithm. 233, 25-35.
- BALAKRISHNAN, S., CACCIOLA, M., UDPA, L., RAO, B. P., JAYAKUMAR, T. and RAJ, B. 2012. Development of image fusion methodology using discrete wavelet transform for eddy current images. *NDT & E International*, 51, 51-57.
- BATISTA, G. E. and MONARD, M. C. 2003. An analysis of four missing data treatment methods for supervised learning. *Applied artificial intelligence*, 17, 519-533.
- BENFERHAT, S. and SOSSAI, C. 2006. Reasoning with multiple-source information in a possibilistic logic framework. *Information Fusion*, 7, 80-96.
- BERESTYCKI, J., BRUNET, É., HARRIS, J. and HARRIS, S. C. 2010. The almost-sure population growth rate in branching Brownian motion with a quadratic breeding potential. *Statistics probability letters*, 80, 1442-1446.
- BERESTYCKI, J., BRUNET, É., HARRIS, J. W., HARRIS, S. C. and ROBERTS, M. I. 2015. Growth rates of the population in a branching Brownian motion with an

inhomogeneous breeding potential. *Stochastic Processes and their Applications*, 125, 2096-2145.

- BOGDANOFF, J. L. and KOZIN, F. 1985. *Probabilistic models of cumulative damage*, John Wiley & Sons.
- BROWN, P. E., ROBERTS, G. O., KÅRESEN, K. F. and TONELLATO, S. 2000. Blur generated non - separable space - time models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62, 847-860.
- BROWN, R. L. 1994. Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of five methods. *Structural Equation Modeling: A Multidisciplinary Journal*, 1, 287-316.
- BUCK, S. F. 1960. A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society. Series B (Methodological)*, 302-306.
- BULLEN, F. P., HENDERSON, F. and WAIN, H. L. 1970. Crack-branching in heavily drawn chromium. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 21, 689-699.
- BURSTEIN, G., LIU, C., SOUTO, R. and VINES, S. 2004. Origins of pitting corrosion. *Corrosion Engineering, Science and Technology*, 39, 25-30.
- CALDER, C. A. 2007. Dynamic factor process convolution models for multivariate spacetime data with application to air quality assessment. *Environmental and Ecological Statistics*, 14, 229-247.
- CALEYO, F., VELÁZQUEZ, J., VALOR, A. and HALLEN, J. 2009. Markov chain modelling of pitting corrosion in underground pipelines. *Corrosion Science*, 51, 2197-2207.
- CARLIN, B. P., GELFAND, A. E. and BANERJEE, S. 2014. *Hierarchical modeling and analysis for spatial data*, Chapman and Hall/CRC.
- CAWLEY, N. R. and HARLOW, D. G. 1996. Spatial statistics of particles and corrosion pits in 2024-T3 aluminium alloy. *Journal of Materials Science*, 31, 5127-5134.
- CHENG, T. and PANDEY, M. 2012. An accurate analysis of maintenance cost of structures experiencing stochastic degradation. *Structure and Infrastructure Engineering*, 8, 329-339.
- CHOOKAH, M. 2009. Structuring a probabilistic model for reliability evaluation of piping subject to corrosion-fatigue degradation. University of Maryland, College Park.
- CHRISTIE, D. and KOLSKY, H. 1952. The fractures produced in glass and plastics by the passage of stress waves. *Journal of the Society of Glass Technology*, 36, 65-73.
- CLARK, A. and IRWIN, G. 1966. Crack-propagation behaviors. *Experimental Mechanics*, 6, 321-330.

- CONGLETON, J. and DVORAK, G. 1973. Practical applications of crack-branching measurements. *Proceedings of an international conference on Dynamic Crack Propagation*. Springer, 427-438.
- CRESSIE, N. and HUANG, H.-C. 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94, 1330-1339.
- CULLIN, M., HARLOW, D. and WEI, R. 2011. Modeling and simulation of constituent particle clustering and distribution in rolled aluminum alloys. *Modelling and Simulation in Materials Science and Engineering*, 19, 035004.
- DAVID, M., LITTLE, R. J., SAMUHEL, M. E. and TRIEST, R. K. 1986. Alternative methods for CPS income imputation. *Journal of the American Statistical Association*, 81, 29-41.
- DAVIS, J. R. 2000. Corrosion: Understanding the basics, ASM International.
- DE LA FUENTE, D., CASTAÑO, J. G. and MORCILLO, M. 2007. Long-term atmospheric corrosion of zinc. *Corrosion Science*, 49, 1420-1436.
- DE LA FUENTE, D., DÍAZ, I., SIMANCAS, J., CHICO, B. and MORCILLO, M. 2011. Long-term atmospheric corrosion of mild steel. *Corrosion Science*, 53, 604-617.
- DE LEON ESCOBEDO, D. and FLORES MACÍAS, O. 2005. Effect of spatial correlation on the failure probability of pipelines under corrosion.
- DEMPSEY, J. and BURGERS, P. 1985. Dynamic crack branching in brittle solids. International journal of fracture, 27, 203-213.
- DEMPSTER, A. P. 1968. A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30, 205-247.
- DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society*. *Series B (methodological)*, 1-38.
- DESIGNATION, A. 2009. Standard test methods for pitting and crevice corrosion resistance of stainless steels and related alloys by use of ferric fhloride solution. *ASTM International*.
- DESTERCKE, S., DUBOIS, D. and CHOJNACKI, E. 2009. Possibilistic information fusion using maximal coherent subsets. *IEEE Transactions on Fuzzy Systems*, 17, 79-92.
- DICESARE, G. 2006. Imputation, Estimation and Missing Data in Finance. UWSpace.
- DIN, M. M., NGADI, M. A. and NOOR, N. M. 2009. Improving Inspection Data Quality in Pipeline Corrosion Assessment. *Int C Comput Eng App*.
- DOMINÉ, M. and PIEPER, V. 1993. First passage time distribution of a two-dimensional Wiener process with drift. *Probability in the Engineering and Informational Sciences*, 7, 545-555.
- DONDETI, S. and TIPPUR, H. 2019. A Comparative Study of Dynamic Fracture of Soda-Lime Glass Using Photoelasticity, Digital Image Correlation and Digital Gradient Sensing Techniques. *Experimental Mechanics*, 1-17.
- DOROGOV, V. and CHISTYAKOV, V. 1988. Probabilistic models of the transformation of particles. Nauka, Moscow.
- EBARA, R. 2007. Corrosion fatigue crack initiation in 12% chromium stainless steel. *Materials Science and Engineering: A*, 468–470, 109-113.
- EFRON, B. 1994. Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, 89, 463-475.
- ELSAYED, E. and LIAO, H. 2004. A geometric Brownian motion model for field degradation data. *International Journal of Materials and Product Technology*, 20, 51-72.
- ELSAYED, E. A. 2012. Reliability engineering, John Wiley & Sons.
- ERNST, L. 1980. Variance of the estimated mean for several imputation procedures. the Proceedings of the Survey Research Methods Section, Alexandria, VA: American Statistical Association. 716-720.
- EYRING, H. 1935. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3, 107-115.
- FAN, Y., SUN, F. and JIANG, T. 2014. An improved ri for missing degradation data based on Brownian motion. *Reliability and Maintainability Symposium (RAMS)*, 2014 Annual. IEEE, 1-4.
- FAN, Z.-Y., JU, H. and SUN, F.-B. 2015. Improved Gamma process for degradation analysis under nonlinear condition. *International Journal of Reliability, Quality and Safety Engineering*, 22, 1550030.
- FANG, M. and ZEITOUNI, O. 2012. Slowdown for time inhomogeneous branching Brownian motion. *Journal of Statistical Physics*, 149, 1-9.
- FRANKEL, G. 1998. Pitting corrosion of metals a review of the critical factors. *Journal of the Electrochemical Society*, 145, 2186-2198.
- FREUDENTHAL, A. M. 1968. Statistical approach to brittle fracture. *Fracture, An* Advanced Treatise, II, 591-619.
- FUENTES, M., CHEN, L., DAVIS, J. M. and LACKMANN, G. 2003. A new class of nonseparable and nonstationary covariance models for wind fields. NORTH CAROLINA STATE UNIV AT RALEIGH DEPT OF STATISTICS.
- GEBRAEEL, N. 2006. Sensory-updated residual life distributions for components with exponential degradation patterns. *IEEE Transactions on Automation Science and Engineering*, 3, 382-393.
- GEBRAEEL, N. Z., LAWLEY, M. A., LI, R. and RYAN, J. K. 2005. Residual-life distributions from component degradation signals: A Bayesian approach. *IiE Transactions*, 37, 543-557.

- GHAHARI, S. M., DAVENPORT, A. J., RAYMENT, T., SUTER, T., TINNES, J.-P., PADOVANI, C., HAMMONS, J. A., STAMPANONI, M., MARONE, F. and MOKSO, R. J. C. S. 2011. In situ synchrotron X-ray micro-tomography study of pitting corrosion in stainless steel. 53, 2684-2687.
- GHOSH, S., DAVENPORT, A., CONNOLLY, B. J., VIGNAL, V. and OLTRA, R. 2007. Effect of Stress on Initiation and Propagation of Localized Corrosion in Aluminium Alloys. *Meeting Abstracts*. The Electrochemical Society, 1420-1420.
- GIORGIO, M., GUIDA, M. and PULCINI, G. 2011. An age-and state-dependent Markov model for degradation processes. *IIE Transactions*, 43, 621-632.
- GNEITING, T. 2002. Nonseparable, Stationary Covariance Functions for Space-Time Data. *Journal of the American Statistical Association*, 97, 590-600.
- GOLDSTEIN, E. M. 1960. The Corrosion and Oxidation of Metals: Scientific Principles and Practical Applications (Evans, Ulick R.). *Journal of Chemical Education*, 37, 662.
- GRAHAM, J. W. 2009. Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, 60, 549-576.
- GRAHAM, J. W. and DONALDSON, S. I. 1993. Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of follow-up data. *Journal of Applied Psychology*, 78, 119.
- GRAHAM, J. W., HOFER, S. M. and MACKINNON, D. P. 1996. Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, 31, 197-218.
- GROS, X. E., LIU, Z., TSUKADA, K. and HANASAKI, K. 2000. Experimenting with pixel-level NDT data fusion techniques. *IEEE Transactions on Instrumentation and Measurement*, 49, 1083-1090.
- GUEDES SOARES, C., GARBATOV, Y. and ZAYED, A. 2011. Effect of environmental factors on steel plate corrosion under marine immersion conditions. *Corrosion Engineering, Science and Technology*, 46, 524-541.
- GUO, J., WANG, C., CABRERA, J. and ELSAYED, E. A. 2018. Improved inverse Gaussian process and bootstrap: Degradation and reliability metrics. *Reliability Engineering & System Safety*, 178, 269-277.
- HAN, L. X. and SURESH, S. 1989. High Temperature Failure of an Alumina Silicon Carbide Composite under Cyclic Loads: Mechanisms of Fatigue Crack - Tip Damage. *Journal of the American Ceramic Society*, 72, 1233-1238.
- HARLOW, D. and WEI, R. 1995. Probability modeling for the growth of corrosion pits. ASME AD-Vol. 47, Structural integrity in aging aircraft. *Proceedings of the ASME Winter Annual Meeting*.
- HARLOW, D. G. 2012. Constituent particle clustering and pitting corrosion. 43, 2832-2838.

- HARLOW, D. G. and WEI, R. P. 1998. A probability model for the growth of corrosion pits in aluminum alloys induced by constituent particles. *Engineering Fracture Mechanics*, 59, 305-325.
- HARRIS, J. W. and HARRIS, S. C. 2009. Branching Brownian motion with an inhomogeneous breeding potential. Annales de l'IHP Probabilités et statistiques. 793-801.
- HAYNIE, F. and UPHAM, J. B. 1971. Effects of atmospheric pollutants on corrosion behavior of steels. *Materials Protection*.
- HOEPPNER, D. 1979. Model for prediction of fatigue lives based upon a pitting corrosion fatigue process. *Fatigue Mechanisms*. ASTM International.
- HONG, H. 1999a. Application of the stochastic process to pitting corrosion. *Corrosion*, 55, 10-16.
- HONG, H. P. 1999b. Inspection and maintenance planning of pipeline under external corrosion considering generation of new defects. *Structural safety*, 21, 203-222.
- HORNER, D. A., CONNOLLY, B. J., ZHOU, S., CROCKER, L. and TURNBULL, A. 2011. Novel images of the evolution of stress corrosion cracks from corrosion pits. *Corrosion Science*, 53, 3466-3485.
- HU, P., MENG, Q., HU, W., SHEN, F., ZHAN, Z. and SUN, L. 2016. A continuum damage mechanics approach coupled with an improved pit evolution model for the corrosion fatigue of aluminum alloy. *Corrosion Science*, 113, 78-90.
- HUA, D., AL-KHALIFA, K. N., HAMOUDA, A. S. and ELSAYED, E. A. 2013. Multisensor degradation data analysis. *Chemical Engineering Transactions*.
- HUANG, G., WANG, S., XIAO, F. and SUN, Y. 2009. A data fusion scheme for building automation systems of building central chilling plants. *Automation in Construction*, 18, 302-309.
- HUANG, X.-G. and XU, J.-Q. 2013. 3D analysis for pit evolution and pit-to-crack transition during corrosion fatigue. *Journal of Zhejiang University SCIENCE A*, 14, 292-299.
- HUBBELL, M., PRICE, C. and HEIDERSBACH, R. 1985. Crevice and pitting corrosion tests for stainless steels: a comparison of short-term tests with longer exposures. *ASTM Special Technical Publication*. 324-336.
- HUGHES, A. E., MOL, J. M., ZHELUDKEVICH, M. L. and BUCHHEIT, R. G. 2016. Active Protective Coatings.
- ISO, E. 1992. 9225. Corrosion of metals and slloys: Corrosivity of atmospheres: measurement of pollution. *ISO, Geneva, Switzerland*.
- JIANAN, X. and KAI, Y. 1997. Upper and lower bounds of stress-strength interference reliability with random strength-degradation. *IEEE Transactions on Reliability*, 46, 142-145.

- KAGAN, Y. 1982. Stochastic model of earthquake fault geometry. *Geophysical Journal International*, 71, 659-691.
- KAGAN, Y. Y. 1994. Observational evidence for earthquakes as a nonlinear dynamic process. *Physica D: Nonlinear Phenomena*, 77, 160-192.
- KALLEN, M.-J. and VAN NOORTWIJK, J. M. 2005. Optimal maintenance decisions under imperfect inspection. *Reliability Engineering & System Safety*, 90, 177-185.
- KALTHOFF, J. 1971. On the characteristic angle for crack branching in brittle materials. *International Journal of Fracture Mechanics*, 7, 478-480.
- KARIYAWASAM, S. and WANG, H. 2012. Useful trends for predicting corrosion growth. 2012 9th International Pipeline Conference. American Society of Mechanical Engineers, 245-256.
- KAWAI, S. and KASAI, K. 1985. Considerations of allowable stress of corrosion fatigue (focused on the influence of pitting). *Fatigue & Fracture of Engineering Materials* & Structures, 8, 115-127.
- KERR, S. and ROSENBERG, K. 1958. An index of cavitation erosion by means of radioisotopes. *Trans. Am. Soc. Mech. Engrs.*, 80.
- KING, G., HONAKER, J., JOSEPH, A. and SCHEVE, K. 2001. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American political science review*, 95, 49-69.
- KIOUMARSI, M. M., HENDRIKS, M. A., KOHLER, J. and GEIKER, M. R. 2016. The effect of interference of corrosion pits on the failure probability of a reinforced concrete beam. *Engineering Structures*, 114, 113-121.
- KIRCHNER, H., GRUVER, R., SWAIN, M. and GARVIE, R. 1981. Crack Branching in Transformation - Toughened Zirconia. *Journal of the American Ceramic Society*, 64, 529-533.
- KLEIN, L. A. 1999. Sensor and Data Fusion Concepts and Applications, Society of Photo-Optical Instrumentation Engineers (SPIE).
- KNOTKOVA, D., DEAN, S. W. and KREISLOVA, K. 2010. ISOCORRAG, International Atmospheric Exposure Program: Summary of Results: Developed by ISO/TC 156/WG 4, Atmospheric Corrosion Testing and Classification of Corrosivity of Atmosphere, ASTM International.
- KOBAYASHI, A., WADE, B., BRADLEY, W. and CHIU, S. 1974. Crack branching in homalite-100 sheets. *Engineering Fracture Mechanics*, 6, 81-92.
- KONDO, Y. 1989. Prediction of fatigue crack initiation life based on pit growth. *Corrosion*, 45, 7-11.
- KONDO, Y. and WEI, R. P. 1989. Approach on quantitative evaluation of corrosion fatigue crack initiation condition. Int'l. Conf. on Evaluation of Materials Performance in Severe Environments, Kobe, Japan. 135-142.

- KRAWCZYK, B., COOK, P., HOBBS, J. and ENGELBERG, D. L. 2017. Corrosion Behavior of Cold Rolled Type 316L Stainless Steel in HCl-Containing Environments. *Corrosion*, 73, 1346-1358.
- LAIRD, N. M. 1988. Missing data in longitudinal studies. *Statistics in medicine*, 7, 305-315.
- LANDOLT, D. 2007. Corrosion and surface chemistry of metals, CRC Press.
- LEON-GARCIA, A. 1994. *Probability and random processes for electrical engineering*, Pearson Education India.
- LI, D., DEOGUN, J., SPAULDING, W. and SHUART, B. 2004. Towards Missing Data Imputation: A Study of Fuzzy K-means Clustering Method. Berlin, Heidelberg. Springer Berlin Heidelberg, 573-579.
- LI, D., GU, H. and ZHANG, L. J. S. C. 2013. A hybrid genetic algorithm–fuzzy c-means approach for incomplete data clustering based on nearest-neighbor intervals. 17, 1787-1796.
- LI, L., LI, X., DONG, C. and HUANG, Y. J. E. A. 2009. Computational simulation of metastable pitting of stainless steel. 54, 6389-6395.
- LIANG, K.-Y. and ZEGER, S. L. 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13-22.
- LIAO, C.-M. 1998. Particle-induced pitting corrosion of aluminum alloys.
- LIAO, C.-M., OLIVE, J., GAO, M. and WEI, R. 1998. In-situ monitoring of pitting corrosion in aluminum alloy 2024. *Corrosion*, 54, 451-458.
- LIBICKI, E. and BEN-ZION, Y. 2005. Stochastic branching models of fault surfaces and estimated fractal dimensions. *pure and applied geophysics*, 162, 1077-1111.
- LIN, P., AUST, K., PALUMBO, G. and ERB, U. 1995. Influence of grain boundary character distribution on sensitization and intergranular corrosion of alloy 600. *Scripta Metallurgica et materialia*, 33.
- LING, Y., XU, X., SHEN, L. and LIU, J. 2008. Multi sensor data fusion method based on fuzzy neural network. 2008 6th IEEE International Conference on Industrial Informatics. 153-158.
- LITTLE, J., GOLDSTEIN, M., JONATHAN, P. and DEN HEIJER, K. 2004. Spatio temporal modelling of corrosion in an industrial furnace. *Applied Stochastic Models in Business and Industry*, 20, 219-238.
- LITTLE, R. J. and RUBIN, D. B. 2014. *Statistical analysis with missing data*, John Wiley & Sons.
- LIU, J. and YE, B. 2008. Fuzzy Fusion of Multi-sensor Data for Tool Wear Identifying. 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery. 570-573.

- LIU, K., CHEHADE, A. and SONG, C. 2015. Optimize the Signal Quality of the Composite Health Index Via Data Fusion for Degradation Modeling and Prognostic Analysis. *IEEE Transactions on Automation Science and Engineering*, PP, 1-11.
- LIU, K., GEBRAEEL, N. Z. and SHI, J. 2013. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. *IEEE Transactions on Automation Science and Engineering*, 10, 652-664.
- LIU, K. and HUANG, S. 2016. Integration of data fusion methodology and degradation modeling process to improve prognostics. *IEEE Transactions on Automation Science and Engineering*, 13, 344-354.
- LIU, X., GOPAL, V. and KALAGNANAM, J. 2018a. A spatio-temporal modeling framework for weather radar image data in tropical southeast Asia. *The Annals of Applied Statistics*, 12, 378-407.
- LIU, X., YEO, K. and KALAGNANAM, J. 2018b. A statistical modeling approach for spatio-temporal degradation data. *Journal of Quality Technology*, 50, 166-182.
- MA, B., SHUAI, J., WANG, J. and HAN, K. 2011. Analysis on the latest assessment criteria of ASME B31G-2009 for the remaining strength of corroded pipelines. *Journal of failure analysis and prevention*, 11, 666-671.
- MA, Y., LI, Y. and WANG, F. 2010. The atmospheric corrosion kinetics of low carbon steel in a tropical marine environment. *Corrosion Science*, 52, 1796-1800.
- MAES, M. A., FABER, M. H. and DANN, M. R. 2009. Hierarchical modeling of pipeline defect growth subject to ILI uncertainty. ASME 2009 28th International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers, 375-384.
- MARKER, D. A., JUDKINS, D. R. and WINGLEE, M. 2002. Large-scale imputation for complex surveys. *Survey nonresponse*, 329341.
- MARSEGUERRA, M. and ZIO, E. 1997. Modelling the transport of contaminants in groundwater as a branching stochastic process. *Annals of Nuclear Energy*, 24, 625-644.
- MCCUEN, R. H., ALBRECHT, P. and CHENG, J. 1992. A new approach to power-model regression of corrosion penetration data. *Corrosion Forms and Control for Infrastructure*. ASTM International.
- MCLACHLAN, G. and KRISHNAN, T. 2007. *The EM algorithm and extensions*, John Wiley & Sons.
- MEEKER, W. Q. and ESCOBAR, L. A. 1998. *Statistical methods for reliability data*, New York, John Willey & Sons, Inc.
- MING-YUAN, H. and HUTCHINSON, J. W. 1989. Crack deflection at an interface between dissimilar elastic materials. *International journal of solids and structures*, 25, 1053-1067.

- MITCHELL, H. B. 2007. *Multi-sensor data fusion: an introduction*, New York, Springer Science & Business Media.
- MORCILLO, M., CHICO, B., DÍAZ, I., CANO, H. and DE LA FUENTE, D. 2013. Atmospheric corrosion data of weathering steels. A review. *Corrosion Science*, 77, 6-24.
- MORCILLO, M., FELIU, S. and SIMANCAS, J. 1993. Deviation from bilogarithmic law for atmospheric corrosion of steel. *British Corrosion Journal*, 28, 50-52.
- MORRISON, T. and WORTHINGHAM, R. 1992. Reliability of high pressure line pipe under external corrosion. PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON OFFSHORE MECHANICS AND ARCTIC ENGINEERING. American Society of Mechanical Engineers, 401-401.
- MOSCHOPOULOS, P. G. 1985. The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics*, 37, 541-544.
- MURER, N. and BUCHHEIT, R. G. 2013. Stochastic modeling of pitting corrosion in aluminum alloys. *Corrosion Science*, 69, 139-148.
- MURPHY, R. R. 1998. Dempster-Shafer theory for sensor fusion in autonomous mobile robots. *IEEE Transactions on Robotics and Automation*, 14, 197-206.
- NAKAI, T., MATSUSHITA, H. and YAMAMOTO, N. 2004. Effect of pitting corrosion on local strength of hold frames of bulk carriers (2nd report)—lateral-distortional buckling and local face buckling. *Marine Structures*, 17, 612-641.
- NETTO, T. A., FERRAZ, U. S. and ESTEFEN, S. F. 2005. The effect of corrosion defects on the burst pressure of pipelines. *Journal of Constructional Steel Research*, 61, 1185-1204.
- NOOR, N., OZMAN, N. and YAHAYA, N. 2011. Deterministic prediction of corroding pipeline remaining strength in marine environment using DNV RP-F101 (Part A). J. Sustainability Sci. Manage, 6, 69-78.
- OLDFIELD, J. W. 1988. Electrochemical theory of galvanic corrosion. *Galvanic Corrosion*. ASTM International.
- OSSAI, C. I., BOSWELL, B. and DAVIES, I. 2016. Markov chain modelling for time evolution of internal pitting corrosion distribution of oil and gas pipelines. *Engineering Failure Analysis*, 60, 209-228.
- OTIENO, M., ALEXANDER, M. and BEUSHAUSEN, H. 2008. Corrosion propagation in cracked and uncracked concrete. *Concrete Repair, Rehabilitation and Retrofitting II: 2nd International Conference on Concrete Repair, Rehabilitation and Retrofitting, ICCRRR-2, 24-26 November 2008, Cape Town, South Africa.* CRC Press, 157.
- PAJARES, G. and DE LA CRUZ, J. M. 2004. A wavelet-based image fusion tutorial. *Pattern recognition*, 37, 1855-1872.

- PANCHENKO, Y. M. 2013. Kinetics gravimetric parameters corrosions process. *Corrosion Mater Prot*, 4, 1-7.
- PANCHENKO, Y. M. and MARSHAKOV, A. 2016. Long-term prediction of metal corrosion losses in atmosphere using a power-linear function. *Corrosion Science*, 109, 217-229.
- PANCHENKO, Y. M., MARSHAKOV, A., IGONIN, T., KOVTANYUK, V. and NIKOLAEVA, L. 2014. Long-term forecast of corrosion mass losses of technically important metals in various world regions using a power function. *Corrosion Science*, 88, 306-316.
- PANDEY, M., YUAN, X. and VAN NOORTWIJK, J. 2005. Gamma process model for reliability analysis and replacement of aging structural components. *Proceedings ICOSSAR*. Rome, Italy.
- PARK, C. and PADGETT, W. 2005. Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes. *Lifetime Data Analysis*, 11, 511-527.
- PARMIGIANI, J. and THOULESS, M. 2006. The roles of toughness and cohesive strength on crack deflection at interfaces. *Journal of the Mechanics and Physics of Solids*, 54, 266-287.
- PELCKMANS, K., DE BRABANTER, J., SUYKENS, J. A. K. and DE MOOR, B. 2005. Handling missing values in support vector machine classifiers. *Neural Networks*, 18, 684-692.
- PENG, W., LI, Y.-F., YANG, Y.-J., HUANG, H.-Z. and ZUO, M. J. 2014. Inverse Gaussian process models for degradation analysis: A Bayesian perspective. *Reliability Engineering & System Safety*, 130, 175-189.
- PENG, W., LI, Y.-F., YANG, Y.-J., ZHU, S.-P. and HUANG, H.-Z. 2016. Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas. *IEEE Transactions on Reliability*, 65, 624-639.
- PHILLIPS & JOHNSTON, I. P. T. ASTM A312 PIPE PRESSURE RATINGS [Online]. Available: [Accessed].
- PROVAN, J. and RODRIGUEZ III, E. 1989. Part I: Development of a Markov description of pitting corrosion. *Corrosion*, 45, 178-192.
- QIN, S. and CUI, W. 2003. Effect of corrosion models on the time-dependent reliability of steel plated elements. *Marine Structures*, 16, 15-34.
- QIU, Y., WU, J., HUANG, H., WU, H., LIU, J. and TIAN, J. 2005. Multi-sensor image data fusion based on pixel-level weights of wavelet and the PCA transform. *IEEE International Conference Mechatronics and Automation*. IEEE, 653-658.
- RABINOVITCH, A. and BAHAT, D. 2011. Differentiation of crack branching types in fractured glass. *EPL (Europhysics Letters)*, 95, 16003.

- RAHIMOV, I. 1995. *Random sums and branching stochastic processes*, Springer Science & Business Media.
- RAHMAN, G. and ISLAM, Z. 2011. A decision tree-based missing value imputation technique for data pre-processing. *Proceedings of the Ninth Australasian Data Mining Conference-Volume 121*. Australian Computer Society, Inc., 41-50.
- RAHMAN, M. G. and ISLAM, M. Z. 2014. FIMUS: A framework for imputing missing values using co-appearance, correlation and similarity analysis. *Knowledge-Based Systems*, 56, 311-327.
- RAMULU, M. and KOBAYASHI, A. 1985. Mechanics of crack curving and branching a dynamic fracture analysis. *Dynamic fracture*. Springer.
- RAMULU, M., KOBAYASHI, A. and KANG, B.-J. 1984. Dynamic crack branching—a photoelastic evaluation. *Fracture Mechanics: Fifteenth Symposium*. ASTM International.
- RAVI-CHANDAR, K. and KNAUSS, W. 1984. An experimental investigation into dynamic fracture: III. On steady-state crack propagation and crack branching. *International Journal of fracture*, 26, 141-154.
- REICH, B. J., EIDSVIK, J., GUINDANI, M., NAIL, A. J. and SCHMIDT, A. M. 2011. A class of covariate-dependent spatiotemporal covariance functions. *The annals of applied statistics*, 5, 2265.
- RIBEIRO, C. and WEBBER, N. 2004. Valuing path-dependent options in the variancegamma model by Monte Carlo with a gamma bridge. *Journal of Computational Finance*, 7, 81-100.
- RICE, J. A. 2006. Mathematical statistics and data analysis, Cengage Learning.
- ROBERGE, P. R. 2000. Handbook of corrosion engineering, McGraw-Hill.
- ROBINS, J. M., ROTNITZKY, A. and ZHAO, L. P. 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the american statistical association*, 90, 106-121.
- SAKAI, T., RAMULU, M., GHOSH, A. and BRADT, R. 1991. Cascadating fracture in a laminated tempered safety glass panel. *International journal of fracture*, 48, 49-69.
- SAWYER, S. 1976. Branching diffusion processes in population genetics. *Advances in Applied Probability*, 8, 659-689.
- SCHAFER, J. L. 1997. Analysis of incomplete multivariate data, CRC press.
- SCHAFER, J. L. and GRAHAM, J. W. 2002. Missing data: our view of the state of the art. *Psychological methods*, **7**, 147.
- SCHIEBER, S. J. 2005. A comparison of three alternative techniques for allocating unreported social secrity income on the survey thelow income aged and disabled. *American Statistical Association, Proceedings of the Section on Survey Research Methods.* 212-218.

- SCHNEIDER, T. 2001. Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. 14, 853-871.
- SCHWEITZER, P. A. 1983. Corrosion and corrosion protection handbook, Marcel Dekker. 483-484.
- SEEM, J. E. 2007. Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy and Buildings*, 39, 52-58.
- SHAFER, G. 1976. A mathematical theory of evidence, Princeton, Princeton University Press
- SHI, H., XU, L., SONG, C., ZHONG, F., DONG, H. and SU, R. 2012. Multi-satellite monitoring SST data fusion based on the adaptive threshold clustering algorithm. *Journal of Computers (Finland)*, 7, 2593-2598.
- SHREVE, S. E. 2004. *Stochastic calculus for finance II: Continuous-time models*, Springer Science & Business Media.
- SIGRIST, F., KÜNSCH, H. R. and STAHEL, W. A. 2015. Stochastic partial differential equation based modelling of large space–time data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77, 3-33.
- SILLS, R. and THOULESS, M. 2015. Cohesive-length scales for damage and toughening mechanisms. *International Journal of Solids and Structures*, 55, 32-43.
- SINGH, Y., ORFINO, F. P., DUTTA, M. and KJEANG, E. 2017. 3D failure analysis of pure mechanical and pure chemical degradation in fuel cell membranes. *Journal of The Electrochemical Society*, 164, F1331-F1341.
- SOARES, C. G., GARBATOV, Y., ZAYED, A. and WANG, G. 2009. Influence of environmental factors on corrosion of ship structures in marine atmosphere. *Corrosion Science*, 51, 2014-2026.
- STEINBERG, A. N., BOWMAN, C. L. and WHITE, F. E. 1999. Revisions to the JDL data fusion model. *Proceedings of SPIE*. 430-441.
- STEWART, M. G. and AL-HARTHY, A. 2008. Pitting corrosion and structural reliability of corroding RC structures: Experimental data and probabilistic analysis. *Reliability Engineering & System Safety*, 93, 373-382.
- STEWART, M. G. and MULLARD, J. A. 2007. Spatial time-dependent reliability analysis of corrosion damage and the timing of first repair for RC structures. *Engineering Structures*, 29, 1457-1464.
- STROUD, J. R., MÜLLER, P. and SANSÓ, B. 2001. Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 673-689.
- SUNDARAM, B. M. and TIPPUR, H. V. 2018. Dynamic fracture of soda-lime glass: a full-field optical investigation of crack initiation, propagation and branching. *Journal of the Mechanics and Physics of Solids*, 120, 132-153.

- SURESH, S., ZAMISKI, G. and RITCHIE, D. R. 1981. Oxide-induced crack closure: an explanation for near-threshold corrosion fatigue crack growth behavior. *Metallurgical and Materials Transactions A*, 12, 1435-1443.
- SUTER, T. and ALKIRE, R. C. 2001. Microelectrochemical studies of pit initiation at single inclusions in Al 2024-T3. *Journal of the Electrochemical Society*, 148, B36-B42.
- SYED, S. 2013. The effect of atmospheric pollution on materials damage. *Materials and Corrosion*, 64, 633-644.
- SZKLARSKA-SMIALOWSKA, Z. and ZS-SMIALOWSKA 2005. *Pitting and crevice corrosion*, NACE International Houston, TX.
- TEIXEIRA, A., SOARES, C. G., NETTO, T. and ESTEFEN, S. 2008. Reliability of pipelines with corrosion defects. *International Journal of Pressure Vessels and Piping*, 85, 228-237.
- THOMPSON, E. 2007. Behavior of proton exchange membrane fuel cells at sub-freezing temperatures. PhD diss., University of Rochester.
- TRZASKOMA, P. 1990. Pit morphology of aluminum alloy and silicon carbide/aluminum alloy metal matrix composites. *Corrosion*, 46, 402-409.
- TUKEY, J. W. 1975. Mathematics and the picturing of data. *Proceedings of the international congress of mathematicians*. 523-531.
- TURCOTTE, J., RIOUX, P. and LAVOIE, J.-A. 2016. Comparison corrosion mapping solutions using phased array, conventional UT and 3D scanners. *19th world conference on non-destructive testing*.
- TURNBULL, A., MCCARTNEY, L. and ZHOU, S. 2006a. A model to predict the evolution of pitting corrosion and the pit-to-crack transition incorporating statistically distributed input parameters. *Corrosion Science*, 48, 2084-2105.
- TURNBULL, A., MCCARTNEY, L. and ZHOU, S. 2006b. Modelling of the evolution of stress corrosion cracks from corrosion pits. *Scripta materialia*, 54, 575-578.
- TURNBULL, A. and ZHOU, S. 2004. Pit to crack transition in stress corrosion cracking of a steam turbine disc steel. *Corrosion Science*, 46, 1239-1264.
- VALOR, A., CALEYO, F., ALFONSO, L., RIVAS, D. and HALLEN, J. M. 2007. Stochastic modeling of pitting corrosion: A new model for initiation and growth of multiple corrosion pits. *Corrosion Science*, 49, 559-579.
- VALOR, A., CALEYO, F., HALLEN, J. M. and VELÁZQUEZ, J. C. 2013. Reliability assessment of buried pipelines based on different corrosion rate models. *Corrosion Science*, 66, 78-87.
- VAN BUUREN, S. 2012. Flexible imputation of missing data, CRC press.
- VAN NOORTWIJK, J. 2009. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94, 2-21.

- VAN NOORTWIJK, J. M., VAN DER WEIDE, J. A., KALLEN, M.-J. and PANDEY, M.
  D. 2007. Gamma processes and peaks-over-threshold distributions for timedependent reliability. *Reliability Engineering & System Safety*, 92, 1651-1658.
- VAN NOORTWIJK, J. M. and VAN GELDER, P. H. 1996. Optimal maintenance decisions for berm breakwaters. *Structural Safety*, 18, 293-309.
- VANAEI, H., ESLAMI, A. and EGBEWANDE, A. 2017. A review on pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models. *International Journal of Pressure Vessels and Piping*, 149, 43-54.
- VANOVERLOOP, D. 1990. An investigation of corrosion on integrated circuits via pressure-temperature-humidity-bias stressing. *IEEE Transactions on Reliability*, 39, 30-34.
- VATUTIN, V., TELEVINOVA, T. and CHISTIAKOV, V. 1985. Probability methods in the physical researches. *Science*, 207.
- VELÁZQUEZ, J. C., CALEYO, F., VALOR, A. and HALLEN, J. M. 2009. Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines. *Corrosion*, 65, 332-342.
- VUKMIROVIC, M., DIMITROV, N. and SIERADZKI, K. 2002. Dealloying and corrosion of Al alloy 2024 T 3. *Journal of The Electrochemical Society*, 149, B428-B439.
- WANG, C. and ELSAYED, E. A. 2020a. Stochastic Modeling of Corrosion Growth. *Reliability Engineering & System Safety*, 107120.
- WANG, C. and ELSAYED, E. A. 2020b. Stochastic Modeling of Degradation Branching Processes. *IISE Transactions*, 1-17.
- WANG, F. J. 1980. The convergence of a branching brownian motion used as a model describing the spread of an epidemic. *Journal of Applied Probability*, 17, 301-312.
- WANG, X. 2010. Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis*, 101, 340-351.
- WANG, X., LI, A., JIANG, Z. and FENG, H. 2006. Missing value estimation for DNA microarray gene expression data by Support Vector Regression imputation and orthogonal coding scheme. *BMC bioinformatics*, 7, 32.
- WANG, X. and XU, D. 2010. An inverse Gaussian process model for degradation data. *Technometrics*, 52, 188-197.
- WEI, M., CHEN, M., ZHOU, D. and WANG, W. 2011. Remaining useful life prediction using a stochastic filtering model with multi-sensor information fusion. 2011 Prognostics and System Health Managment Conference. IEEE, 1-6.
- WEI, R. P. 2001. A model for particle-induced pit growth in aluminum alloys. *Scripta Materialia*, 44, 2647-2652.
- WHITESIDE II, J. D. 2008. A practical application of Monte Carlo simulation in forecasting. *AACE International Transactions*, ES41.

- WIKLE, C. K. and CRESSIE, N. 1999. A dimension-reduced approach to space-time Kalman filtering. *Biometrika*, 86, 815-829.
- WOTHKE, W. 2000. Longitudinal and multigroup modeling with missing data.
- WRANGLÉN, G. 1985. An introduction to corrosion and protection of metals.
- WU, X. 2010. Branching processes with biological applications. PhD diss., Rice University.
- WU, X., KUMAR, V., QUINLAN, J. R., GHOSH, J., YANG, Q., MOTODA, H., MCLACHLAN, G. J., NG, A., LIU, B., PHILIP, S. Y. J. K. and SYSTEMS, I. 2008. Top 10 algorithms in data mining. 14, 1-37.
- WU, Y., XIE, L., WU, N. and LI, J. 2011. Time-dependent reliability model of components with strength degradation based-on gamma process. *Reliability, Maintainability* and Safety (ICRMS), 2011 9th International Conference on. IEEE, 363-368.
- YAGER, R. R. and LIU, L. 2008. *Classic works of the Dempster-Shafer theory of belief functions*, New York, Springer Berlin Heidelberg.
- YAN, H., LIU, K., ZHANG, X. and SHI, J. 2016. Multiple Sensor Data Fusion for Degradation Modeling and Prognostics Under Multiple Operational Conditions. *IEEE Transactions on Reliability*, 65, 1416-1426.
- YANG, S., WU, J. and PECHT, M. G. 2009. Reliability Assessment of Land Grid Array Sockets Subjected to Mixed Flowing Gas Environment. *IEEE Transactions on Reliability*, 58, 634-640.
- YAO, J., LUO, R. and YIN, X. 2014. Research on degradation behavior propagation model based on wear. *Reliability and Maintainability Symposium (RAMS)*, 2014 Annual. IEEE, 1-7.
- YATES, F. 1933. The analysis of replicated experiments when the field results are incomplete. *Empire Journal of Experimental Agriculture*, 1, 129-142.
- YE, Z.-S. and CHEN, N. 2014. The inverse Gaussian process as a degradation model. *Technometrics*, 56, 302-311.
- YU, L. and JATA, K. V. 2015. Review and study of physics driven pitting corrosion modeling in 2024-T3 aluminum alloys. 94372E-94372E-8.
- YUAN, X.-X., PANDEY, M. and BICKEL, G. 2008. A probabilistic model of wall thinning in CANDU feeders due to flow-accelerated corrosion. *Nuclear Engineering and Design*, 238, 16-24.
- YUAN, X. X. and PANDEY, M. D. 2009. A nonlinear mixed-effects model for degradation data obtained from in-service inspections. *Reliability Engineering & System Safety*, 94, 509-519.
- ZADEH, L. A. 1999. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 100, 9-34.
- ZAMAN, K. B. M. Q. and WEI, R. P. 1994. Probability approach for prediction of corrosion and corrosion fatigue life. *AIAA Journal*, 32, 2073-2079.

- ZHANG, S. 2014. Development of probabilistic corrosion growth models with applications in integrity management of pipelines. PhD diss., The University of Western Ontario.
- ZHAO, X., WANG, R., GU, H., SONG, G. and MO, Y. 2014. Innovative data fusion enabled structural health monitoring approach. *Mathematical Problems in Engineering*, 2014, 1-10.
- ZHOU, J., PAN, Z. and SUN, Q. 2010. Bivariate degradation modeling based on gamma process. *Proceedings of the World Congress on Engineering*.
- ZHOU, W., HONG, H. P. and ZHANG, S. 2012. Impact of dependent stochastic defect growth on system reliability of corroding pipelines. *International Journal of Pressure Vessels and Piping*, 96, 68-77.
- ZOIA, A., DUMONTEIL, E., MAZZOLO, A., DE MULATIER, C. and ROSSO, A. 2014. Clustering of branching Brownian motions in confined geometries. *Physical Review E*, 90, 042118.